Assessing schizophrenia in daily life : the experience sampling method

Citation for published version (APA):

Delespaul, P. A. E. G. (1995). *Assessing schizophrenia in daily life : the experience sampling method*. [Doctoral Thesis, Maastricht University]. Datawyse / Universitaire Pers Maastricht.
<https://doi.org/10.26481/dis.19950504pd>

Document status and date:

Published: 01/01/1995

DOI:

[10.26481/dis.19950504pd](https://doi.org/10.26481/dis.19950504pd)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Assessing Schizophrenia in Daily Life

— The Experience Sampling Method —

Philippe A.E.G. Delespaul

IPSER Series in Ecological Psychiatry

M.W. de Vries

Editor

* * *

Vol.1 Assessing Schizophrenia in Daily Life

— The Experience Sampling Method *Philippe A.E.G. Delespaul*

The International Institute for Psycho-social and Socio-ecological Research (IPSER)
Collaborating Center of the World Federation of Mental Health
Associated with the Rijksuniversiteit Limburg

Technical realisation: Cover painting: Frank Vreeswijk/ ΦΑΝΤΑΣΤΙΚΗ — Maastricht
Cover DTP: Unigraphic
DTP/Grafisch werk: PhD
Printing: Datawyse | Universitaire Pers Maastricht

Sponsored by: IPSER Foundation
Solvay-Duphar B.V. Weesp

Assessing Schizophrenia in Daily Life

— The Experience Sampling Method —

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Rijksuniversiteit Limburg te Maastricht,
op gezag van de Rector Magnificus, Prof. Mr. M.J. Cohen,
volgens het besluit van het College van Dekanen,
in het openbaar te verdedigen
op donderdag 4 mei 1995 om 16.00 uur.

door

Philippe Amand Etienne Ghislain Delespaul
geboren te Mol (België) in 1958

promotor: Prof. Dr. M.W. de Vries

beoordelingscommissie: Prof.dr. H.M. van Praag (voorzitter)
Prof.dr. R.J. van den Bosch (Rijksuniversiteit Groningen)
Prof.dr. M.A. van den Hout
Prof.dr. A.F. Lehman (University of Maryland, USA)
Prof.dr. H. Philipsen

IPSER Foundation
Postbus 214
6200 AE Maastricht
The Netherlands
tel -31-43-299770
fax -31-42-299708

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Delespaul, Philippe Amand Etienne Ghislain

Assessing schizophrenia in daily life: the Experience
Sampling Method / Philippe Amand Etienne Ghislain
Delespaul — Maastricht : Universitaire Pers Maastricht.
— Ill. — (IPSER series on ecological psychiatry; Vol.1)
Thesis Rijksuniversiteit Limburg Maastricht. — with ref.
— with Dutch summary.
ISBN 90-5278-182-6
Subject headings: schizophrenia; experience sampling
method

Table of Contents

Psychological Assessment, 2000, 10, 1-10. doi:10.1023/A:1009400020001

Preface	v
Assessing subjective experience in research and practice	vii
Personal account	vii
Study themes	viii
Searching for an assessment technique	ix
Scientific rigor and clinical decision making	xi
Conclusion	xii
Structure and contents	xiii
Acknowledgment	xv
 Part 1: Setting the Field	 1
Chapter 1: Schizophrenia: Current Status	3
Diagnosing schizophrenia	3
Treating schizophrenia	11
Chapter 2: Alternative Assessment of Psychopathology	21
Issues in the traditional psychiatric assessment	22
Towards ecological validity — from laboratory to real life	27
Functional analysis: a model to interpret daily life experience	35
Summary and discussion	47
 Part 2: The Experience Sampling Method	 49
Chapter 3: Self-Observed Mental State and “Time Sampling”	51
Introduction	51
Proposed changes in diagnostic practice	55
Data collection by intensive repeated observation	56
The ‘Experience Sampling Method’ (ESM)	61
Discussion: Understanding ESM data	73
Chapter 4: Technical Note: Devices and Sampling Procedures	83
Choice of the sampling device	83
Choice of the sampling process	87
Chapter 5: Reliability and validity	93
Traditional reliability and validity	94
Classification of ESM data	95
Reliability and validity in ESM studies	96
Summary and discussion	99
Chapter 6: Empirical Studies in ESM Reliability and Validity	101
A: Stability of ESM over time (test/retest reliability)	103
B: Definition of the “ESM reliability time-window”	107
C: Comparing ESM with diary assessments	110
D: ESM time and RIR event sampling in social interactions	116
E: Comparing ESM with retrospective assessments	133
F: Discussion: sampled experience in outcomes of clinical trials	140
Chapter 7: Analyzing ESM data	147
Formulation of research questions	148
Analyzing research questions	150
Interpreting results	159
Final comments	159
 Part 3: Studying Schizophrenic Symptoms and Daily Life Adaptation	 161
Chapter 8: Context and subjective experiences in schizophrenia	163
Variability in schizophrenic symptomatology	164
Study	166
Case presentations	170
Discussion	172

Chapter 9: Daily Life in Chronic Mental Illness	173
Method	173
Results	177
Discussion	197
Chapter 10: Hallucinations in Daily Life	201
Method	203
Results	206
Discussion	216
Chapter 11: Assessing "Quality of Life" in Schizophrenia	221
Defining "Quality of Life"	223
Method	225
Results	232
Individual cases and quality of life	236
Conclusion	237
Appendix: Preliminary reference information of ESM-QoL	238
 Part 4: Experience Sampling in Daily Clinical Practice	 239
Chapter 12: ESM in psychiatric rehabilitation	241
ESM in clinical practice	242
ESM and behavioral assessment	247
ESM software to assess daily life variability	249
Discussion	253
Chapter 13: Case Studies in "Personalized Rehabilitation"	255
Case 1: "against cultural expectations"	256
Case 2: "the challenge of parenthood"	265
Case 3: diagnostic fuzziness	277
Negotiated medical care	287
 Epilogue	 291
Symptoms of Daily Life Adaptation	293
Varying symptomatology	293
Daily life scene	294
Discussion	295
Towards an Innovation in Psychiatry	297
Ecological psychiatry	297
Negotiated mental care	298
 Appendices	 299
Appendix 1: ESM protocol	301
Appendix 2: ESM sampling device	305
Appendix 3: ESM sampling scheme	307
Appendix 4: ESM booklets	311
Appendix 5: Codebook information (including re-coding)	317
Appendix 6: ESM file structure	327
Appendix 7: ESM output for clinical practice: Personalized Rehabilitation	331
 Bibliography	 347
 Summaries	 375
Summary	377
Samenvatting	381
Curriculum Vitae	385

Preface

Assessing subjective experience in research and practice

PERSONAL ACCOUNT

My first contacts with psychiatry were not very successful. I was unprepared for the experience I had to face. I had great ambitions to help people but found myself lacking the necessary skills to meet this challenge. I did my internship on the chronic mental patient ward of the mental hospital in Maastricht, the same ward I now head 12 years later. During those months I submerged myself in the world of chronic mental patients, adjusting myself to their daily life routines. The most important objective was to synchronize the rhythm of my life to the slow circular experience of time of the schizophrenic patients on the ward, who seemed to have lost all purpose in life. This was a provocative experience, a confrontation with my own restlessness, my eagerness to learn, to understand and to help. It was provocative because it brought me in close contact with the experiential world of psychopathology, the suffering of patients, their families and the (nursing) staff who cared for them.

One of my first days in the mental hospital, I had a memorable experience. Once each week the patients on the ward cooked their own meal in order to develop and preserve skills necessary for community survival. They planned the menu together, composed a grocery list and matched this with a budget. Three patients — Jane, Don and George — had volunteered to do the shopping. I escorted them to the supermarket that was 1 kilometer away. I remember having a tremendous problem keeping them together. George and Don walked restlessly far ahead of us while I was talking to Jane, who enjoyed the personal attention. When we finally arrived at the supermarket, George and Don were already waiting for us. I went with them to the various sections to collect the items which were on the grocery list, leaving Jane by the entrance. After some minutes we heard a noise. Jane had tried to pick a tin can of vegetables from a pyramidal display. Not being able to reach the top one, she took one from the middle, thereby causing the whole structure to collapse. Faced with the chaos she had created, she panicked and started to throw the cans around. I was able to calm her down and diverted her outside the shop, wondering what the attendant would charge for the mess created. Unexpectedly, he charged nothing and helped me rather sympathetically to gather the groceries and to leave the shop quickly. Don and George were pacing up and down outside. They seemed untouched by what had happened. Jane, on the other hand, was becoming increasingly psychotic, screaming at invisible persons around us. On the way back I had to stop her from repeatedly throwing rocks at windows in which she saw a person, which scared her tremendously.

I still do not know if the nursing staff sent me on this mission on purpose to show the naive “would-be-psychologist” the real world. The emergence of a “psychotic breakdown” over a 10-minute interval, however, is not observed very often, even in the very ill subjects I was escorting that day. That episode made a deep impression on me.

This and a number of other experiences confronted me with the fact that an academic curriculum had left me unprepared to “understand” or “help” schizophrenic subjects.

STUDY THEMES

How can we understand what is happening in patients suffering from schizophrenia? This question haunted me for a long time and still intrigues me today. However, while *understanding schizophrenia* is an important first step in the therapeutic process, it offers no guarantee that this will facilitate change or alleviate suffering. Therefore, an equally important question is how we can *optimize treatment choices* using the knowledge available to us. Both themes are central to the questions of this study.

Understanding schizophrenic pathology

One can learn much about schizophrenia by studying major textbooks. Recent developments in the psychiatric assessment systems allow reliable case finding strategies. The resulting epidemiologic studies teach us that schizophrenia is one of the major illnesses, with a lifetime prevalence of 1.5% — almost 100 million people on earth. We also know that schizophrenia is, in part, genetically determined and can be characterized by neurotransmitter imbalances and rather aspecific structural brain abnormalities. Furthermore, psychotic symptoms respond to neuroleptics. None of these biologically related facts, however, are specific for the disease and environmental influences are considered important for both etiology and course. Neuropsychologic functioning promises insight into the interactions between brain and environment as it is reflected in behavior. Schizophrenic subjects and their relatives show gross deficits but a large number of neuropsychologic test conditions are not affected. Globally, schizophrenics have trouble working under conditions of distraction, but the ability to learn under focused conditions seems not to be disturbed. As a mediating factor, neuropsychologic functioning is not very successful and a number of problems remain unresolved. It is not clear if and how neuropsychologic deficits are related to structural and functional brain deficits. Also, we do not know how results of neuropsychologic tests are related to success in skills training situations and to general daily life functioning.

Although the literature on schizophrenia is so large that it is almost impossible to read it all or even to stay updated, our current etiologic knowledge is limited. A large number of empirical studies have demonstrated differences between schizophrenic and non-schizophrenic individuals, but no comprehensive theory allows a synthesis of the results gathered from the different fields of research. The “missing link” is undoubtedly an adequate description of the relation between the malfunctioning brain and the actual daily life adaptation. This link is largely unexplored. Part of the problem is caused by the heterogeneity of the disease and our inappropriate description of it, both from the phenomenologic and from the biologic points of view. This study will be centered around the *description of the dynamics of change in the subjective experience* of people suffering from schizophrenia. More adequate

descriptions of the daily adaptive process might enhance our knowledge of meaningful subgroupings in schizophrenia and allow a more successful integration of results from biologic and neuropsychologic research.

Helping schizophrenic subjects

In the middle of this century, schizophrenia was largely untreated and its prognosis was poor. The advent of the first neuroleptics in the Fifties created new hope. At the same time, studies demonstrated that spontaneous ameliorations were possible even after long periods of illness. Despite the original optimism, the prognosis for schizophrenia was still poor and as a consequence, a new therapeutic nihilism emerged. Again, schizophrenia became associated with a poor outcome and patients were largely un- or under-served.

In the past decade, however, treatment strategies for schizophrenia have again been proliferating. A new treatment technocracy has replaced the existing therapeutic nihilism. Taking care of schizophrenic patients is again rewarding and, luckily, the interest among mental health professionals is growing. Unfortunately, none of the many proposed interventions have been a panacea. While successes have been demonstrated in large numbers of patients, schizophrenia remains a serious disease with unknown etiology for which our treatment options are far from powerful. Schizophrenia is considered a heterogeneous condition and improvements in diagnostics leading to the definition of subgroups promises the tailoring of treatment interventions to the subjects' individual needs and ultimately enhanced outcomes. However, those subgroups are largely unknown and "trial and error" still seems to be the best treatment selection strategy.

The growing treatment technocracy provides us with tools to sustain the therapeutic relation with the schizophrenic patient and his family over prolonged periods of time. This unspecific effect of the growth of treatment alternatives is therapeutically important because it prevents early treatment termination by dissolving a therapeutic relation after one or a small number of failures. In the absence of therapeutically meaningful subgroupings, however, it is highly unsatisfactory to rely on a "trial and error" selection strategy over prolonged periods of time. In this study we will develop a *tool for treatment selection* that capitalizes on the assessment of the *patient's individual adaptation* to allow a *personalized rehabilitation* process.

SEARCHING FOR AN ASSESSMENT TECHNIQUE

Therapeutic relation (for assessment and treatment)

What possible developments in our knowledge about schizophrenic etiology can we expect in the near future? It is likely that researchers will find a genetic basis of schizophrenia in the next years. This may enhance our knowledge about neurotransmitter systems, facilitating a more targeted synthesis of pharmacotherapeutic compounds. It will hopefully also lead to a better understanding of the information

processing, allowing us to determine the conditions under which learning can be maximized. As in the case of Alzheimer's disease, it is highly unlikely that a genetic breakthrough would resolve all the cognitive, affective and behavioral deficits of schizophrenia. Neither a miracle drug nor enhanced neuropsychologic understanding will change the skills we need in the therapy of schizophrenia in the future. Our challenge then, will not be different from now. We will have to develop skills to communicate with, and relate to, subjects suffering from schizophrenia and to use this relationship meaningfully, both for assessment and for treatment.

The nosologic diagnostic systems currently available in psychiatry are valuable and allow the worldwide integration of scientific knowledge. For schizophrenia, however, the assessment strategies do not allow meaningful subgrouping, in spite of the generally assumed heterogeneity of this disease. Moreover, although we subscribe to the belief that the individual reality of daily life adaptational dynamics reflects psychopathology and should be central to our planning of interventions, we still select subjects for interventions in an indiscriminate way, largely ignoring this information source. Our diagnostic strategies in psychiatry are basically reductionistic approaches that miss the opportunity to radically individualize the assessment for interventions, while we acknowledge that the need for this is high in the population of (chronic) schizophrenic persons. This unfortunate situation is understandable, since no "tools" are currently available to allow us to facilitate this process. The actual communication with a schizophrenic patient, the personal meeting with his inner experience, and the graduation from the level of "knowing" and "understanding" to "skilled help", is still more an art than a technique. We all know clinicians in the settings where we work who have developed the therapeutic skills to relate to those severely invalidated individuals and who are able to obtain the answers that are needed to make the necessary decisions for planning interventions. Unfortunately, their skills are highly personal, and incorporate not only knowledge of the salient questions, but also and especially, the skills of making "moments of communication" from those encounters with schizophrenic subjects. The didactics of transferring those skills are not well-developed. For the novice in the field there seems to be no alternative for this long-lasting process of observing and assimilating, than to learn by "trial and error". This is not an ideal situation.

Structuring observations: bringing order to chaos

This study is an attempt to enhance our understanding of what the daily life reality is for subjects suffering from schizophrenia. Although the methodology involved requires a great deal of data, the understanding being pursued is not intended to be theoretical, but rather very practical. At present, neither our knowledge of schizophrenic etiology nor our understanding of how the suffering can be alleviated allows us to select interventions in a predictable way. Perhaps, our traditional assessment strategies collect data in the wrong place. A professionally conducted diagnostic process yields reliable summaries and predictions of the subject's psychopathology, but fails in validity when it comes to relating the diagnosis to etiology and treatment selection. I think that our current assessment strategies often generate "misplaced concreteness"

(A.N. Whitehead, 1925) in the content and process of the data they collect. This results in an alienating, chaotic experience for both the clinician and the patient. The expression of schizophrenic symptomatology is reflected in the subject's behavior, thoughts, perceptions and emotions. Moreover, it is in the interplay of the schizophrenic individual and his or her environment, that the deficits reveal themselves. Most of these psychopathologic expressions cannot be directly observed by the clinician, as the symptoms occur outside of the therapist's office most of the time — in the subject's living environment ("the other 23 hours"). Also, chances are that the modalities of the occurrence of symptoms are significantly different in the natural environment of the subject's daily life. Therefore we need to make assessments in that situation ("ecological validity"). Finally, most of those psychopathologic symptoms are private phenomena. Symptoms such as cognitions, perceptions and emotions are definitely unavailable for direct observation by the clinician. But while these crucial data are unavailable to the clinician, they can easily be collected from the subjects themselves. Why, then, not just ask them? The conventional answer is that by irrespectively trusting the verbalizations of the schizophrenics, we fear to create another chaos, capitalizing on unreliable data.

Subjective experience and the Experience Sampling Method

The Experience Sampling Method (ESM) was developed to study psychiatric symptomatology with ecologic validity, both to enhance our treatment potential in the individual case and to understand the relation between biologic or neuropsychologic deficits and handicaps in daily life. ESM is a random time sampling strategy of mental state and context, a repeated self-observation conducted and embedded within the natural flow of daily life experience. Subjects carry a digital watch that signals them randomly, 10 times each day, to fill in a pocket size questionnaire that assesses their behaviors, thoughts, perceptions and emotions at the moment of the beep. In addition, the context of the experience is described in terms of performed activities, persons involved, time and place.

Although the complexities of the schizophrenic experiences are real and meaningful, I do not believe that we need to cherish this complexity and to avoid each reduction as a kind of religion "mission" to safeguard the integrity of the individual experience. Equally, I do not believe that subjective reports are reliable or valid by definition. The authenticity of the subject's narrative ensures its relevance, but does not automatically generate the subjective experiences we need for studying etiology and customizing treatment selection. This necessitates an in-depth study of the genesis of subjective experiences and of the strategies for generating personal experiential reports. ESM partly reflects the acquired knowledge that has come from the study of this process.

SCIENTIFIC RIGOR AND CLINICAL DECISION MAKING

We should always make sure that the individual daily life adaptation as reflected in the subject's own cognitions, affects and behaviors are the basis of our assessment. It

is our assumption that schizophrenic subjects form a heterogeneous group of which, in fact, no relevant subgroups are known. Therefore, if we base our knowledge and understanding of an individual case on research data collected from other subjects, we risk using irrelevant information for the particular subject we want to treat. In contrast to the reductionistic approach of traditional psychiatric assessment and as a consequence of the assumption that schizophrenia is a heterogeneous disease we advocate the use of data collected from the subject him or herself and aggregating the repeated observations, generating rules that allow predictions of the subject's own behaviors, cognitions and emotions (individualized laws: an "idio-nomothetic" process). As a test of relevance, we think that the subjects should be able to recognize the observations that are used in the assessments as being an authentic description of their own life. The data should be sensitive to the real context of their daily life, the environmental niche in which symptomatology emerges ("ecologic validity"). Subjects should finally acknowledge that the conclusions derived from those observations apply to themselves ("member validity") — even if they hold for other subjects as well.

As a tool for clinical practice, ESM is a development towards a model of "negotiated medical care". It choreographs the patient-doctor encounter, allowing individualized assessments and customized interventions. It is not the doctor's expertise, but the elucidation and structuring of the normal daily life behavior and experience of schizophrenic patients and those who live with them, which form the basis of the negotiation towards change. The therapeutic task is "to bring order to the chaos" of the schizophrenic experience. Our task as clinicians is to structure and organize the observations in order to "understand" what is happening and what creates the suffering. To "understand" is being able to "predict". We should know how a specific schizophrenic subject typically thinks, feels and behaves in different situations or under different conditions. These kinds of predictions are a prerequisite for "helping". The professional's role is not to be an expert in the scientific knowledge about schizophrenia through creating meaning, inserting cultural concepts and psychologic interpretation, but, first and foremost, to help aggregate, summarize and negotiate the interpretation of the subject's own data. We do not aim for an maximal fit between data and summaries, but for a "good enough" optimal fit. I think that this analytic strategy, primarily inspired by clinical practice, is of added value to science itself.

CONCLUSION

This book is an attempt to make a contribution to the study of subjective experiences of schizophrenic subjects. It is my conviction that an adequate assessment of the daily life dynamic environmental adaptation in the experience of schizophrenic subjects is essential, both for fundamental research into the etiology and course of the disorder and for clinical work, to personalize treatment and rehabilitation programs.

From biologic and social psychiatry to neuropsychology, our scientific knowledge about schizophrenia is scattered. We assume that an undocumented heterogeneity causes this chaotic picture. Therefore, it is an important endeavor to make

improvements in the phenomenologic description of the schizophrenic illness process. In contrast to somatic illness, symptoms of psychiatric illnesses cannot be reduced to their biologic substrate. It is in the individual daily life adaptation — reflected in the subject's own cognitions, affects and behaviors — that a psychiatric illness reveals itself. It should therefore also be in those descriptions of experience and daily life adaptation that we should anchor our assessment strategies. This improved assessment of environmental adaptation is the target of this study. We hope that such an improvement can fertilize the multidisciplinary fundamental research into the heterogeneity of the etiology and course of schizophrenia. Figure A illustrates the multidisciplinary collaboration that integrates achievements on the biologic substrate (the structure and physiologic functioning of the brain), the brain related behavior (neuropsychologic functioning) and environmental adaptation (sociopsychologic functioning).

Figure A Multidisciplinary research in etiology and course of schizophrenia based on an improved assessment of brain, brain behavior and environmental adaptation.

Our aim is to generate reliable and valid reports of subjective experience over time. Because instruments that allow such a study are not currently available, this study is about method development. The *raison d'être* of this method development is the scientific description and understanding of the dynamics of environmental adaptation within the time frame of a day for human subjects in general, and for schizophrenics specifically. This understanding is only important when it offers us the scientific knowledge for treatment selection. At the group level this means that the method generates useful parameters which can be applied as outcome variables in clinical trials. At the subject level it allows optimization of interventions in individual cases.

Structure and contents

The work presented in this dissertation is the result of almost ten years of study. Over these years I was involved in a large number of research projects, both in the development of innovative assessment strategies in psychiatry and in the field of chronic mental health and schizophrenia. These collaborative endeavors have resulted in a

number of articles, peer reviewed for a Cambridge Press monograph edited by Marten deVries (deVries, 1992) as well as for national and international journals¹. Furthermore, our scientific work was presented at many major conferences worldwide and have resulted in lively discussions and collaborations with distinguished national and international colleagues and their institutes.

In Part I the frame of reference for this work is set by means of two introductory overview chapters of the research fields covered in this dissertation. Chapter 1 reviews the current state of our understanding of schizophrenia, from a diagnostic and a treatment point of view. Chapter 2 discusses assessment in psychiatry and gives a rationale for measuring experience in daily life in the study of psychopathology.

Part 2 describes the "Experience Sampling Method" (ESM) in detail. It contains a compilation of articles written and published collaboratively over the past years. Chapter 3 is a general introduction to the assessment of daily life experience and introduces ESM. An abbreviated version of the topics discussed in this chapter was published in the Dutch "Yearbook of Psychiatry and Psychotherapy" under the title "Self-Observed Mental State and Intensive Time Sampling". Chapter 4 was published in "The Experience of Psychopathology". It contains a highly technical description under the title "Devices and Sampling Procedures" which covers the implementation of different sampling strategies and their consequences for the quality of the collected data. Chapter 5 is written collaboratively with Prof. Arthur Stone of Stonybrook University in New York during his sabbatical stay at our department and covers the complex reliability and validity issues in ESM assessment strategies. Chapter 6 combines a number of previously unpublished empirical studies — although presented at international congresses — on ESM reliability and validity. The test-retest reliability study that is included, was published in *Schizophrenia Research*. The large subsection comparing event and time sampling data-collection strategies was written with Prof. Harry Reis of the University of Rochester during his Fulbright Scholarship at our section of Social Psychiatry and Psychiatric Epidemiology. Chapter 7 is the last chapter in Part 2 and was also published in "The Experience of Psychopathology". It was written some years ago in collaboration with Prof. Reed Larson of the University of Illinois at Urbana while he was a visiting professor at our department. It covers the analysis of ESM data, stressing the need for an accurate formulation of research questions and advising the use of unsophisticated statistical techniques, especially when the impact of sophisticated techniques on this data is unknown.

Part 3 covers the analyses of the schizophrenics' daily life using ESM. Chapter 8 was originally published in *Schizophrenia Bulletin* under the title "Social Context and Subjective Experience in Schizophrenia". The next three chapters expand our original work published in the *Journal of Nervous and Mental Disease* using new data on the daily lives of schizophrenic subjects. They were presented in congresses over the last year. In Chapter 9 I discuss the daily life adaptation of chronic mental patients, comparing schizophrenics and depressive subjects. It assesses the traditional questions in ESM research: how do subjects divide their time, how is their moment-to-moment mental state in daily life; and how is this experience related to contextual changes.

¹ Full references of the different publications are included when the Chapters are presented.

Part of this data was presented at the Bern International Schizophrenia Symposium in 1993. Chapter 10 assesses the issue of psychopathological symptoms, in cases of auditory and visual hallucinations, in the daily life of chronic schizophrenics. Preliminary results of this study were presented at the International Conference on Schizophrenia in Geneva in 1991 and published in *Schizophrenia Research*. Finally, Chapter 11 covers the development of aggregated ESM indices to assess "Quality of Life" and was presented at the WHO/WPA conference in Vienna.

In Part 4 I discuss the use of Experience Sampling Method in the treatment of individual subjects. Chapter 12 is a theoretical introduction of the use ESM data for clinical practice with an emphasis on behavioral assessment. These ideas were presented at the Behavioral Therapy Conference of Edinburgh and published in the *Tijdschrift voor Klinische Psychologie* and *Schizophrenia Research*. Finally, Chapter 13 contains three case studies. The first case was published originally in "The Experience of Psychopathology", the second case was presented at the APA Congress "Symposium on Daily Life and Psychopathology" in Washington and is in press in a book on psychiatric rehabilitation edited by Guido Pieters. The last case is new.

Acknowledgment

This study would not have been possible, without the collaboration of a large number of persons. First of all, I want to thank Professor Marten deVries. Marten. During all these years we have discussed psychiatry and developed the ESM in a working environment that was never without strains. In this stress we were able to keep a focus on the content of our research topic, to keep up with the norms of productivity but also to *remain creative scientifically* and to work in a broadly defined and heterogeneous field towards a clear well-defined goal: describing and understanding the processes underlying moment-to-moment adaptation in illness and well-being for fundamental research in psychopathology and for use in clinical practice.

I thank the direction of the Vijverdal psychiatric hospital and the RIAGG mental health center for the faith they have put in me, both as a clinician and as a researcher. I want, especially, to thank my colleagues of Woonunit A. Dear fiends, together we developed the therapeutic environment in which I could learn my expertise in schizophrenia. We spent a lot of time together, discussing therapeutic options, evaluating strategies, trying to understand what was going on. Without this collaboration this study would never have received its clinical depth.

I want to address a special thank to the academic staff and nursing crew of the services and wards who collaborated in the research projects that generated the data used in this dissertation: "Woonunit A", "Woonunit B", the "Psychiatric Foster Care Program", "Afdeling 06", the residential facility "Afdeling H" and the "Social Psychiatric Center" all from the PMS Vijverdal, and the "Social Psychiatric Service" of the RIAGG in Maastricht. I thank Steven Radstake and Egbert Van der Poel who co-authored in the clinical cases.

This dissertation grew in the scientific environment of the IPSER Foundation, the Academic Psychiatric Centre — Department of Psychiatry and Neuropsychology — Section of Social Psychiatry and Psychiatric Epidemiology — and the “Academisch Centrum voor Chroniciteitsonderzoek” (ACC). I thank the members of the research staff who came and went over the years. I will probably forget some of them: Jan Theunissen, Jan Scholte, Maarten Dormaar, Chantal Dijkman, Leny Meertens, Marcel Hilwig, Helga Nauta, Sandra Gransier, Anette Hofmann, Frank Kerpentier, Peter Habets, Truda Driesen, Carolien Dijkman, Maaïke Luxemburg, Frieda Van Goethem, Helene Visser and Armin Ashaad. Lex Volovics and Henk Bayer were partners in the animated discussion on statistical options for ESM and Bart Leunissen helped over the last months to conduct the complex multilevel random regression techniques continually rechecking on the data and rerun the analyses. I thank Marten deVries of course, but also Herro Kraan, Charles Kaplan, Nancy Nicolson and Joop deJong — senior staff members of our department — as well as the other Ph.D. students in the field of ESM research: Chantal Dijkman, Leny Meertens, Piet Portegijs, Marleen van Eck and Daniela Barge-Schaapveld for the stimulating scientific discussions. Other researchers at the University of Limburg used ESM in their studies: Rob Van Diest, Ton Schmidt, Marius Romme, Meindert Haveman and Richelle Lousberg. Finally, national universities (Rotterdam and Utrecht) and international collaborations (Liverpool, Milano, Rochester, Stonybrook, Nevada, Urbana, Chicago, Los Angeles, Newcastle) have stimulated us and helped formulate the ESM technique.

I further thank the different external consultants we have hosted at the IPSER institute and who have collaborated in and stimulated the current publications: Mike Csikszentmihalyi, Lyman Wynne, Ian Falloon, John Strauss, Michael Goldstein, Russel Hurlburt, Richard Lamb, Leona Bachrach, Helene Kraemer, Don Hedeker, Warren Ten Houten and Justin Kenardy. I specially thank Harry Reis, Arthur Stone and Reed Larson co-authored some of the chapters.

I thank Peter Emonds (of ECS International B.V.) for his help in the development of the ESM database and the feed-back software for clinicians. I also thank Kit Gerould who corrected the final draft versions of the text. Thanks to Frank Vreeswijk, Bertil Soogele and Eric Wijnands of the Creative Therapy of Vijverdal for the cover painting.

Finally, I thank the IPSER collaborators Anke, Tessa, Sara, Jan and Yvonne for the support in the most chaotic moments fighting against deadlines over the previous years. A special thanks to Harry Pomerantz. Thanks also to the secretaries of “Social Psychiatry”, Trees and Marie-José.

This research would not have been possible without the financial support of the Letten F. Saugstad Foundation, Solvay-Duphar International, The National Fund for Scientific Research of The Netherlands, The National Fund for Mental Health and the Field Intensive Research Stimulation Fund.

I finally want to thank my family. Dear father, you gave me an education fostering personal development and social concern. Without you, I would never have been able to study. No words can express my gratitude.

I especially want to thank my wife Frieda. Clinical work as well as research both take a lot of time and energy. This, of course, happened at the detriment of our time together. Now this work is finished, I do not believe and you know that things will calm down in the near future. You will know by now that deadlines and late nights of work are part of our live. I promise, however, to slow down a little in the near future and promise once again to enjoy more time together. I'm grateful that you shared this experience with me, stimulated me and sometimes brought me back to our daily lives. Martine and Karel, I was not always there. I hope we will find some more time in the future.

Part 1: Setting the Field

Chapter 1: Schizophrenia: Current Status

DIAGNOSING SCHIZOPHRENIA ²

Historical overview

What is schizophrenia and where does the concept come from? Some authors claim that schizophrenia is an old disease dating back for at least 4000 years. Others, on the contrary, say it is a disease of recent onset with a viral origin related to modern life (Howells, 1991). An impressive historical account of schizophrenia is presented by Celsus (25 B.C. - A.D. 50) in his “De Re Medica” cited by Howells (p. xi): “There is a third kind of madness, the longest of all so that it does not injure life itself, and which is accustomed to be a disease of a strong body. But there are two kinds of this; for some are deceived by false images, not in their judgment: such as poets report the raving Ajax or Orestes to have perceived. Some are disordered in their judgment”³. This grouping of both delusions and (visual) hallucinations (not imagination) in one disease entity separate from melancholia and phrenitis (meningitis) is impressive for someone living almost 2000 years ago and sounds almost modern. Of course, nowadays a diagnosis of schizophrenia based only on delusions and hallucinations would sound suspect because of the possible differential diagnostic aspects that have to be considered. As a confirmation of this suspicion Aretaeus (A.D. 135-200), cited by Roccatagliata (1991), adds the following to his description of psychotic symptoms: “this disease of young people may also be caused by lust with venereal excesses and excessive use of wine” (p. 17).

The case descriptions of madness from the Middle Ages usually lack the necessary information to allow diagnostic specificity. The concept of melancholia is central in the Medieval diagnostic vocabulary. Psychotic symptomatology is usually seen as a extreme form of derealization related to melancholia, while excited schizophrenics might have been classified under mania (Howells, 1991). The notion of hallucinations was intermingled with the concept of mystic experience and was not generally considered pathologic, although a demonic (pathologic) phenomenon was also acknowledged. It is unclear but possible that the depicted experiences were partly caused by sensory deprivation, starvation, the use of alcohol and drugs and self-castigation (Slade & Bentall, 1988). By our modern standards we would probably not consider Joan of Arc to be a schizophrenic (see Johnson, 1978).

From 1580 through 1850 ideas about madness evolved (Skultans, 1979; Sedler, 1991). The changes did not primarily affect the diagnostic system itself, but they influenced the thinking about etiology and even more importantly the concept of patient manage-

² Parts of this paragraph contains material from: Delespaul, Ph. A.E.G. (1982) *Literatuuronderzoek Extramuralizatie*. Vijverdal, Maastricht.

³ Celsus A.C. (1853) *The First Four Books of De Re Medica*. Transl. Steggall J., London Churchill.

ment. Mystical experiences were becoming less important. Romanticism reinstalled the role of emotions in action and — psychiatric symptomatology being reflected in the person's action — in psychopathology. Demonic attributions (possession) were left behind and moral considerations came to the fore, leading to Pinel's freeing of the patients of the Salpêtrière followed by Conolly in England and Tukey's establishment of what is now called the "moral treatment". Historical accounts of that period (1700-1850) are primarily sociologic critiques of the important changes in management of the care for the insane that took place over that period of time: the rise of the medical profession in the care of abnormal mental states and the birth of the mental hospital as an institutionalized form of seclusion of the mental patients from society (Foucault, 1972; Jones, 1972; Scull, 1977). The authors of these sociohistorical interpretations were all writing in a period of intense criticism of mental hospitals. They tend to evaluate these developments as highly negative and to romanticize about more remote times when patients were taken care of by their relatives or the community they lived in. It may come as a surprise, but the seclusion of mental patients was already regulated in Ancient Greece (Plato, *The Laws* XI, 934): "if anyone be insane, let him not be seen openly in the town, but let his kinfolk watch over him as best they may, under penalty of a fine". Mental hospitals were built in the 18th century as asylums, refuges for people in need of professional care and as an answer to the seclusion and exploitation that existed in local communities. Overcrowding and a failing therapeutic instrumentarium, more than the fact itself of bringing those suffering people together in one place, has led us to reject the mental hospital as a place to give appropriate care.

The second part of the 19th century gave us a large number of clinical descriptions of insanity (e.g. the work of Kahlbaum, Charcot, Korsakov) but no systematized diagnostic taxonomy. The systematization of schizophrenia as a disease entity comes from Novalis and Kraepelin. Novalis stressed the process-like development and context-relatedness of the disease in contrast with the primarily French tradition that describes only the symptomatic "status praesens". Kraepelin deserves credit for combining existing disease descriptions in one group, the so-called "dementia praecox". He repeatedly placed emphasis on the heterogeneity of the illness as it presents itself in clinical practice (Peters, 1991a). Eugen Bleuler later introduced the name "schizophrenia", herewith reactivating Hoffman's concept of split personality that formed no part of Kraepelin's definition. Bleuler's most important contribution to the study of schizophrenia is not the choice of the name. The concept of split personality as a characteristic of schizophrenia has been abandoned today. In contrast to Kraepelin, Bleuler (1911, 1920) stressed the non-deterministic course of the illness without, however, being too optimistic about the possible outcome⁴. The study was later continued over longer periods of time by his son Manfred (Bleuler, 1978). Eugen Bleuler is also remembered for structuring the study of symptoms and introducing the central pathogenic concepts of "loosening of (normal) associations" ("...daß das Denken unzusammenhängend wird.") and of "affective blunting" ("affektive Verblödung"), now respectively related to the positive and negative symptomatology clusters (Kay,

⁴ "Diese Krankheit kann in jedem Stadium stillstehen, und manche ihrer Symptome können sich sehr weit oder ganz zurückbilden; aber wenn sie weiter schreitet, führt sie zu einer Verblödung bestimmten Charakters." (Bleuler, 1920, p. 279)

1991). Schneider (1939, 1959) published his diagnostic criteria along the lines set by Bleuler and arranged them as “first” and “second” rank symptoms, stressing the precedence of clinical manifestations such as delusions, hallucinations and loss of ego-boundaries in making the diagnosis of schizophrenia. The work of Schneider helped to systematize diagnostic criteria and became the model for further diagnostic developments. Unfortunately, the objective definition of symptoms was also a clear departure from the German phenomenologic tradition. It resulted in a concretistic diagnostic practice to the detriment of a more appropriate globalized approach to diagnosis.

Modern nosology: the DSM - and ICD-systems

The United States-United Kingdom Cross-National Project of 1965-1970⁵ showed that the American psychiatrists applied the diagnosis of schizophrenia more often than their British colleagues. The greatest disagreement was over patients whose main complaints were motoric retardation, depressed mood and anxiety, as well as mania. Those were almost invariably diagnosed as schizophrenic in New York (Peters, 1991b). The International Pilot Study on Schizophrenia of the World Health Organization (data collection in 1966: WHO, 1973) was set up to define a set of universally accepted behavioral referents that would form the basis for an international diagnostic system. Since the publication of this study much has been achieved in the development of case detection instruments that can reliably be applied cross-culturally in research and clinical practice. Carpenter et al. (1974) defined criteria using discriminant functions that minimize the existing confusion in the core symptoms needed for the diagnosis of schizophrenia. Feighner et al. (1972) and Spitzer et al. (1978) have added information about the illness course. Information both about present symptoms and course forms the basis of our current diagnostic systems, the DSM IV (APA, 1994) and ICD-10 (WHO, 1992).

Modern nosologic systems converge. Schizophrenic diagnosis is actually made using a list of symptoms including signs such as hallucinations, delusions, thought disturbances (incoherence — loosening of associations), inappropriate affect and inappropriate behavior (catatonic). Symptoms of the acute phase should last for at least 1 month and the diagnoses of mood disorders, substance abuse or organic psychosis should be ruled out. Generally accepted subtypes are paranoia (primarily delusions and hallucinations), hebephrenia (primarily blunted affect, thought disorders) and catatonia (bizarre and inappropriate behavior). Things get more confused when we try to set criteria for other subtypes (“undifferentiated”, “residual”, “simple”, “other” and “unspecified”) and other related disorders (schizophreniform disorder, brief reactive psychosis, schizoaffective disorder and delusional disorders) (McGlashan & Fenton, 1991). Confusion arises because no comprehensive list of well-defined symptoms is used. Furthermore, the time frame over which the symptoms must be active is not well-defined and is applied confusingly. Terms such as “life-long development”, “acute”, “prodromal” or “residual” phase are used inconsistently. Moreover, the phase descriptions such as those needed for the diagnosis of schizophrenia rely on the

⁵ Professional Staff of the US-UK Cross-National Project: The Diagnosis and Psychopathology of Schizophrenia in New York and London. *Schizophrenia Bulletin*, 11: 80-102, 1974.

erroneous assumption that symptoms are invariably active over a specific period of time and that those periods can be defined precisely down to a day.

Although a wide range of alternative assessment criteria for schizophrenia were considered by the Task Force for the DSM IV (APA, 1991; Flaum & Andreasen, 1991; Andreasen & Flaum, 1991), the criteria that were finally selected after the field trials for the draft version of DSM IV still have the same structure as in the previous versions: a list of characteristic symptoms for the active phase; social and occupational dysfunction compared to previous levels of expectation; overall duration of at least 6 months; exclusion of depressive syndromes; and exclusion of organic mental syndromes (APA, 1993; 1994, see Table 1.1).

Diagnostic criteria for schizophrenia:

A Characteristic symptoms: Two (or more) of the following, each present for a significant period of time during a 1- month period (or less if successfully treated):

- (1) delusions;
- (2) hallucinations;
- (3) disorganized speech (e.g. frequent derailment or incoherence);
- (4) grossly disorganized or catatonic behavior;
- (5) negative symptoms, i.e., affective flattening, alogia, or avolition.

[Note: only one Criterion A symptom is required if delusions are bizarre or hallucinations consist of a voice keeping up a running commentary on the person's behavior or thoughts, or two or more voices conversing with each other].

B Social/occupational dysfunction: for a significant portion of the time since the onset of the disturbance, one or more major areas of functioning such as work, interpersonal relations or self-care are markedly below the level achieved prior to the onset (or when the onset is in childhood or adolescence, failure to achieve expected levels of interpersonal, academic or occupational achievement).

C Duration: Continuous signs of the disturbance persist for at least 6 months. This period must include at least 1 month of symptoms (or less if successfully treated) that meet criterion A (i.e., active-phase symptoms) and may include periods of prodromal or residual symptoms. During these prodromal or residual periods, the signs of the disturbance may be manifested by only negative symptoms or two or more symptoms listed in criterion A present in an attenuated form (e.g. odd beliefs, unusual perceptual experiences).

D Schizoaffective and mood disorder exclusion: Schizoaffective and mood disorder with psychotic features have been ruled out because either (1) no major depressive, manic or mixed episodes have occurred concurrently with the active phase of symptoms; or (2) if mood episodes have occurred during active phase symptoms, their total duration has been brief relative to the duration of the active and residual periods.

E Substance/general medical condition exclusion: The disturbance is not due to the physiological effect of a substance (drug abuse, medication) or a general medical condition.

F Relationship to a pervasive developmental disorder: If there is a history of Autistic Disorder or another Pervasive Developmental Disorder, the additional diagnosis of Schizophrenia is made only if prominent delusions or hallucinations are also present for at least a month (or less if successfully treated).

Classification of longitudinal course (can be applied only after at least 1 year has elapsed since the initial onset of active-phase symptoms):

- *Episodic with interepisode residual symptoms* (episodes are defined by the reemergence of prominent psychotic symptoms); also specify if: *with prominent negative symptoms*;
- *Episodic with no interepisode residual symptoms*;
- *Continuous* (prominent psychotic symptoms are present throughout the period of observation); also specify if: *with prominent negative symptoms*;
- *Single episode in partial remission*; specify if: *with prominent negative symptoms*;
- *Single episode in full remission*;
- *Other or unspecified pattern*.

Table 1.1 DSM-IV criteria for schizophrenia (APA, 1994).

The DSM-IV criteria are more consistent than DSM-III-R. Important differences between DSM-III-R and DSM-IV are in the definition of the active phase. The symptom list is more homogeneous now. Furthermore, the duration of active symptoms is changed from 1 week to 1 month. Aspects that still need improvement are the definition of negative symptoms as well as symptom criteria for the prodromal and residual phases. Also, the time criteria for the active phase — significant period of 1 month or less if successfully treated — is understandable, but still confusing. The modification of the symptom criterion of the active phase (A) in the note as well as the redefinition of the active phase as part of the duration criterion in the part on exclusion of depression (D) are also problematic.

Biologic subtypes

As many authors have ascertained, schizophrenia is a heterogeneous disease, in phenomenology (Bleuler, 1911; Arieti, 1974; Strauss & Carpenter, 1981) but also in etiology (Carpenter et al., 1987; Johnstone, 1991) and course (Bleuler, 1972; Ciompi & Müller, 1976; Miller & Cohen, 1987). Thus, the absence of a valid subtyping scheme may impede our efforts to identify etiologic factors (Walker, 1987). Different classification schemes have been applied based on symptomatology (paranoid/non-paranoid) and course (chronic/acute and process/reactive). The most recent scheme is the positive/negative distinction, a terminology that originated at the end of the previous century and is attributed to the British neurologist Hughlings-Jackson. It was introduced in modern psychiatry through the work of Strauss et al. (1974). The concept was standardized and studied by Crow (1980a,b) in Great Britain and by Andreasen (Andreasen & Olson, 1982; Andreasen et al., 1982) and Kay (Kay, 1991) in the United States. It was assumed that positive, or productive, symptoms prevail in the acute stage and represent hyperdopaminergic phenomena, which responds to neuroleptics and leads to a positive prognosis, while negative symptoms predominate in chronic stages and signify structural brain abnormalities. Prognosis is poor. The positive-negative typology was promising because it was a phenomenologic subdivision that assumed a clear etiologic biological hypothesis and prognostic prediction. In a monograph, Kay (1991) presented data that challenge this original typologic conceptualization. He found an orthogonal factor structure of positive and negative symptom clusters — therefore independent and not mutually exclusive as in a typology. Furthermore, longitudinal studies show that the positive symptom cluster does not predict better prognosis (Pogue-Geile & Harrow, 1984, 1985, 1987; Kay & Lindenmayer, 1987; Kay & Murrill, 1990; Fenton & McGlashan, 1991a,b). Lastly, the underlying biologic etiology was not confirmed (Andreasen et al., 1990).

While the dualistic typology with its biological etiologic hypothesis might seem a naive reduction, unable to capture the complexities of schizophrenia, the behavioral aspect of the positive-negative distinction has been repeatedly replicated and seems phenomenologically valid. The proposed procedure of optimizing the symptomatic assessment and reorganizing the symptoms dimensionally in the research of etiology is still considered a fruitful strategy. Alternatives such as the pyramid model (Kay & Sevy, 1990) are interesting work hypotheses for such an approach. The current

symptom list, however, might be improved with assessments of brain functioning (Gur & Pearlson, 1993; Lieberman & Koreen, 1993), neuropsychologic (cognitive) functioning (Frith & Done, 1989; Frith, 1992; Braff, 1993) or assessments of global adaptation in daily life (deVries, 1992). While measurement techniques are now available, the construction of reliable parameters and the testing of the validity of these data is currently under study. The dynamic or process of adaptation, measured at different levels, might be the most productive.

Saugstad (1989) has pioneered a model that allows us to understand the mutual influence of constituent biologic, demographic and social variables in the emergence of schizophrenia. Her arguments are based on epidemiological and demographic findings and results in a neurodevelopmental theory that predicts differences between schizophrenia and manic depression using maturation rate differences over time within this century but also, for instance, in subpopulations defined on social class. This theory allows us to bridge the gap between nature and nurture etiological arguments that haunts our understanding of major mental illnesses (see Plomin, 1994).

The “vulnerability-stress” model

The most prominent current theory of schizophrenia — a working model is a better description — is the so-called “vulnerability-stress” theory. It is rooted in two basic observations related to schizophrenic etiology. Firstly, whatever the actual working mechanisms are, schizophrenia runs in families and responds to psychotropic medication. Because the prevalence of schizophrenia diminishes when genetic links are looser — from monozygotic twins (35-50%), to children of two (39%) or one schizophrenic parent (10%), to the unrelated members of the general population (1%) (Torrey, 1983) — it is assumed that schizophrenia has an important genetic root (Smeraldi, 1988; Kendler & Diehl, 1993; Reveley, 1994). Interestingly, this higher prevalence is also found in even higher frequencies for other diagnoses such as depression. Other, non-genetic biologic explanations have also been put forward, of which most notably the early viral infection hypothesis (Torrey, 1988; Kirch, 1993). A second root for schizophrenic etiology is the observation that when genetic influences are controlled for, the prevalence of schizophrenia is clearly higher in children raised in the same environment (Goldstein, 1987). This strengthens the hypothesis of a combined biologic/environmental influence on schizophrenic prevalence.

The “vulnerability-stress” theory states that certain individuals are predisposed to develop schizophrenia and that serious environmental factors influence the expression of that predisposition (Zubin & Spring, 1977). Vulnerability is defined primarily as biologic/genetic, but some authors have also included other influences and build more complex models (Nuechterlein & Dawson, 1984; Nuechterlein, 1987). The necessary complexity of models that explain stress influences on schizophrenic symptomatology and more generally the mechanisms that control central nervous system functions have been illustrated by Falloon & Fadden (1993). While simple models have appealing explanatory power, they probably fall short in explaining the real life variability in vulnerability and stress.

Vulnerability in a more general form is a relatively stable characteristic of the individual. Depending on the level at which the "vulnerability-stress" theory is applied, it explains lifelong emergence of schizophrenia or the occurrence of psychotic relapses. The definition of "vulnerability" changes, then, from a restriction to genetic determinism to the inclusion of slowly changing influences such as a growing strength through learned coping strategies, higher invalidation through repeated relapses and even the long-lasting protection of depot neuroleptics. Recent developments in neurology have demonstrated how maturational changes in the central nervous system can occur in later life (Birchwood et al., 1988). PET scan and cerebral blood flow studies have shown localized stimulation of specific brain regions when subjects are involved in different cognitive tasks (Nuechterlein, 1983; Buchsbaum & Haler, 1987). It can be assumed that future research might demonstrate structural changes caused by environmental (psychologic) influences. If this appears to be true the traditional difference between psychologic and biologic structures becomes less important. And, both the sociopsychological study as well as the biopsychologic study of human adaptation might become congruent and equally legitimate approaches.

Unfortunately, we still have a long way to go. Reviews of research literature on the relationship between psychosocial stress and schizophrenic symptomatology have generally concluded that the relationship is weak (Rabkin, 1980; Tennant, 1985; Norman & Malla, 1991). On the other hand, the idea that neurodegenerative structural brain changes occur in the life course of schizophrenics has been challenged by Crow (Crow et al., 1984). They assumed that if brain tissue degenerates and neurons die off, this is marked by gliosis. Post-mortem studies of schizophrenic brains show that while some brains show evidence of gliosis, most do not (Roberts & Bruton, 1990). Furthermore, brains with enlarged ventricles — a traditional indicator of schizophrenic brain abnormality — show no more sign of gliosis (Bruton et al., 1990). Based on this and other research, it is now generally accepted that the brain abnormalities are fixed at birth — genetically or pre-natally — and pathologic expressions become manifest clinically during the normal maturation of the brain (Breslin & Weinberger, 1990).

Discussion

The search for more reliable nosologic assessment instruments of schizophrenia is not a success story (Pope & Johnson, 1987). In an attempt to enhance reliability⁶ we have sacrificed clinical validity (Klerman et al., 1984). Further developments in our diagnostic systems have resulted in higher diagnostic purity but lost most practical meaning. Our current knowledge of schizophrenia has not gained from this diagnostic sophistication. Some authors even argue that our current diagnostic systems are responsible for the standstill in scientific improvements in the field (Van Praag, 1992; Bentall, 1990, 1992). The diagnosis of schizophrenia assumes homogeneity, but the observed heterogeneity in phenomenology and possible etiology and course means that more meaningful subgroupings may be hidden (Andreasen & Carpenter, 1993). The diagnosis of schizophrenia is primarily determined through the observation of the behavior of schizophrenic subjects. The underlying core disease entity or biologic

⁶ Andreasen & Flaum (1991) reviewed of DSM III reliability and schizophrenic symptomatology.

vulnerability of schizophrenia is actually still a working hypothesis with no hard etiologic or causal proof. Evidence for the existence of an underlying biologic ground, however strong, is circumstantial. The assessment of the (biologic) vulnerability of schizophrenics is based on indirect observations: ratings of behavioral expressions. Those human behaviors — normal as well as pathologic — are rarely reflexive or mechanistic expressions of a biological substructure. On the contrary, behaviors are (mal)adaptive actions of individuals interacting with their environments and pursuing personal goals. This assertion holds for all kinds of behaviors, the primary (illness-related) symptoms, as well as secondary symptoms (the disabilities caused by the illness). We are not able to measure this failing adaptational process that characterizes the illness directly in the biologic substrate. Our current diagnostic system may be the cause of this and we, therefore, will have to develop better observations of schizophrenic behavior. Before this can be realized, the only gain we get from models such as “vulnerability-stress”, is the conclusion that relapse-anticipating events must have disturbed the “vulnerability-stress” balance for that subject. No individual predictions on vulnerability or stressful events are possible. Nor is, the replication of relapse in similarly stressful situation certain for a specific individual. Different contextual circumstances or changed emotional reactions to events can lead to another result. The “vulnerability-stress” model is therefore, of poor clinical usefulness. The only actual application is as a didactic tool in educational programs.

Even without reference to an underlying behavioral substrate, the study of daily life adaptive behaviors is meaningful because those behaviors are expressions of what is experienced by the subjects and their relatives. They are also the expression of what can actually be realized by the subjects as a result or despite of their illness. The assessment of behaviors, cognitions and emotions is, therefore, not only a path toward determination of the illness, but probably, even more so, worthwhile in itself. However, both have to be improved: the assessment of the biological illness as well as the whole field of behavioral, cognitive and emotional adaptation.

We do not know if the subjects we call schizophrenics carry schizophrenic vulnerability. It is of no use to add more restricted criteria to our diagnostic instruments as long as we do not really know how to make our choices. A good research strategy is to start with a broad, but still specific, range of criteria for including subjects in the research group. From that group we should collect in depth descriptions of biologic markers as well as behavior using multidimensional systems (which needs further improvement!). Those criteria should not be restricted only to core symptomatology (we do not know if the actual itemset is useful) but also be based on additional behavioral, experiential and contextual information.

There is a positive aspect in this rather aspecific research strategy. It allows us to develop useful interventions. If a biologic marker were to be found, and if this marker could be influenced by a biologic intervention, the basic rules of human behavior would still apply. The biologic changes will facilitate behavior, but relational and motivational aspects will regulate its actual expression. If we do not find such a biologic marker, or if we are not able to influence it satisfactorily, we must keep in mind that behavioral changes are still possible. Similarly, as it is possible for a one-legged person to learn to jump even higher than I could, it is equally possible to

change the illness-related behaviors (primary as well as secondary symptoms) of patients who suffer from a biologically determined mental disease.

In summary, we need a new kind of typology that reflects a biologic reality and constitutes a better description of the actual experience of our subjects with sensitivity for the daily life interactions with the environment (deVries, 1994).

TREATING SCHIZOPHRENIA

Schizophrenia clearly is a major disorder leading to suffering both for the patients and their relatives. What can we do alleviate the illness? The history of psychiatric interventions is interwoven with the history of treatment of schizophrenia. Many new intervention techniques for mental health were originally intended for the treatment of schizophrenia. Intriguingly, a number of those approaches have been abandoned in the care of schizophrenic patients and are used now for other psychopathologic states. Unfortunately, no single strategy is accepted as the treatment of choice for schizophrenia. Neuroleptics are a necessary, but not sufficient, option.

The major psychiatric fields have developed theories which offer a rationale for the schizophrenic symptom formation and therapeutic change. With no ambition to be exhaustive, we will discuss psychoanalysis, behavior therapy and biologic psychiatry. The first two are psychotherapeutic disciplines that do not aim at changing the biologic substrate. Those theories are attempts to understand symptomatic behavior and to make assumptions about how those behaviors can be changed. We will discuss how psychoanalysis and behavior therapy have influenced the field of interventions in schizophrenia and illustrate the basics of clinical practice in the field. In the final section we will discuss the field of biologic psychiatry, which has introduced important treatments to supplement the more relation-based approaches.

Psychoanalysis

Psychoanalysis has always had an ambivalent relationship with schizophrenia. The first interpretations were formulated by Freud (1911) in the Schreber case. Schizophrenia was assumed to be caused by conflicts or intrapsychic trauma, especially of repressed homosexuality. Symptoms were ascribed to the withdrawal of libidinal cathexis from the outer world and from the representation of the object. Therapy, therefore, is the recathexis of the objects. Later Freud (1917) described the thought processes of schizophrenia to be dominated by primary process thinking. In this formulation unrestricted communication is central. After Freud's formulation of personality structures (Freud, 1923) the discussion on the differentiation between neuroses and psychosis was reactivated. The conflicts in schizophrenic behavior are not restricted to ego-id conflicts, but also reflect the more fundamental loss of ego/external world boundaries (Freud, 1924a,b).

Fenichel (1945) reiterated Freud's theories. Summarizing, Fenichel argued that "the following formulae mean only the same thing, only varying in point of view: the schizophrenic has regressed to narcissism; the schizophrenic has lost his objects; the

schizophrenic has parted with reality; the schizophrenic ego has broken down" (p. 415). Sullivan (1929, 1962) linked the emergence of psychopathology to the human life cycle that evolves in interaction with the environment. He stressed the role of anxiety in the fragmentation of schizophrenic subjects. Hartmann (1953) later demonstrated that the role of aggression is as important as libidinal withdrawal in the schizophrenic process. He argues that the ego-apparatus might be defective, lacking the necessary energy to neutralize impulses leading to panic and an uncontrolled flow of aggressive energy. Klein (1975) relates this to oral experiences of gratification and frustration. Anxiety in that conflict is very profound because it is related to survival. Ordinarily the care (breast-feeding) of the mother helps to establish the objects, reducing the need for annihilation. But in the absence of such a mother the ego is constantly beset with anxiety. The child feels that "he cannot rely on an external and internal good object, nor can he rely on his own self". In its attempt to disperse fear ("paranoid-schizoid anxiety") the ego simply disintegrates. Fairbairn (1952) accepts the central role of aggression but chooses the resulting feelings of futility as the primary focus of therapy.

Freud formulated psychosis as "narcissistic neuroses" in contrast to "transference neuroses". He acknowledged that the psychoanalytic technique, in which object formation is a prerequisite to realize change in the transference situation, was useless in the cases of psychoses. The pre-oedipal characteristics of the psychotic regression do not allow the (re)constructive use of ego-functions in the interpersonal therapeutic relation. Later this extremely pessimistic position was abandoned because the regression to narcissism is not total in schizophrenia and the therapist can still utilize the residues of reality relationships as well as the patient's spontaneous attempts at recovery in the analysis (Fenichel, 1945). Federn (1952) stresses the quality of transference the schizophrenic patients are able to achieve but also acknowledges that this transference is quite unstable. To establish and maintain a positive transference Federn accepts "acting", even physical gratification of the patient's oral needs. Sullivan (1956, 1962) was active and confrontational, he told his patients his feelings about them and gave them corrective feedback about their cognitive distortions of reality. More generally speaking, "Psychoanalytically oriented intensive psychotherapy" (Fromm-Reichman, 1950; Searles, 1965; Kohut, 1971) is not classic psychoanalytic therapy but goes beyond the prescribed rules by adopting a face-to-face set-up and sometimes allows physical contact. Pao (1979) stressed that "we must be alert to the fact that (the schizophrenic's) capacity to communicate varies from moment to moment and we must continually monitor that capacity to make our communication meaningful to him" (p. 207). Furthermore, in line with the tradition developed at Chestnut Lodge, Pao acknowledges the vital importance of the total caregiving environment as contributing to a part of the patient's sense of being cared for in the therapeutic relationship — a "good enough holding environment" (Feinsilver, 1986).

The therapeutic success of psychoanalysis has been criticized for a long time (e.g. Eysenck, 1960) and doubts are formulated about the theory and technology of pre-packaged interpretations (Timpanaro, 1976; Castoriadis-Aulagnier, 1975; Castel, 1973). This criticism applies to the technique, content and interpretative structure of psychoanalysis, and also, more recently, to the scientific merits of Freud's writings.

All this criticism does not diminish the impact of psychoanalysis on modern thought (Mujeeb-ur-Rahman, 1977; Fisher & Greenberg, 1978). Moreover, the psychoanalytic literature is the richest source of clinical expertise in a skill that has been lost in the technocratic therapeutic world: the communicational relation with schizophrenia (see Searles, 1965; Selzer et al., 1989, 1990; Robbins, 1993; Steiner, 1993). It shows that the seemingly erratic behavior of schizophrenics can be understood as meaningful and that schizophrenic communication can be trusted as real and applied constructively.

Behavioral therapy

The basic rule of behavioral therapy is that abnormal behavior is not different from normal behavior in its development, maintenance, or in the procedures to change it (Ullman & Krasner, 1975, p. 32). Two principles — derived from laboratory findings in experimental (animal) learning theory — explain the occurrence and changeability of behaviors. In behavioral therapy, these experimental principles are applied in clinical practice. The first principle — classical learning — applies to the learning of new meanings for stimuli (S-S- relations). This model has become specially related to the therapeutic change of emotions. The second principle is operant learning and describes the learning of behaviors, how someone can change one situation into another (S-R-C - transitions). Operant learning is primarily applied in skills training.

The universal changeability of behaviors — while criticized from an ethical point of view — was overoptimistic. The promise was that every behavior could be learned or unlearned by changing the context in which the behaviors originally occur. This was not true in practice. Contemporary animal learning research demonstrates the limitations of the changeability of behavior (Dickinson, 1980). Not all stimulus meanings nor all stimulus - response associations can be changed equally. There is further evidence that learning principles cannot be transposed irrespectively from the animal laboratory to human real life. Motivational factors as well as the role of verbal instructions make the difference (Wearden, 1988).

The most important contribution of behaviorism to the field of psychopathology is behavioral assessment. It links diagnosis with therapeutic change. The meaning of environmental cues and the occurrence of actual behavior can be understood by describing its contingencies. Interventions are planned disruptions in the contingencies in order to obtain behavioral change. The effect prediction is that situations that co-occur will borrow each other's meaning and the behavior followed by the expected effects will be performed more often (see Wood, 1988).

While some of the behavioral techniques that are generally applied in clinical practice were originally developed for the treatment of schizophrenia, this field has been largely abandoned by behavioral therapists (Bellack, 1986; Scotti et al., 1993). The classic learning paradigm is primarily used to change emotions (e.g. in anxiety and depression) while operant techniques were traditionally used to change behaviors, e.g. in chronic psychiatry. Behavioral intervention techniques to change behaviors can be divided into three subgroups (Delespaul & deVries, 1988; Delespaul, 1989). Subjects that are disoriented in place and time, such as some brain-damaged patients, have lost the recognition of the appropriate stimulus-configuration to perform the behavior.

They can benefit from interventions that maximize the difference between stimuli, which are S-approaches such as "prompting". Subjects who have trouble learning complex sequences of behaviors, such as most mentally retarded patients, can be helped by the progressive learning of new behaviors. R-approach strategies such as "shaping" and "chaining" are appropriate. Finally, the subjects who have trouble sustaining a certain level of behavior due to motivational disturbances should be stimulated by maximizing the consequences of the behavior. C-approaches include "operant conditioning programs", "Token Economies" and "Response Cost" systems⁷. Due to attentional and cognitive deficits in schizophrenia (Oades, 1982; Van den Bosch, 1982; Harrow & Quinlan, 1985; Frith, 1992; David & Cutting, 1994) these patients have difficulties in learning and discriminating between salient aspects of S and C. Because both the appropriate context for behaviors, as well as the effect of behaviors, are not adequately perceived by the schizophrenic subjects, they have difficulties learning the appropriate behaviors. This basic cognitive deficit of subjects with schizophrenia has bothered clinicians and was held responsible for the limited effect or failure of behavioral interventions, especially first generation social skills training (Goldstein et al., 1976; Liberman et al., 1975). As a consequence, attempts were made to regress the interventions to the level of neuropsychologic functioning. These attempts were handicapped because compared, for example, to brain-injured subjects, neuropsychologic deficits are not circumscribed and limited in schizophrenia. Nevertheless, neuropsychological assessment potentially offers the basis for a well-grounded remedial program for psychiatric patients (Erikson & Binder, 1986). Recent developments in fundamental neuropsychologic research have reactivated interest in the field (Frith & Done, 1989; Spring & Ravdin, 1992; Green, 1993; David & Cutting, 1994). Clinical applications were developed in a growing number of research settings (Spaulding et al., 1989; Stuve et al., 1991; Roder et al., 1992; van der Gaag, 1993). Behavioral therapy has not been very successful in the training of neuropsychologic functions of chronic mental patients, however. This training situation is rather abstract and the skills do not generalize well. Token Economies (Kazdin, 1976; Paul & Lentz, 1977) have been the treatment of choice for the adaptation of psychiatric patients to the institutional environment. Here, too, the generalization of the behavioral change outside the hospital grounds is a problem. Delespaul (1989) argued that the design of a Token Economy enhances or hampers the generalizability of behaviors. Because behaviors in the operant paradigm are learned as a means to an end, one should avoid making the token the reason for the behavioral change. This can be done by allowing the subjects to freely select the behaviors for which they want to earn tokens. If the subjects have realistic alternatives, the intrinsic value and natural contingencies of the behavior for the subjects will persist in the behavioral selection. When tokens are withdrawn, those natural contingencies will still persist and the behavioral change will not be lost. A Token Economy that avoids forced behavioral change but allows room for individual freedom makes it also possible to learn higher order behaviors such as

⁷ Depue & Zahd (1993) have described the relation between reward-recognition and the dopaminergic system, which is disturbed in schizophrenia. They argued that different levels of environmental stimulation can alter the working of the dopaminergic system over time.

making choices and, more generally, planning. Those ward systems are also grow-environments that will be less criticized by modern ethical norm.

Another achievement of behavioral therapy for chronic mental patients is the protocolizing of skills training procedures as propagated by, for example, Liberman (1988; Liberman & Eckman, 1989; Liberman et al., 1989). When comparing second generation skills training programs there are two important differences with their predecessors. Firstly, the training packages from the first generation were general purpose assertiveness programs while the new techniques were specially developed for chronic mental patients and are therefore targeted at this very difficult population. Secondly, the content of the skills training is based on topics that directly concern chronic mental patients. Subjects, for example, are not trained to converse on an abstract topic in which they are not involved, but are taught to use conversational skills to negotiate medication with their psychiatrist.

Social skills training was originally proposed as an alternative to the HOUND⁸ subjects who lacked the necessary skills for verbal therapy (Goldstein, 1973; Beekers, 1982; Beckers, 1986). Recent studies have demonstrated that behavioral social skills training cannot be applied universally. The best outcomes in social skill training were obtained for those subjects who were good performers in serial verbal learning or verbal memory, who were less distracted and more vigilant (as rated using neuropsychologic assessments). No relation was found with "conceptual flexibility" or the severity of their psychotic symptomatology (Kern et al., 1992; Corrigan et al., 1993). The empirical overall effectivity of these approaches has yielded mixed results (Shepherd, 1986; Bellack & Mueser, 1993).

Recently cognitive behavioral techniques have been applied to the treatment of schizophrenia (McNally, 1994; Alford & Correia, 1994; Kingdon & Turkington, 1994). They focus primarily on coping with symptoms in schizophrenia (see Liberman, 1988). Specific interventions were developed for auditory hallucinations (Haddock et al., 1993; Bentall et al., 1994; Chadwick & Birchwood, 1994); for delusions (Himaldi et al., 1993; Chadwick et al., 1994a,b; Bentall et al., 1994) and for residual psychotic symptoms (Tarrier, 1992).

Behavioral therapists claim to incorporate experimental findings in their therapy programs. Behavioral family therapy (BFT: Falloon, 1988) builds on the observation that schizophrenic subjects in remission who are discharged to families with high Expressed Emotion⁹ (EE) have a higher relapse rate (Schene, 1986; Leff & Vaughn, 1985). Successful attempts were made to influence the EE using the knowledge accumulated in social skills training programs (McFarlane, 1983; Anderson, et al., 1986; Asselbergs, 1986, 1989; Van Meer, 1991; Barrowclough & Tarrier, 1992, 1994). The most comprehensive and best evaluated interventions come from Falloon

⁸ Homely, Old, Unattractive, Non-verbal and Dumb (HOUND) as contrasted to YAVIS people, who are Young, Attractive, Verbal, Intellectual and Social.

⁹ Expressed Emotion is assessed using the Camberwell Family Interview (Brown & Rutter, 1966; Rutter & Brown, 1966). Responses of family members on questions about events and activities / attitudes and feelings concerning the patient are rated on the following dimensions: critical comments; hostility; emotional overinvolvement; warmth; and positive remarks (Hirsch & Leff, 1975). The negative criteria are used for the computation of the EE-index. Thresholds are different in different cultures and for different illnesses (e.g., depression).

(Falloon et al., 1984, 1985, 1993; Falloon & Fadden, 1993). Bellack (1993) presented preliminary data on a large NIMH study on behavioral family therapy. The results show the gains in adding family intervention to the medication alone condition in the prevention of relapse, but fail to demonstrate a difference between structured (Falloon's model) and non-structured (counseling) family therapy.

Biological therapy

Organic theories of schizophrenia assume that structural and functional brain anomalies cause the schizophrenic symptomatology. The original biological therapies were all developed in the first part of the 20th century and heralded the biologic revolution in psychiatry. Examples are the insulin coma therapy developed by Manfred Sakel in 1927; psychosurgery, introduced by Egas Moniz in 1936; and electroconvulsive therapy pioneered by Ugo Cerletti and Lucio Bini in 1938 (see Kaplan & Sadock, 1988). All have been applied to treat subjects suffering from schizophrenia but they have come under a lot of criticism and are not currently in use with the exception of electroconvulsive therapy which is still used as a treatment option for nonresponders (Lindenmayer, 1992). The use of ECT is controversial in schizophrenia (Hay et al., 1992), but it has been successful with non-responders (Ries et al., 1981; Fink, 1985).

In contrast to tradition, which defines schizophrenia as a functional psychosis as opposed to organic psychoses, the biologic etiology of the disease is now well-accepted. Judd (1994) states: "It is now unequivocally established that schizophrenia is a brain disorder — possibly one arising very early in development — and we are developing rapidly the scientific and technological sophistication to pinpoint its etiology and the pathophysiological mechanisms that contribute to what most often is a lifetime course (p. xiii)". Andreasen (1994) summarized the state of the art in a recent monograph. Despite the optimistic tone, most of the chapters are stuffed with positive findings that could not be replicated while conclusions disappointingly yield little information for targeting interventions (Clardy et al., 1994). While the biologic etiology of schizophrenia is not disputed anymore, the strongest evidence is circumstantial and its specific etiology is still unknown. For instance, we know the genetic risks of schizophrenia but not the causative genes because we lack the adequate description of attributes of schizophrenic heterogeneity (Moldin & Erlenmeyer-Kimling, 1994). Equally, Crowe (1994) stresses that biologic findings "... can only be as powerful as the clinical assessments ... that go into it (p. 258)" and Strauss (1994) wonders whether biologic psychiatry is building on an adequate base: "In psychiatry, the sophistication of biological psychiatry has far outstripped the sophistication of our research observations of patients (p. 42)".

Circumstantial evidence for a biologic etiology in schizophrenia has, however, been growing rapidly in recent years. Torrey et al. (1994a) published a monograph on a new NIMH twin study. The genetic basis of schizophrenia was reconfirmed, but unexpected predictors such as the amount — not specific — symptoms of neuropsychologic deficits or small physical anomalies were strong predictors of schizophrenic pathology (see also Torrey et al., 1994b; Green et al., 1994). Arguments for aspecific, non-genetic prenatal influences come from epidemiologic

findings. Higher prevalence for schizophrenia was found repeatedly but not systematically in those subjects whose mother was in the second trimester of pregnancy at the worst periods of different influenza epidemics (Torrey, 1986, 1988; Kirch, 1993; Cannon et al., 1994; Selten & Slaets, 1994; Mednick et al., 1994). The long-term follow-up of subjects born during the Dutch "Hunger Winter" at the end of the Second World War shows higher prevalence rates in those subjects whose mother was in the second quarter of pregnancy when the smallest rations were distributed by the German occupants (Susser & Lin, 1992). Epidemiologic studies also have demonstrated convincingly, using blindly rated data collected before the emergence of symptoms — obstetric records, systematized behavioral observations from child guidance centers, school tests such as reading speed and familial video material — that significant pre-illness behavioral changes exist (McNeil et al., 1994; Crow et al., 1994; Jones et al., 1994; Walker et al., 1994; Walker, 1991, 1994). In summary, genetic, pre- and peri-natal influences can influence the prevalence of schizophrenia, but this influence is highly aspecific as is the co-morbid symptomatology of schizophrenia.

Psychopharmacology is the most interesting biologic psychiatric research field. Since the second half of the 20th century, chemotherapy has also represented a major field of treatment for mental illness. Interestingly, most of the pharmacologic compounds that are applied in clinical practice are chance findings that bear little or no relation to the scientific findings acquired through fundamental research. Usually, the empirical evidence of therapeutic effectiveness of an intervention strategy targets new fundamental research in order to understand why it is effective. Research data come, for instance, from functional neuro-imaging studies that show the neurotransmitter occupancy following neuroleptics. Traditional psychopharmacology was looking for the optimal balance between therapeutic effect and side-effects, for which the dopaminergic system (D_1 to D_4 , but primarily D_2 receptors) was held responsible (Meltzer, 1991; Sedvall, 1994). To control this optimally, drugs with a specific profile had to be synthesized or their prescription had to be titrated to stay below the dose that creates side-effects (Carlsson, 1994). The advent of clozapine, a potent antipsychotic drug with a neuroleptic effect on negative symptoms but an atypical receptor binding, reshuffled the whole field and shifted it towards an in-depth study of the serotonergic system ($5HT_2$ receptors). Kane (1994) concludes: "At present there is reason to pursue a variety of interesting leads, but there are insufficient data on which to draw meaningful conclusions regarding what is responsible for the apparent novel effects of particular antipsychotic compounds such as clozapine (p. 183)". Leysen et al. (1994) argued that a predominant and full blockage of central $5HT_2$ receptors (beneficial for treatment of negative symptoms) concomitant with gradually occurring blockage of only part of the central D_2 receptors (for treatment of positive symptoms with minor induction of extrapyramidal symptoms) appears to be required.

Lacking the information to optimize the choice for an individualized selection of a neuroleptic, the clinician will base his first choice on the clinical picture, primarily assessing the need for sedation. He will titrate the dose over time to maximize clinical effects and minimize side-effects. If non-response is observed, the second option is a new trial with a neuroleptic from another group. Next in line is co-therapy with lithium or an anticonvulsive. Due to the side-effects of clozapine, which can be lethal

and warrants careful monitoring, making the drug extremely expensive, this option comes last (see Kane & Marder, 1993; Lindenmayer, 1992; Christison et al., 1991). The optimal choice of a long-term strategy has also be debated. Depot neuroleptics have an important part in this. Low dose strategies have been proposed to minimize the risks of long-term use of neuroleptics. Unfortunately, relapse rates are high when low doses are prescribed (Schooler, 1991; Adler et al., 1992; Meise et al., 1994). Currently, the field is changing rapidly. Different compounds are being tested. Innovations should control the extrapyramidal side-effects of traditional neuroleptics, without the dangers of clozapine kind of drugs. Important is also the targeting of drugs to influence not only the positive, but also the negative, symptoms. For the practice of neuroleptic treatment, it might be important to relate the doses to plasma concentration levels (Van Putten et al., 1991) or to develop assessment instruments that allow the dose to be titrated to the optimal binding balance between receptors, which would protect patients from being overmedicated (Leysen et al., 1994). Neuroleptic treatment of chronic mental patients is increasingly being integrated in psychosocial interventions (Falloon et al., 1993; Liberman et al., 1994). Two basic reasons exist for this trend. Firstly, relapse rates are very high (65% in 12 months) when subjects stop taking their medications (Hogarth et al., 1974; Baldessarini, 1985; Kane, 1985, 1987). Training programs are developed to ensure patients take their medication. Secondly, coping and specific environments can protect patients against the risk of relapse (Leff & Vaughn, 1985). Subjects can learn those psychosocial protections in addition to taking the neuroleptic itself. Liberman and his group demonstrated that neuroleptic doses could be targeted to symptom variation (Eckman et al., 1990; Wirshing et al., 1990) and reduced in non-responding chronic schizophrenics using a combination of plasma level control and behavioral skill training interventions (Liberman et al., 1994).

Discussion

There is a proliferation of therapeutic techniques for chronic mental patients and schizophrenics. But, although a lot of research is invested in the evaluation of those techniques, the results are still inconclusive. The paradoxical choice for interventions is between content (behavior, cognition and emotion) and process (brain structure and chemistry, and neuropsychologic functions). Interestingly, both approaches are valid.

- Tapping in on the subject's experienced life bridges a gap and allows a therapeutic relation. Listening to, but also doing things with, a schizophrenic is an adequate method and a prerequisite to achieve change. However, we should not confuse strategy with scientific truth or relevance. The content of schizophrenic behaviors, cognitions and emotions reflects the success of the subject's adaptation. Influencing adaptive behaviors is the target for interventions. Reliable assessments of those adaptational strategies should therefore be collected.
- On the other hand, the neuropsychological deficits, especially verbal memory problems, are real and are a handicap for implementing treatment strategies. Interventions should therefore also focus on those aspects. Including neuroleptics in the intervention plans is crucial. Unfortunately, here too, an important

breakthrough has still not been achieved. Side-effects, non-responders and the marginal influence on negative symptomatology are the problems here.

The conclusion remains undecided. Should we go from micro-level to macro-level — from the brain and brain function to daily life functioning — as Spring & Ravdin (1992) propagated and Green (1993) described as reasonable? Or should we try to circumvent the fundamental handicaps of schizophrenics by addressing brain functions of a higher order that they can mobilize to substitute for micro-level handicaps in more meaningful contexts and interactions? The latter strategy would allow intervention strategies that are targeted to the perceived problems by the patients and their family.

The therapeutic effectiveness of interventions for chronic mental patients is rather disappointing. New techniques are developed as a kind of antidote to the failures of traditional strategies which led to a “therapeutic nihilism” at the beginning of the 1970s. Modern modular training programs, medication and psychotherapeutic developments are important additions to our therapeutic arsenal. Those techniques sometimes come in happily to save us from frustration. We always will find another technique that can satisfy our “furor therapeuticum”; technocracy as a cover-up for therapeutic nihilism. Nevertheless, while the overall effectiveness remains unclear, this technocracy protects schizophrenics from being forgotten.

Anthony et al. (1990) have propagated the negotiation of means and ends for interventions between patient and therapist. It is generally accepted that psychiatric rehabilitation should be tailored to the subject’s personal aims and should consider the demands of their own living environment and the strengths and weaknesses of the patient. To attain this we should develop communication methods that allow us to relate to schizophrenic subjects and to elucidate the subject’s aims. On the other hand, assessment strategies should be optimized to generate information on those adaptive behaviors — including emotions and cognitions — that can directly be included in the planning and negotiation of interventions. Behavioral assessment can offer us the necessary tools for negotiated care (Delespaul & deVries, 1988, 1991).

Chapter 2: Alternative Assessment of Psychopathology ¹⁰

In Chapter 1 I have described how intervention techniques for subjects with schizophrenia are proliferating. Over the last decade the development of psychoanalysis and behavioral therapy has achieved important gains in protocolizing treatment strategies. In addition, biological psychiatry offers a wide range of new neuroleptics, while social psychiatric interventions have demonstrated possibilities for treatment of patients outside the hospital. But while all those strategies have had some degree of success, the conclusion from the preceding survey of therapy was not optimistic.

We have a limited understanding of the illness process in schizophrenia. Therefore, we do not know which intervention strategy is appropriate. In practice, we can choose from a wide range of interventions that differ in mechanism, but have similar therapeutic effects (Crencevage et al., 1993). In clinical practice, the choice of intervention is often made by trial and error or based upon the therapist's personal preferences. Little empirical data is available to aid in the choice of optimal treatments. Even less is known about the next alternative in line when the first intervention fails to produce change. Often, schizophrenic patients cannot adequately be helped. The subjects who are called "therapeutic successes" often remain a disabled population for which we still need to optimize our intervention strategies (Lehman & Cancro, 1985). Unsuccessful cases — non-responders or hard-core patients (defined primarily according to neuroleptics: Lindenmayer, 1992; Christison et al., 1991; Kane et al., 1988; Meltzer et al., 1993) — fill the largest part of our residential mental health facilities or wander homeless through the streets (Lamb, 1984; Bachrach, 1984; Wright, 1989; Bughra, 1991; Susser & Struening, 1990).

In order to optimize treatment, we need to know which intervention is optimal for which patient in which situation. Therefore we need information obtained from evaluation research as well as better assessment strategies for individual patients. We need to improve our diagnostic tools in order to select those patients who will respond to specific treatments from within the heterogeneous group of schizophrenics. However, selecting only those subjects with the highest likelihood of success is an unsatisfactory strategy. Decisions for optimal treatment should not exclude any patient from care and are, therefore, a problem of indication, not of selection (Delespaul, 1981). Of course, cost efficacy arguments can be relevant in the selection of the most suitable subjects for treatment. However, the costs involved in (chronic) schizophrenia are enormous, and therefore, even seemingly less cost-effective interventions might be the best treatment options in this population. The question remains: which intervention for which subject in which situation? There is no simple answer to that question. The assessment process, relevant for treatment selection, is not limited to a single stage of treatment but is an iterative process in which the interventions are continually updated, based on new information collected for the evaluation of the

¹⁰ With thanks to Justin Kenardy for the useful discussion of the material.

evolving treatment process (Delespaul, 1981). Patients are continually adapting to their changing situation. Therefore, working with patients in therapy should also be a flexible, dynamic process and requires an ecologically valid ongoing assessment. In this chapter we will discuss how assessment can be optimized in order to understand the dynamic process of psychopathology and to optimize treatment selection. The intention is not to solve all the complex problems related to the assessment of psychopathology. Instead, we will question the current diagnostic tradition and propose some changes. We will focus here on the content and form of the assessment process: the role of emotions in psychopathology assessment and the importance of time- and context-related variability.

ISSUES IN THE TRADITIONAL PSYCHIATRIC ASSESSMENT

Subjective versus objective assessments

The data collection strategies of traditional assessments in psychiatry can be divided into the following broad categories: self ratings, ratings by others, behavioral tests such as responses to neuropsychologic tests, and biologic measures. Three of these alternatives are attempts to substitute objective measures for the subjective experiences that form the basis of psychopathologic phenomena.

What are arguments for the use of objective measures in preference to subjective experiences? Intriguingly, subjective experience is central to our concept of psychopathology and there are no good alternatives as valid measures of psychopathology. For instance, we know from empirical studies that muscle movements in the throat or increased cerebral blood flow in specific brain areas are observed during verbal hallucinations (Gould, 1949, 1950; McGuire et al., 1993). However, even with the most sophisticated current or future assessment technique, the experience of verbal hallucinations will never totally match throat muscle movements without outer speech or increased blood flow in the Broca area. It is the subjective experience of "voices coming from outside without sensory stimulation of the ear" that is fundamental to our definition of hallucinations. Likewise, we can observe anxiety based on the subject's avoidance behavior or facial expression, but the experience itself can only be understood using the subject's self-reports. Generally speaking, the so-called objective correlates (observations by others, behavioral tests and even biologic markers) of inner phenomena are often poor analogs of the experienced emotion. Lang (1985) argued, in discussing "anxiety":

"For most patients, the basic datum (...) is their feeling state, i.e., a direct experience of internal apprehension, requiring no further definition. In some cases even the most careful clinical inquiry may elicit from the patients no more elaborate description of their disorder..."(p. 131)

Does the unreliability of self-reported data, warrant the use of less valid, but possibly more reliable, "objective" assessments? In research, validity is more important than reliability. Valid assessments are reliable but reliable assessments are not always valid. Therefore, there is no argument in favor of reliable objective assessments over

valid subjective ones using general decision principles. Of course, the objective markers of schizophrenic psychopathology such as biologic markers and behavioral assessments, are developed with a different aim: they promise to enhance our understanding of the etiology of the disease. Consequently, this is an valuable and important field of research. Unfortunately, the development of phenomenologic assessments of schizophrenia has not kept pace with the development of objective markers (Strauss, 1994). The result of this process is that theoretic models are often based on correlates of objective markers, possibly leading to capitalizations on error.

In conclusion, there is no uniform objective correlate for most psychopathologic symptoms. In addition, while objective measures can be more reliable, their validity is unclear. In view of this, it is intriguing to observe how little has been done to enhance the reliability and validity of subjective assessments.

The problem of enhancing the reliability of self-assessments is an important one. The validity of objective indicators of psychopathology is low. In addition, to assess psychopathology we often have no other alternative than to rely on introspection (e.g. Luriya, 1965; Hurlburt, 1990, 1993). Unfortunately, introspection is unreliable. According to Piaget (1973) subjective reports are incomplete — they grasp the results of mental processes, not their underlying mechanisms — and distorted — the subject projects his or her own philosophy into the reports. The common solution has been either to embrace introspection uncritically and exclude empirical rigor, or to reject the evidence of introspection as false, delusional, or heretical (Laughlin et al., 1990, p. 21). Although this controversy — related to the mind-body problem — has played a role in the history of human science for centuries (Damasio, 1994), the most important bifurcation point for modern history was crystallized in the critique of behaviorism (Watson, 1924) ¹¹ based on the psychologic methodology of Wundt (1896) ¹².

Monson & Hurlburt (1993) reviewed the evidence of the controversy that led to the prominence of behaviorism for half a century. They concluded that the basic critique — that introspection generates unusable material for research because it leads to divergent, sometimes even contradictory, results — is not correct. The problem is not the quality of the raw data generated by the self-reports, but the interpretations, speculations and theory-building based on this raw material.

In line with the Monson & Hurlburt, the disturbing influence of the interpretation should be minimized in self-reports. To avoid corruption of the data collected by introspection and to enhance reliability, we should minimize the interpretations derived from the data and improve the data collection strategy itself. The empiricism of behaviorism can be helpful for this. But, as I will demonstrate (Part 2), Monson & Hurlburt are optimistic when they assume that introspective data are valid by design.

¹¹ "Behaviorism (...) holds that the subject matter of human psychology is the behavior of the human being. Behaviorism claims that consciousness is neither a definite, nor a usable concept (Watson, 1924 (1970, p. 2)).

¹² "Da die Psychologie nicht spezifische Erfahrungsinhalte, sondern die allgemeine Erfahrung in ihrer unmittelbaren subjectiven Beschaffenheit zu ihrem Gegenstand hat, so kann sie sich auch keiner andern Methoden bedienen als solcher, wie sie von den Erfahrungswissenschaften überhaupt zur Feststellung von Tatsachen sowie zur Analyse und kausalen Verknüpfung derselben angewandt werden." Wundt, 1896 (1911, p. 24).

Categorical versus dimensional assessment

Psychiatric nosology such as in the DSM-IV (APA, 1994) continues a medical tradition and assumes a categorical view of illness. In this model schizophrenia is considered different from major depression, dissociative disorders and post-traumatic stress. The exclusive categorical medical assessment model considers shared symptoms to be less relevant because they are nonspecific and blur distinctions between disorders. Illness entities are defined using a set of specific symptoms not shared by other diseases or using a concurrent and specific combination of nonspecific symptoms. The categorical approach has been successful because it facilitates research into the etiology of a syndrome and the evaluation of specific treatment strategies. A noncategorical disease concept would make this kind of research more difficult to conduct and the results more difficult to interpret. In order to enhance the reliability, the categorical approach defines strict rules for syndromal decision making. They include symptom definitions to determine syndrome caseness and exclusion rules to decide upon the differential diagnosis. Exclusion rules created confusion and some of them were abandoned in DSM-III-R. The actual implementation of DSM-III-R is, however, still quite confusing (Van Praag, 1992), a situation that is only marginally enhanced in DSM-IV.

As for syndromes, the symptoms themselves are often categorical in medicine. Anxiety is different from panic, low mood is different from depression and daydreaming is different from hallucinations. This difference is not assumed to be quantitative but qualitative. The presence of symptoms is considered important in the categorical model, while the intensity has less relevance for diagnosis and treatment. The dimensional approach applies even more to symptoms than to syndromes. "Although there appear to be normal counterparts of symptoms such as paranoia (suspicion), there do not seem to be normal counterparts of schizophrenia" (Costello, 1993, p. 16). The study of normal counterparts of individual psychotic symptoms is fruitful (Romme & Escher, 1993; Chapman et al., 1976, 1978, 1984) and shows that symptom frequency and intensity are important for making a distinction between illness and well-being. By allowing a differential rating of illness severity, dynamic assessments of variability become possible. Dimensional ratings can be used to assess illness severity but also to evaluate the illness-related adaptation. The study of normality is crucial for our understanding of psychopathology (Offer & Sabshin, 1991).

In contrast to the medical tradition, psychologic assessment uses a dimensional model to describe the phenomena under study. In this model, different illnesses are interrelated and therefore not discrete entities. The dimensional model allows discrete diagnoses to share a common core of symptoms. Moreover, the intensity of the characteristics is important to weight the syndromal summary. The implicit assumption of overlapping syndromes is a shared etiology.

Dimensional approaches to schizophrenia have led to a wealth of scales. The work of Eysenck & Eysenck (1976) on "psychoticism" is an attempt to operationalize the continuum of the schizophrenic experience into normality. A similar scale, labeled "schizophrenia", has been extracted from the MMPI/MMPI-2 (Graham, 1987, 1990), but is not very successful as the sole predictor of schizophrenic psychopathology.

Unfortunately, scaled constructs are not real experiences in the subject's daily life. They are the result of statistical techniques and we need computers to assess them. The different items are weighted based on empirical correlations and the face validity of the items that form a construct is often difficult to understand. Usually, neither the subjects nor skilled raters are aware of the actual position of an individual on the scale. Due to the complexities involved in the computations, the factors are primarily useful for assessing groups of subjects.

Empirical evidence supports the dimensional model. For instance, the heritability of schizophrenia, using Schneiderian symptoms is low, while broader criteria such as the RDC result in much higher predictions¹³. Interesting also is that the concordance ratio between mono- and dizygotic twins for schizophrenia is much higher when using a broad criterion including schizo-typical personality disorders and psychotic depression compared to the more narrowly defined DSM-III diagnosis (McGuffin et al., 1984, 1987). Such studies show that there is no real empiric support for the use of a narrow definition of schizophrenia in the study of biologic schizophrenic etiologies, nor in treatment research. For this reason the omission — due to ethical restrictions — of MMPI measures for the dimensional assessment of familial history of mental illness in the recent NIMH twin study has been regretted (Torrey et al., 1994). The higher familial linkage, using broad definitions of schizophrenia spectrum disorders, including manic-depressive illness or personality inventory data, makes it more likely that the figures of genetic determinism of schizophrenia are underestimated. All this illustrates the relevance of the observed variability up to normal adaptation for fundamental research in schizophrenia (Frances et al., 1991; Cloninger et al., 1985; 1990).

An intriguing fact is that the inattentive choice of an assessment strategy potentially fools us into believing that syndromes are either categorical or dimensional. As a consequence, the conclusions and theories derived might be contaminated by the hidden assumptions of the data collection strategy, resulting in inaccurate assessments of the occurrence, frequency and intensity of the phenomena under study. If discrete events are assessed, discontinuity is assumed. For instance, when we ask a patient if he has been hallucinating over the previous week (a DSM-III-R schizophrenia criterion) the result will be "yes" or "no". Equally so, when a clinician summarizes observational and interview data to assess symptoms that the subject might not be aware of himself, such as delusions, the answer is categorical: present/absent.

Due to the phenomenology of some symptoms, they can only be rated retrospectively. In the flow of experience, the awareness that the experience has reached the level of a delusion for the past period of time grows over time. After a while the subject realizes that when he thought his food was poisoned, he must have been delusive. Equally so, the starting and end moments of prodromal and residual phases are not well-defined. In retrospect, we draw a line and interpret slight changes that passed unnoticed when they occurred as the beginning of a new phase. This illustrates that in the moment-to-moment experience of subjects, most of the symptoms are experienced in continuum from normal experiences to the intensity level of a clinical symptom. Once again, it is only in retrospect that psychiatric symptoms are described as discrete events.

¹³ Schneiderian criteria were defined by Schneider (1939, 1959) while the Research Diagnostic criteria were defined by Spitzer et al. (1975).

When assessing symptoms using a visual analog scale, a continuum is assumed. When we assess the constituent signs and symptoms of a syndrome on a set of scales — even yes/no alternatives — the resulting syndrome is automatically evaluated as a continuum. We then need a cut-off score to dichotomize it again and to conclude the syndrome is present or absent. Both models — the categorical and the dimensional — can be defended on empiric grounds and are the consequence of divergent theoretical positions. For instance, when hallucinations are rated on a continuum — as in the BPRS (Lukoff et al., 1986; Ventura et al., 1993a,b) or in the SANS (Kay et al., 1992) — or schizophrenia is diagnosed as the summary of a set of yes/no scales — as in the DSM-III-R (APA, 1987) — the studies conducted using the resulting assessments will yield comparable results (see Mulder, 1994). We will draw the same conclusions when using dichotomy ratings of “paranoia” or dimensional ratings of “suspicion”. Equally, because the presence or absence of hallucinations is related to the growing intensity of disturbingly realistic daydreams, we will get the same kind of results in studies that use both assessment models. Although there is no empiric evidence that allows us to decide which model is true, the dimensional approach is more appropriate in prospective research where the assessment of symptom variability is important.

Symptom versus syndrome assessment

The schizophrenic syndrome is heterogeneous and one of its core symptoms (hallucinations) is also shared by other diagnoses such as dissociative states or substance abuse. This warrants a better assessment of symptoms in psychiatry (Klein & Riso, 1993). Empiric observations show that schizophrenic symptoms are interrelated and clusters together. Therefore a core disease entity can be assumed (Liddle, 1987; Liddle & Barnes, 1990). However, the relative salience of different symptoms for the diagnosis of schizophrenia varies widely from one study to another (Zigler & Phillips, 1961; Slade & Bentall, 1988; Costello, 1993c). As a consequence, diverging symptom definitions of the illness were formulated, adding to the confusion. For instance, only 28% of the DSM-II schizophrenics were still schizophrenics under DSM-III (Klein, 1982). Fortunately, the DSM-IV is traditional in its approach to diagnosis and experimental draft options were not withheld (APA, 1991, 1994). The same problem arises when rating the occurrence of different symptoms using standardized instruments such as the PSE and the DIS. The frequency of different symptoms also varies widely (Costello, 1993).

A central problem in schizophrenic research is that the patients are most difficult to study in the acute phase of illness, when the symptoms are most salient. Therefore, Elkins & Cromwell (1991, p. 27) suggest that we should investigate the “neurological basis of schizophrenia (...) which is present not only during acute illness but also (a) premorbidly, (b) in remission, and (c) in healthy first-degree relatives.” It would indeed be very helpful if we were able to enhance our understanding of schizophrenia by extrapolating from adaptive mechanisms in less disturbed situations such as remission states. In such an approach, we should not restrict ourselves to the data collected in the traditional assessment, but redefine the domain of symptomatology (Cromwell, 1984). Knight (1987) suggested switching from the traditional dependent and

independent variables and using "process-oriented" strategies to help generate deficit types as independent variables and symptoms as dependent variables. This process bears similarity with the "ascending cross-classification" (Van Meter, 1990) in which disease-related symptomatology is used as the database for enhanced syndromal clustering. An aim of our research endeavor is to enhance the psychiatric assessment of symptomatology and possibly redefine syndromes in schizophrenia research.

TOWARDS ECOLOGICAL VALIDITY — FROM LABORATORY TO REAL LIFE

Our solution to the issues raised in our discussion of psychiatric assessment strategies is based on three arguments. Firstly, we will defend that the study of emotions can be valuable to understand psychopathology, even in schizophrenia. Secondly, we will stress that the variability of symptoms over time should be assessed to understand how from a process point of view, illness evolves in real life. Finally, we will demonstrate that concurrent observations of the context can be assessed and included in our understanding of the dynamics in the adaptational process. These characterize different experiences of illness and well being.

Studying emotional variability

Our aim is to study the processes underlying the changes in psychopathology, using data collected during periods with less florid symptomatology. We will argue that the assessment of emotions and emotional change allows the charting of important sub-threshold changes in psychopathology. We believe that the study of emotional variability can be used to assess the adaptational mechanisms that regulate the development of psychopathologic symptoms before its full-blown emergence.

As for psychopathology, the study of emotions is plagued by the assumed unreliability of self-reports. Self-reports are constantly under suspicion (Hayes & Cavaioir, 1977; Nelson et al., 1978) due to the expected response tendencies (Wiggins, 1964; Rorer, 1965; Norman, 1967; Fiske, 1971; Wegner & Vallacher, 1980). As a consequence greater weight has been placed on other sets of data such as behavioral observations and physiologic data. In order to avoid the pitfalls of depending on less reliable self reports, the behavioral expressions (Camras et al., 1993; Pittam & Scherer, 1993) and the neurophysiologic mapping (Koukkou-Lehmann, 1987; Debus & Machleidt, 1990; Cacioppo et al., 1993) of emotions have been studied. Others have focused on the cognitive (Strongman, 1987; Ortony et al., 1988; Plutchik, 1993) and adaptive (Frijia, 1985; Izard, 1991, 1993) processes to define separate emotions.

The study of emotions using reliable, objective measures assumes a deterministic model of consciousness, adopted from leading French philosophers such as Descartes (1596-1650) and Laplace (1749-1827). The likelihood of determinism, modeled by mathematic or physical laws has been challenged (Eccles, 1994; Penrose, 1989, 1994). Emotions or consciousness in general is something different than the behavior, for instance facial expressions, that can be observed. When we try to chart the processes that regulate human functioning, the adaptive and integrative assimilation to

the internal and external context of the emotions becomes more important (see Harré & Gillett, 1994). Objective lab assessments of emotions become less adequate and ecologic validity more important (Bronfenbrenner, 1979; Maher, 1991; Chow, 1992).

Emotions and psychopathology

The role of emotions in the study of psychopathology has a long tradition. Using a psychoanalytic frame of reference, Rado (1969) defined psychopathology as an over-reaction to emerging emotions such as fear. Brenner (1975) considered symptoms as (failing) compromise formations. In behavioral therapy, the emotions and the cognitive processing are considered the psychopathologic agents (Eysenck, 1947, 1987; Foa & Kozak, 1986; Eysenck & Martin, 1987). The behavioral model tries to understand emotions through the context in which they occur. Lazarus (1991) gives a synthesis of psychoanalytic and behavioral theories of emotions and conceptualizes psychopathology as a failure in appraisal and coping, which results in emotional states, ego defenses, and self-deception. In normal life the spectrum of different emotions is broad. However, in the study of psychopathology the emotions that have been given important roles are restricted to "anxiety" and "depression". Moreover, those emotions are considered syndrome specific. The observation that emotions, already shared by all "normal" individuals, are also shared over the boundaries of psychiatric syndromes is not self-evident and has to be argued in psychopathology ("co-morbidity" argument: Racagni & Smeraldi, 1987; Maser & Cloninger, 1990; deLisi, 1990). Changes in emotions and selecting environments as well as activities through emotional monitoring have been proposed as central in normal human adaptation (Csikszentmihalyi & Selega-Csikszentmihalyi, 1988; Csikszentmihalyi, 1990). Nevertheless, the changes in the normal spectrum of emotional life is underresearched in mental health patients. Two factors deserve reconsideration: a) the restriction of the spectrum of emotions to anxiety and depression in psychopathology research; and b) the assumption of a static relationships between emotional symptomatology and specific syndromes on symptom intensity, disregarding emotional processing.

Emotions in schizophrenia

In the DSM-III-R description of the schizophrenia, relatively little room is left for the intensity of content specific emotions such as anxiety and depression, but emotions are present as process variables (absence: anhedonia, flat affect; or erratic occurrence: inappropriate affect, sudden mood swings). Much trouble has been taken to draw distinctions between schizophrenia and affective psychosis. Concerns come from the observed differences in scope of the schizophrenic diagnosis between the US and European nosologic systems (WHO, 1973, 1975; Carpenter & Strauss, 1979). These were solved by the APA adoption of DSM-III (1982) and the exclusion of affective psychosis from schizophrenia. The adoption of DSM-III criteria was an important next phase, but the benefits from a scientific point of view are meager (Van Praag, 1992). Although the traditional symptom sets are different for schizophrenia and affective

disorders (Rogers & Winokur, 1990), the genetics (deLisi & Hoff, 1990)¹⁴ as well as the time course (Leff, 1990) cannot be adequately discriminated. Thus, the debate on affective symptoms in schizophrenia is ongoing (deLisi, 1990; Ferrero et al., 1992). Foulds & Bedford (1975) have proposed a hierarchic model of symptoms across psychopathologic syndromes and argue that the symptoms of lower level syndromes, e.g. emotional disturbance, are included in higher order syndromes, e.g. schizophrenia and organic brain disease. Affective symptoms, therefore, become part of the schizophrenic syndrome. Ciompi (1982) also challenged the overemphasis on first rank symptoms and the exclusion of affective symptomatology in the schizophrenic syndrome. In his conceptualization, affective structures such as confusion, tension, anxiety and changes in mood are central symptoms. Machleidt (1992) also defines schizophrenia as a syndrome characterized by a hyperexcitability of the anxiety system. Schizophrenia, however, differs from anxiety psychosis in its phenomenology and course. In his study of schizophrenic symptomatology, Kay (1991) includes affective states in the general pathology scale of the PANSS (Kay et al., 1992). The PANSS is almost always used to assess the dichotomy between subjects with predominant positive or negative symptoms. The more complex pyramid model that can also be assessed with the PANSS, however, was preferred by Kay for the description of the symptom structure of schizophrenia (Kay & Sevy, 1990). In this model, negative, positive and depressive features constitute divergent points of a triangular base that represent the schizophrenic symptomatology, while excitement makes up a separate vertical axis and completes the pyramid structure. Once again, the proposals indicate the possible importance of emotions in explaining schizophrenic heterogeneity.

Problems in studying emotions in schizophrenia

The assessment of emotions relies primarily on self-monitoring. The assessment of emotions using self-monitoring techniques is problematic generally, but even more challenging in schizophrenia. Van der Gaag & Haenen (1990) have demonstrated how schizophrenics are impaired in the monitoring of facial expressions of emotions. In his book on the "Cognitive Neuropsychology of Schizophrenia", Frith (1992) describes the impairment in self-monitoring as the central deficit in schizophrenia. He argues that the inability to monitor willed intentions can lead to delusions of alien control and the inability to monitor the beliefs of others leads to delusions of reference, paranoid delusions and incoherence, sometimes even hallucinations. According to Frith, the inability to make reliable meta-representations — representations of mental and physical states — is the missing skill in schizophrenia. Self-monitoring by schizophrenics should therefore be used with caution.

We challenge this assumption. Of course, schizophrenic symptomatology can be described as a problem in meta-representation. However, meta-representation requires perceptual as well as integrative and synthesizing skills. The self-monitoring problems of schizophrenics, described by Frith, primarily emerge in situations where the

¹⁴ With the current state of the art and rapid changes in genetic research, we expect to see a lot of new evidence in the near future in this field. However, considering that behaviors, emotions, and even aesthetic preferences have been localized recently on human genomes, chances are that the actual boundaries between schizophrenia and affective psychosis will be redefined.

subjects are asked to summarize and integrate emotional self- and other-assessments. Birchwood et al. (1994) demonstrated the variability in this reflective deficit by showing a inverse relation of insight and acute phase schizophrenia. Blanchard et al. (1992) challenged the uselessness of self-reports by demonstrating that negative mood assessments by schizophrenics were better predictors of outcome at follow up, than observations by a clinician. We have demonstrated that emotional self-assessments collected using non-interpretative techniques are reliable and valid in schizophrenia (Delespaul & deVries, 1987; deVries & Delespaul, 1989).

The observation of variability and co-variability of symptoms allows us to interpret the nature of the illness. We need observations of changes in frequency and intensity of symptomatology to draw conclusions on the cause and maintenance of the symptoms. When no changes in symptomatology are observed, it is difficult to make causal inferences about the occurrence of symptoms. The most appropriate observations of variability should be made in the natural context of emotions, the context in which the normal environmental contingency structure applies. The observations of patterns in time and their contextual determinants allow us to understand the processes at hand.

Time effects

"Human life unfolds in time. All human activity has a temporal dimension, and all of it takes a place in a temporal as well as a spatial context. Time, therefore, would seem to be a fundamental dimension of every field of the social and behavioral sciences, as it has been for most of the physical and biological sciences" (McGrath, 1988, p. 7).

Time is the dimension that orders the sequence of consecutive experiences. Time is also the structure that defines the causality direction when we interpret change (Kelly & McGrath, 1988). Assessments of change assume a (linear) time structure underneath. Without this time structure it would be very difficult to interpret data. Linear time is an important methodology-induced reduction used in psychology research to structure data (Slife, 1993). Because time is so embedded in our daily life we see it as a natural thing and forget that the measurement of time is man-made. However, by structuring observations in linear time we can distort the real relations of the phenomenon under study (see section on "Functional analysis"). We should therefore be careful in our interpretations.

The experience of time

In his essay on time, Elias (1987) stresses the cultural roots of the concept of time. Time is traditionally measured by scaling a motion in space. Developments in the measurement strategies have always influenced our perception of time, as we realize when we compare time experience between cultures. The precision of time measurements based on oscillations within a quartz crystal is now universally available in digital watches and has dramatically changed our experience of time (Toulmin & Goodfield, 1965; Landes, 1982; Deleuze, 1989). Even so, the "opening" of the world through TV, radio and telecommunications has broadened the scale of time checking

and has created the need for worldwide synchronization of time. The result is a measurement of time that reaches much farther than the concept of experienced time or social time (Jones, 1988).

The absolute nature of modern measured time does not fit our day-to-day experience. From the point of view of subjective experience, time moves fast or slowly according to the circumstances and moments. The basic characteristic of experienced time seems to be that it structures — sequences — experiences. Seriation, or the concepts “before” and “after”, is learned concepts according to Piaget (1966). As Elias puts it, “the unchanging recurrence of the same sequence patterns, such as the cycle of the seasons usually looms much larger in the knowledge of people at this stage, than the succession of years which never return” (Elias, 1987, p. 7). The concept of time that never comes back, so prominent in modern culture, is not the only concept of time that organizes our experience. Linear time is a concept that emerges from the experience of finiteness. The necessity to order experiences such as “before” and “after”, as “what was”, “is” and “has to be”, comes from the experience that the available time has boundaries (“tempus nascendi — tempus moriendi” and “tempus fugit”). Without reference to an end, however, the experience of time is recurrent. Patterns of previous experiences are reactivated and the moment, with the same meaning, repeats itself.

Subjects give meaning to themselves in the environment by associating and discriminating between elements of their living context (see the section on Functional Analysis). Meaning is extracted by identifying repeating sets of salient cues. For the subject, adaptation assumes prior knowledge of the meaning of the situation. Known aspects of the situation reactivate those memory sets and allow adaptive behaviors. Contrary to the experience of the irreversible flow of time, the recurrent identification of known patterns is important for daily life adaptation. Life, as it is experienced, assumes recurrent time. Without this, no prior knowledge would exist and no purposeful behavior would be possible.

It is our conviction that the experience of subjects is modeled through their experience of time (see also Bergson, 1907; Heidegger, 1962). Also, the assessment of experienced time is derived from the way subjects relate to their environment and cope actively. To understand the process through which situations become meaningful to subjects, it is important to evaluate how the subjects experience time.

Studies have demonstrated that psychopathology influences the experience of time (Edlund, 1987). The model for studying time perception is the rating of time intervals under standardized conditions of disruptive activities or disturbing emotions. The concept of timelessness has been put forward in relation to chronic mental illness by many authors (Polstra, 1988; Crow & Stevens, 1978). This timelessness translates into the fact that the world is perceived as unchanged over time. Chronic mental patients do not perceive differences in situations adequately. Richard Lamb challenges this moment-to-moment experience of timelessness: “chronic mental patients are eternally moving to nowhere”. In other words, the moments affect subjects with schizophrenia, but the consequent behaviors have no purpose and the whole process is perceived as unproductive — meaningless. Therefore, timelessness in chronic mental patients is not only a momentary but even more a generalized experience.

For acute schizophrenic subjects we assume that the flow of time is hazardously disrupted. Past and present are mixed up. Because of this, behaviors and emotions should be unrelated to the environment. For the observer, the salient cues on which the subject reacts are unrelated to the actual context and therefore the resulting behaviors or emotions are difficult to understand. In depression, however, time is slowed down and this, contrary to chronic mental patients, affects momentary experience as well as generalized long-term evaluations. For those subjects environmental changes will not be recognized or behavioral and emotional responsibility will be low.

The structure of time

The acknowledgment of the limitations of the linear/Newtonian concept of time is not restricted to mental phenomena or human experiences (Casti & Karlqvist, 1989). The structuring of observational information on a linear time frame has been very productive and has given us important laws in physics, e.g. "gravity". However, the traditional concept of time has now largely been abandoned (see Prigogine & Stengers, 1984; Hawking, 1988; Coveney & Highfield, 1990). Although Hawking (1993) warns against transposing the laws of science to human behavior, it is intriguing to observe how the concept of deterministic causality that is applied in psychology and psychiatry has not evolved beyond state of the art of the last century, while in physics non-linear time models and probabilistic assessments have become more prominent (Slife, 1993).

Musès (1985) introduced the term "chronotopology" to distinguish the interactive connectedness of human systems with time. He describes how human actions and systems can only be understood when we include information on past, present and (anticipated) future in the equation. Equally, Edelman (1987, 1988, 1989, 1992) argues that the development of the self needs a subjective model building of the environment in terms of the present awareness, the past and the projection into the future. Symbolic memory and human speech allow the formation of this social self (Bruner, 1990). The self is socially constructed with culture and influenced by narratives. The study of the self should therefore be a study of the subjective experience — represented in human symbolizations (narratives) — as it evolves over time, including its anticipation of the future. For this line of research we need a new methodology that allows the appropriate collection of experiential data (Lake, 1993). To derive meaning from those data we also need a methodology that structures the observations thereby mirroring the fuzzy non-deterministic aspects of human reasoning and attribution of meaning (see Kosko, 1993; McNeill & Freiburger, 1993). In psychology we still usually use deterministic schemes to interpret our data. We hope to find the personality trait that would predict the behavioral reactions and are disappointed when we see that this predictive power is far from perfect. When we are not able to predict behaviors from specific stimuli, we think our assessments have failed. We regard non-fitting observations as falsifications from our hypothesis. This procedure, however, does not fit with the fuzzy nature of the relations under study. This leads to non-deterministic assessments at best, and therefore predictions of individual behaviors are never perfect.

Environmental effects

The environment is central to our understanding of changes over time. We feel different when we are alone compared to in social situations. Ecology is the field that studies the complexity of the behavior-environment interactions from a dynamic-adaptive point of view. Concepts and methods were developed in ethology, the study of animal behavior (Tinberger, 1965, 1972; Lendrem, 1986; Slobodchikoff, 1988). The application of ecologic principles to the field of human adaptation stems from the 1940's and was propagated by leading psychologists of that time such as Kurt Lewin (1936, see Patnoe, 1988) and Egon Brunswik (1956, see McDonald & House, 1992). This has resulted in major epidemiologic studies on psychiatric prevalence in the community, for instance the "Midtown Manhattan study" (Srole et al., 1962), the "Nova Scotia" study in a rural environment (Leighton et al., 1963) and the study on social class conducted in New Haven (Hollingshead et al., 1958). Later, authors such as Bateson (1972, see also Hammersley & Atkinson, 1983) have developed models to describe the individual person-environment interactions. To explain the salient environments in which we can understand and predict the occurrence of human behaviors more meaningfully, I will explore two ecologic concepts related to those environments: "behavioral settings" and "perceiver-environment reciprocity".

Behavioral settings — The notion of a simple stimulus explaining behavior applies only in relatively rare situations in the natural flow of daily life (van Heck, 1989). The adaptive behaviors of subjects need a context, a set of concurrent environmental aspects that makes the environment meaningful (see the section on "Functional analysis"). Environments can be described by the place, the persons involved, the activities performed and the time. When we think about those issues we quickly acknowledge that places have different meanings for an individual, depending on the activities that are performed and the persons that are involved. The living room is different during a party, for a housewife while cleaning or for a couple discussing an insurance contract with a salesman.

Appropriate behaviors vary with situations (Wicker, 1979). The concept that applies to those meaningful contexts that synthesize the place, persons, activities and time aspects into a meaningful moment, is the "behavioral setting". Procedures to define which settings have the same meaning — inter-setting congruence — have been developed by Barker (1968, 1978). The concept of behavioral settings was originally a static concept. Wicker & King (1988) have described the developmental dynamics of "behavioral settings". According to their research, during the life cycle of a behavioral setting, the meaning and appropriate behaviors change.

Reciprocity of perceiver and environment — The occurrence of behaviors can be explained by describing the context in which they occur. Ecologic psychology has introduced the notion of complex environments ("behavioral settings") to explain the similarities in the behavioral repertoire of individuals acting and interacting in the same kind of situations. However, the "behavioral setting" does not explain the systematic patterns in the differences in behavioral reactions of different individuals

in the same environment, nor the variability of the behavioral patterns within individuals but over time, in those behavioral contexts.

Stokols argues: "... although much progress has been made over the last two decades in achieving a more contextually oriented approach to the study of environment and behavior, an important challenge for future work in this field is to give greater expression to the reciprocal, dynamic, and fortuitous aspects of people-environment transactions" (Stokols, 1988, p. 236). The reciprocity of perceiver and environment is a central notion of Gibson's Ecological Psychology (Gibson, 1979). Based primarily on a study of illusions, Gibson demonstrated that the subjective characteristics of the moment of perception, such as viewpoint and prior knowledge, drastically alter what is seen. This line of research resulted in an ecologic theory of perception (Lombardo, 1987). We argue that the reciprocity between perceiver and environment is central to our relationship to the world. Stimuli or environmental cues are not objective physical entities, but are perceived differently by each individual and at each moment. A "stimulus" is always a "stimulus-for-the-subject". This does not mean that stimuli cannot be studied. It only illustrates why a dog will not always start eating when he sees food. The food is a different stimulus when the dog is hungry. This can be determined by observing its previous deprivation from food.

Environments and experience — The study of stimuli has primarily been a study of behavioral consequences of environments. The effect of context on subjective emotion is less well understood. However, Walter (1988) defined a place as "...a unity of experience, organizing the intercommunication and mutual influence of all beings within it." Places and more generally the ecologic realities in which subjects behave and adapt, are related to experience. Subject and environment always interact. Environments model the subject, while the subject also builds the environment. The observation of this interaction between subjects and environments — both behaviorally and experientially — are reflections of culture, self or subjective identity and momentary experience.

Cultures have traditionally been studied by describing special moments such as rituals. Recently, however, more emphasis has been placed on normal, daily life activities. The cultural reality of daily life behaviors is studied by describing the "time use patterns" of the persons (Szalai et al., 1972). Time use patterns are the distribution of time for a specific individual or group of individuals over places, persons, activities or over different mental states. Time use patterns are different between cultures and within cultures over time (Robinson, 1988). The study of time use patterns is primarily based on a diary technique. This has methodologic implications (Ås, 1978), some of which will be described in the next chapters.

Time use is an important parameter to describe how persons experience their life. Grønmo (1982) has given an indication of how time use patterns can be related to the concept of "quality of life". More important than the mere description of activities is the concurrent monitoring of emotional variability with activities, places and persons the subject interacts with. Using a concurrent monitoring of experience and activities Hormuth (1990) shows that the environment models the self, and drastic changes in that environment can create clashes in the self: the culture shock.

The ecologically valid assessment techniques that have been applied in cross-cultural studies are also useful in the study of human development over the life span (Bronfenbrenner, 1979). Csikszentmihalyi & Larson (1984) have described the time use patterns and mental experience of adolescents. The reality of daily life has proved productive in describing the emotional turmoil of adolescents in what is considered a period of transition in human development. Altergott (1988) has used the daily life to understand the problems of the elderly. Her work especially focuses on the cross-cultural differences in the daily life experience in late life (Altergott & Duncan, 1988). Other authors have studied social relations over time and operationalized them as a rhythmic process to emphasize a continuous, smooth evolution (Warner, 1988). The process is not autonomous but evolves in interrelation between experience in environment — in which both the external and internal processes can take the lead to redefine the interaction — and the flow of experience. Kelly (1988) used the term “entrainment” — described in the biological rhythm discussion above — to describe this process of mutually exchanging meaning.

In the field of psychiatry the effect of environment and environmental change has primarily been studied in relation to the hospital environment (Goffman, 1961; Barton, 1976). Later studies have carried out long term behavioral observations of psychiatric symptomatology in hospital wards (Reynolds, 1965; Paul et al., 1987, 1988; McGuire & Polsky, 1979). Meanwhile, ecology has become a separate field in psychiatry (McGuire & Fairbanks, 1977; Glaeser, 1989; Kruse et al., 1990; Andresen et al., 1992; Tschacher et al., 1992). Ecologic studies of mental state related mood, behaviors and cognitions in non-hospital environments have not been performed to my knowledge.

FUNCTIONAL ANALYSIS: A MODEL TO INTERPRET DAILY LIFE EXPERIENCE

The “functional analysis” (FA) is the most important part of the behavioral assessment process (Kanfer & Phillips, 1970; Haynes & Wilson, 1979; Nelson & Hayes, 1986; Ciminero et al., 1986). It offers useful models for structuring and interpreting the concurrent observations of behaviors (including cognitions and emotions) and context. The monitoring techniques developed for this purpose assess the varying degrees of intensity of the symptom under study, concurrently with contextual information. Temporal and situational contiguity are crucial notions in this model to interpret observed differences in symptom intensity. In this section I will use the FA as a model to describe the hermeneutics of a process that intends to give meaning to the temporal and environmental variability in mental state.

The objective of the FA, which forms the basis of the behavioral assessment, is to describe the functional similarities in contexts and behaviors for a specific subject. The result of a FA is a set of salient contextual cues in which a target behavior occurs or a set of functionally similar behaviors that the subject performs in specific circumstances. A FA also assesses the results of specific behaviors and tries to understand the occurrence of those events through their effects. The functional similarities of behaviors and environments form the basis for planned change.

I will recapitulate the ingredients and interpretative structures of the FA-model using a non-traditional, cognitive operationalization. The traditional elements of the FA were derived in the experimental psychologic lab situation and are too reductionistic to be applied in real life situations. We think this modified operationalization facilitates more ecologically valid assessments of symptom variability in real life situations.

Introductory concepts

The basic reference scheme of the functional analysis is the “S-o-R-k-C”-scheme (Figure 2.1). The elements are [S] for “stimulus”, ‘o’ for “organism”, [R] for “response”, ‘k’ for “contingency” and [C] for “consequence”. The model is sometimes presented as the ABC-model in which the “stimulus” is called the “antecedent situation”, the “response” is the target “behavior”, and the “consequence” the result of it. I will now present some introductory principles that modify this scheme into a model that allows more ecologically valid assessments of symptoms in real life situations.

Figure 2.1 Basic reference scheme for behavioral assessment.

A stimulus is a complex situation - The circularity involved in the traditional concept of “stimulus” has been criticized by Gibson (1960, 1967). Because a subject responds systematically after its occurrence, we call an antecedent cue a “stimulus”. On the other hand, the concept “stimulus” assumes causality and “explains” why the consequent behavior occurs. The observational ground for defining a stimulus is also its explanation. The concept “stimulus” is adequately chosen for the laboratory situation where the experimenter selects and controls the antecedent situations, usually by concurrently avoiding the occurrence of other environmental eliciting cues. The traditional reductionistic concept of “stimulus” originated in this situation.

Contrary to the traditional representation and to match the ecologic reality, we conceptualize the “stimulus” as a complex situation in which different elements co-occur and in this complex situation, the concept of “stimulus” loses meaning. In the functional analysis elements of the complex situation gain meaning as the situational element (discriminatory stimuli — S^Δ) or cluster of co-occurring stimuli that eliciting behavior (Figure 2.2). Further functional similarities are derived from behavioral observations. The conceptualization of the antecedent stimulus configuration as a complex stimulus is not new. Pavlov (1927) already described the complexity of the stimulus situation:

“... that the animal at first generalizes any definite individual stimulus in the outer world, but that with repetition the stimulus becomes more and more specialized as a result of the development of an inhibitory process.” (p. 131)

In a review of Pavlov’s work, Pickenhain (1970) summarizes the stimulus situation in which conditioned reflexes occur as follows:

“As a rule, these conditioned reflexes are not formed to isolated, single stimuli, but rather to complex stimuli and to various stimulus relations.” (p. 126)

Figure 2.2 Stimuli are complex and cluster together when they acquire functional similarity through learning processes over time.

We will discuss the so-called learning principles later and describe how situational elements become cathected with meaning — acquire functional similarity — over time.

A stimulus is always “a stimulus for someone” (subjective) — Lang (1979, 1985) stresses the informational value of the stimuli. Contrary to the traditional behavioral assumption, the stimulus elements that constitute the complex situation are not primarily characterized by physical characteristics. Their relevance is related to the informational value for the subject. Part of this “information for the subject” can be reconstructed by observing behavioral changes over time. The functional or informational value is the result of the learning process, a result that can be checked by observing behavior. When we consider the complex stimulus in the antecedent situation, some elements are salient cues for that subject, while others are not. These can be differentiated by observing responses. This is not easy because often one functional set of elements is not enough to explain the emergence of behaviors.

The aim here is to understand individual adaptive reactions. To do this, we should try to understand the informational values of the stimulus elements and their interactions for that subject. Once again, we are fooled by the reductionistic experimental methods developed for the study of learning processes in the psychologic labs. The unraveling of complex situations for a concrete subject at a specific time is basically the core of the behavioral assessment in clinical practice.

Encoding occurs in a continual flow of observations — The laboratory situation repeatedly replicates the same S-R sequence, herewith maximizing and controlling the effect of this relation. As a consequence, we come to believe that in real life situations as well, the salient S-R sequences have no historical context. This is not true. Behaviors unfold in continuous sequences over time, and while we might think that a

subject systematically reacts to a specific stimulus, the complex environments at different moments in time are never the same. A closer look at the environments in which subjects behave, shows that behavioral consequences act as environments for the next series of behaviors as Watson (1924) already acknowledged. The understanding of functional relations in natural environments is only possible through the inspection of those sequences of behaviors. The observation of change over time from one moment to the next focuses the interpretations in the functional analysis.

Encoding of behaviors comprises multiple levels — The encoding of behaviors comprises multiple levels. “When an individual reacts to an object or situation, his whole body reacts. For us this means that manual organization, language organization (after it begins) and visceral organization all function together — each and every time the body reacts” (Watson, 1924, p. 252). It may come as a surprise that this point of view was part of the first operationalization of “behaviorism”. A more recent operationalization of this “whole body reaction” is the three-factor theory with behaviors (CAR — “conditioned avoidance reaction”), emotions (CER — “conditioned emotional reaction” including physiologic changes) and cognitions (COV — “coverants”) (see Mowrer, 1950; Dollard & Miller, 1950).

Because traditionally we are interested in describing learning processes the point of reference in the operationalization of the FA is the performing subject. We observe changes starting at a moment in time in which the target behavior is not yet performed and will assess “learning” as the change in behavior over time. In natural situations, however, we are seldom in a position to assess these learning processes because they usually started a long time before our observation. Due to the continuous flow of experience and behaviors we often only assess the (partial) result of learning processes. The functional analysis in an ecologic context has the assessor as the point of reference. Because the behaviors of the actors are only one of the things that happen, the difference between environment and actor vanishes. For the observing assessor, the behaviors of the subjects are part of his observational environment. The need to define “S-R”-schemata becomes less prominent. Instead, we assess “S-S”-schemata of which all the elements are simultaneously part: the stimulus and the three modalities of “behavior” (R) — behaviors, emotions and cognitions. Remember, a stimulus is a “stimulus for the subject” and learning processes change the meaning of those events for the subject in predictable ways.

Interpreting the time-interval between consecutive observations — It is now time to redefine the ‘o’ and ‘k’ factors in the traditional model. The traditional conceptualization of the ‘o’ factor is the “organism”. Processing of information within the subjects, unknown to the researcher, causes the emergence of the behavior in a non-deterministic way. The S does not always cause R. We have learned to interpret such an observation as a failure in the assessment model — we overlooked salient aspects of the complex S. It is, however, a myth to think that behaviors can always be predicted. On the contrary, random factors, impossible to control or to predict, influence the occurrence of R. A “black box”-term — modeled by the ‘o’ from “organism” — point to this indeterminism. “Organism”, however, is not the best term to describe the “black box” because sometimes there is no person — sometimes more

than one — between the consecutive observations of the complex S's (traditionally between S and R). Basically, the 'o' and 'k' (contingency) factors are the same. In the operant model, the 'k' factor is the 'o' of the therapist who rewards the subject at a specific rate. Both 'o' and 'k' are inserted in the model to stress that the observed sequence of complex stimuli in time is not deterministic. Both 'o' and 'k' can best be operationalized as "black boxes". No content but only the strength of the probabilistic relation is known. Moreover, in this operationalization content assumptions for the intermediate variables become part of the complex stimulus or the antecedent structure. When the subject has learned how to change an antecedent situation into a consequent one, the informational value of that stimulus element has changed for that subject. This is the meaning of the "situation for the subject". Cognitions (COV - coverants), emotions (CER - conditioned emotional responses) and the subject's information processing capacities, which are traditionally part of the 'o', are antecedents — therefore part of the complex S — because they are part of the context from which R's originate. According to Lang (1979) fear exists as an information structure (prototype) in memory that includes information about the feared stimuli, the fear responses and their meaning (see also Foa & Kozak, 1985). Accessing this fear memory-prototype, using whichever method is available, generates the behavior. The expectation of the consequent fear behavior changes the meaning of the stimulus for that subject. The antecedent structure integrates physical stimuli, known or learned emotional and cognitive changes, but also the habitual responses which are a consequence of the information processing capacities of the subjects (brain structure and function). The integration of the results of a learning process into the antecedent complex situation seems a little weird. However, for years rats have been deprived of food for prolonged periods of time before submitting them to learning tasks in a lab. Using such an approach we can be sure that the rats are hungry and ascertain that responses are triggered by the hunger drive. By analogy, each emotion (but also each cognition) can become part of the antecedent complex stimulus. The analysis of the learning history shows which cues have acquired informational value for a specific subject.

Causality assumptions are not allowed — The concurrent observation of behaviors and environmental elements does not warrant interpretations of causality. This is because we observe stimuli and responses concurrently over time. In a lab situation the experimenter knows which element he manipulates. As a consequence, there is experimental control over the events and the observed changes in the responses can be attributed to this manipulation. In natural environments stimuli occur in a predictable, meaningful way, often in response to or as an interaction with the subject's own experience. As a consequence, the environmental stimuli should not be interpreted as the cause of change. The subjects seek environments that match his mental state. The mental state is not caused by the specific environment. To make assessments that warrant causal interpretations we can use two different strategies. The first is a planned and experimental control of environments, the second is the randomization of situations within the normal flow of experience. Observations made during this kind of manipulation would yield more powerful control over the independent variable (see Campbell & Stanley, 1963; Kratochwill, 1978; Hersen & Barlow, 1976).

Summary — The model that will be used for the FA of R's — behaviors, cognitions and emotions — occurring in naturalistic environments is more complex (compare Figures 2.1 to 2.3). It shows the continuous flow of experience in interaction with the environment over time.

Figure 2.3 Behavioral assessment in complex situations including [S]timuli, [B]ehaviors (self and other-observed), [E]motions and [C]ognitions. Over time elements appear and disappear.

Depending on the discrimination required for our analysis, slices of time with different intervals can be selected. We can argue that the assessment of emotions needs smaller intervals while behaviors need larger ones. As a general rule the observational density (frequency) should result in at least two assessments for each process under study (Gottman, 1981). Often, repeated observations of the same process in comparable situations will not be perfect replications. Therefore a so-called “black box” symbolized by percentage signs is inserted between each time slice. They replace the “organism” and “contingency” of the traditional FA-operationalization. In this model, everything that is known or can be derived from the observations is part of the observation at the moment. For instance, results from learning processes are part of the stimulus situation, because the “stimulus for the subject” changes during learning.

Learning principles

Behavioral analysis is based on the observation of changing complex stimuli over time. Some moments are conceptualized as antecedent stimuli, some as behaviors or responses and some as consequences. For the interpretations we center our attention on those parts of the complex stimulus that change over time. The complex situation contains the outer and known inner environment (cognitions and emotions) of the individual. The functional analysis selects some elements in the continual evolution of complex situations over time as cause of the change (behaviors or events). To understand which contingencies are important for a subject, we have to know how environmental elements become meaningful for a subject over time. Learning theory explains this.

Learning the meaning of stimuli — The study of conditioned reflexes has evolved from the observation of dogs in the natural environment to the controlled laboratory situation. The original observation was that the salivary excretion (UCS ¹⁵), which is

¹⁵ UCS: “unconditioned stimulus”, or the stimulus with meaning before the learning process starts;

expected in the sight of food, was also observed in an unexpected antecedent (CS): the bell signal. The principle behind the realization of this learning involves the manipulation of the contingent occurrence of both stimulus elements. The traditional presentation of this is an “UCS - CS” association. Those associations are often weak in natural environments (which means that there is a low frequency of co-occurrence) but are maximized in the lab, to optimize the learning speed. To check if learning occurred, we observe behavioral changes: the UCR, which normally follows the UCS, follows also on the CS, but in a slightly modified form as CR. The emergence of the CR, means that the CS has received signal value for the UCS. The information value of the UCS was carried over to the CS. Later research has shown that this behavioral change cannot be universally realized. Some stimuli can be associated more easily than others (the notion of “preparedness” (see Merckelbach, 1989)).

Figure 2.4a Traditional scheme for “stimulus”-learning (classical conditioning).

Unknown elements of a complex situation acquire meaning from a known stimulus through association. When a unknown new stimulus element emerges or a set of stimulus elements tend to co-occur in the complex situation the observer starts a scanning process. The known meaning of the old associated elements transfers to the new stimulus. This process is described as “classical learning” and is triggered by the arousal which is the consequence of the observation of change in the complex environment. Sometimes, however, the scanning process does not result in a match with a co-occurring event with known meaning. The new elements can then not be associated with anything and will not change their original meaning or may even acquire no meaning at all. In this situation the original arousal diminishes and the subject learns that the stimulus element is unimportant. This process is called “habituation”.

Figure 2.4b Modified scheme for “stimulus”-learning (classical conditioning).

CS: “conditioned stimulus”, or the stimulus gains meaning through the learning process;
 UCR: “unconditioned response”, or the response that followed the UCS before the learning process;
 CR: “conditioned response”, or the response that follows the CS after the learning process.

Although co-occurrence in time is a powerful way to form those associations, it is not the only one. The whole spectrum of characteristics of the stimulus element (such as color, form, similarity, functionality) can be used to facilitate the learning of meanings of the previously unknown stimulus elements. The cues that have acquired the same meaning are called “functional identical”.

Learning behaviors — While habituation and classical learning are related to the learning of the meaning of stimuli based on the principle of contingency, operant learning is related to the learning of behaviors. Subjects learn behaviors through the evaluation of the difference between the antecedent and consequent complex environments and the assumption that the behavior or behavioral sequence observed in the intermittent time has caused this change. The behaviors that are learned are purposeful behaviors, i.e., strategies to change the environment purposefully from one moment to the other. The learning principle is operant learning. The law of effect predicts that the frequency of purposeful behaviors will rise, while the expression of ineffective behaviors will reduce over time.

Figure 2.5 Modified scheme for “response”-learning (operant conditioning).

The study of operant change, as it is conducted in the lab, has been reduced to a study of experimentally induced behavioral change. Moreover, the experimentally induced environmental changes following the behavioral performance of the subject in the lab were usually extreme forms of clear-cut consequences: rewards and punishments. In this very structured situation, the researcher has to deprive the subject in order to be sure of the direction of the behavioral change.

In the natural environment, where the complex ecologic situation creates complex antecedent and consequent situations, behavioral scientists have learned that universal rewards and punishments do not exist. Behavioral changes do not always evolve in the expected direction. To change behaviors you have to manipulate their consequences. However, even the global definition of these consequences as punishments or rewards, cannot be predicted in advance, but have to be derived from the observed behavior change. This is a circular definition. From a cognitive point of view we can understand this process. Consequences are situational elements. As with all situational elements their meanings change from one moment to the next and from one subject to another. As a consequence a poor prediction of meaning is no surprise.

Considering the difficulties in defining rewards and punishments in controlled situations, the meaning of context and its importance for understanding behaviors is

even more confusing in ecologic situations. To avoid the unpredictability of the complex stimulus environment as a rewarding or punishing consequence, operant treatment techniques have introduced methods to maximize control in planned behavioral change. Token Economies, for instance, are techniques that introduce an artificial generalized reinforcer — such as money — to ascertain behavioral change. This procedure, however, is of little value in an ecologic environment because behavioral changes that are realized that way do not generalize well after token withdrawal. The poor generalization can be understood because behaviors are purposeful strategies to change the environment over time. The change strategy that is learned by the subject in a Token Economy is not the purposeful behaviors but is how to earn tokens.

We know that chronic mental patients and schizophrenic subjects in particular are poor assessors of the consequences of their behaviors. The task is not an easy one: environmental consequences occur irregularly and with low frequency in the natural environment. Patients with grooming problems, for instance, have only a marginal chance at picking up the fact that people avoid them, and relating this information to their own poor grooming behavior is even more unlikely.

Generalizations in Token Economies should start from the natural consequences of behaviors and ascertain that these consequences — which are originally meaningless for the patient — acquire meaning through a systematic linkage with a known and valued stimulus element (the token). The meaningless natural consequence of the behavior will acquire meaning (a classic learning procedure) and act as reinforcer — reason to perform behaviors — on its own, without the help of tokens.

Causal assessments in complex stimuli

Referring back to the model presented in Figure 2.3, we can observe and define different moments of the evolving complex environment and make interpretations based on changes occurring between the repeating observations over time. In this section I present a taxonomy of the most important sequences of observations and the relation of these sequences to the interpretation of change over time in research and clinical practice. At each time moment we assess the complex situation and compare this with previous and consecutive moments in order to define the aspects of the complex stimulus that changed and might play a role in understanding the behavioral evolution. We will present models to interpret the differences that can be observed.

Antecedence

The first model that explains changes in situations over time is the “antecedence”. In this situation the occurrence of a target event at T_i is preceded and understood by a prior event (T_{i-1}). Conclusions are derived from the fact that a specific characteristic of the antecedent situation has a higher probability of occurring before a target event than do other concurrent antecedent characteristics. In practice, however, it will often happen that different situational cues co-occur. In addition, the frequency of occurrence of target events in natural situations is usually low making an appropriate assessment and discrimination problematic. Experimental manipulation of the events or antecedent situations can be helpful to generate enough observations to allow

reliable, assessments of the salient antecedent situation as well as to break the natural co-occurrence of cues, adding experimental control and explanatory power.

An example can illustrate how in natural situations a number of salient situational characteristics co-occur. Assuming we assess a subject at random chances are that all the not-alone observations will be with a particular person at a specific place and time of the day. For instance, an unemployed married woman is alone from 9 to 5 on working days but not in the evening hours when her husband is home. In order to be able to draw conclusions or to enhance the credibility of the assumptions — in our example: does the time of the day or the presence of her husband explain a change in mood — it is important to disconnect the natural flow of events. Two options are possible: the experimental occurrence of an event at prescribed times or at random times. In both cases neither the subject nor his environment selects the occurrence of the event, and thus the occurrence of those events are uncorrelated — and therefore discernible — from other situational aspects. Such a procedure, however, comes at additional costs for the ecological validity.

Interestingly, we know that co-occurring situational cues will acquire the same meaning over time. This is what the classical learning paradigm teaches us. Therefore, we should not be too concerned about the co-occurrence of antecedent situational terms. Co-occurrence in daily life makes it unlikely that the meaning, from a behavioral and emotional point of view, will be different. The whole set of stimulus elements that co-occur in a context will result in similar behavioral, cognitive and emotional consequences.

Consequence $T_i \rightarrow T_{i+1}$

The “consequence”-model assesses the environmental results or consequences of a target event. We focus our attention on the changes in the complex stimulus after the occurrence of a target event. The purposefulness of the target event is derived from its consequent situation. The same remarks made in the previous paragraph apply to the “consequence” situation. We cannot derive causal conclusions from the observation of consequent changes as a result of the target event. Here, too, hidden unrelated motives can cause the occurrence of events. For instance, a schizophrenic subject can choose to withdraw to a quiet environment when he is troubled by hallucinations. When the hallucination intensity drops, we can conclude it is an effective coping strategy. However, chances are that nothing else was possible for the subject or that he was commanded by his voices to leave the room.

The “consequence” situation mirrors the operant learning paradigm. The occurrence frequency of the target event (T_i) will increase or decrease over time, depending on the purposefulness of the environmental change. This environmental change will be discussed below.

Process $T_{i-1} \rightarrow T_i \rightarrow T_{i+1}$

When assessing a “process”, we compare antecedent (T_{i-1}) and consequent (T_{i+1}) situations, relating the changes between both observed moments to the intermediate events (T_i). As was the case in the “antecedent” and “consequent” models, we caution

against causal interpretations of co-occurring situations in natural environments. Causal interpretations are restricted to independently occurring events. In the natural flow of daily life this independence is not guaranteed.

The assessment of a “process” relates the consequent situation ($T_i \rightarrow T_{i+1}$) to a specific, discriminatory antecedent situation (T_{i-1}). A process contains therefore more richness than the consequent model. Even so, the antecedent model ($T_{i-1} \rightarrow T_i$) gains explanatory power when it is followed by specific events (T_{i+1}). Those events are crucial to the understanding of the role of the behaviors, cognitions and emotions that occur between antecedent and consequent moments and to the creation of the situational change ($T_{i-1} \rightarrow T_{i+1}$). The operant paradigm teaches us that the observation of high frequency behaviors should be related to the perceived purposefulness for the acting subject of the situational change. Emerging cognitions and emotions reflect the signal value of the antecedent situation and/or the anticipation of what will come for the subject. Lasting, low frequency, events on the contrary, are a result of the consequences not being perceived as purposeful by the subject. For a chronic mental patient who stays in his house all the time, being home should be considered rewarding. Equally so, when a chronic mental patient repeatedly forgets to take his medication, we should conclude that taking medication is not considered positively.

Concurrence

The traditional behavioral assessment model (S-o-R-k-C) assumes a linear temporal pattern in the antecedent stimulus. The antecedent situation is past history when the target behavior is performed. Most of the laboratory operationalizations are bells or lights that function as signals for the appropriate response. The word “stimulus” suggests this discrete property of the antecedent event. The antecedent situation, once emitted, stimulates the organism to perform a response.

In our conceptualization the antecedent situation is a complex stimulus. Aspects of the situation present themselves in the natural flow of life and are usually active for longer periods of time. Basically, there is no difference between the antecedent, behavior and consequent situations, because due to the time pattern of each of these elements, they all can be present at the same time while there is no way to know what came first. As a consequence, when we make interpointions — time slices — in this continuous flow of perceived and experienced situational aspects, often the target antecedent stimuli, the behaviors and consequences are confusingly intermingled. This is the situation of concurrent assessment.

An example can illustrate this. When we observe a subject in the company of friends in a bar, feeling happy, those elements are observed concurrently at a specific moment in time. We can hypothesize that being with friends in a leisure context is rewarding for that individual, but we cannot prove that. The traditional operationalization would define the company and the bar as the discriminating stimulus, the target subject attending the bar as the acting individual and the resulting happy feeling as the consequence. However, all observational elements of the S-o-R-k-C model are collapsed in our example — and often in daily life — into one observation.

Interpretations cannot be drawn easily from this situation because it is very difficult to reconstruct the process of change in the complex situation in a concurrent analysis — define what is antecedent, behavior and consequence — and therefore to draw conclusions. The problem is not restricted to this model only. None of the other assessment strategies described in this taxonomy has experimental control over the salient and possibly causal aspects of the complex stimulus. Therefore the analysis of the changing complex situation does not allow causal interpretations.

In our example, the explanation would gain power if we had observed the subject feeling less happy in a non-social situation before the visit to the bar or, even more, if the sequence of observations showed happy feelings increasing in the bar when specific individuals joined the party.

Rhythmicity $T_1 \rightarrow \dots \rightarrow T_{i-1} \rightarrow T_i \rightarrow T_{i+} \rightarrow \dots \rightarrow T_n$

The last model describes the evolution of the process over time, for which time itself is the explaining variable. This situation applies to the assessment of diurnal rhythms or the continuous monitoring of events over time (recovery curves, for instance). The same processes apply to this situation, except that in the study of rhythmicity, we do not restrict ourselves to data collected over one, two or three moments.

Kelly & McGrath (1988) have tried to define different options for the temporal shape of changes caused by events: a) a gradually increasing linear process; b) an all-at-once change that is then maintained over time; c) an all-at-once change that is not maintained over time; d) a delayed effect; and e) a cyclic effect. More complex taxonomies are also available (Gottman, 1981; Bakeman & Gottman, 1986). The interpretation of change over time can differ in relation to the number of observations involved. Because two points always describe a line we need at least one extra third observation to confirm a linear trend in the data. That additional observation can be on or off the line and confirm the model. When a linear trend is not assumed, the assessment becomes even more difficult. As a general rule we can say that a model with the same power — in mathematical terms — as the number of observations, always fits the data. To have explanatory power, the model should at least have one degree less, than the number of observations: 3 data points for a linear trend, 4 for a quadratic (parabolic) equation, etc. In the fortunate situation where we have a large number of observations, we can model cyclic processes. In this situation we can only describe those cycles, for which we have at least observations for two full periods and with a periodicity of at least 2 times the between-observation interval. To make this more concrete: to observe a diurnal rhythm we should have at least two observations per day and observe for at least two days. Equally so, when we observe once each hour it is impossible to draw conclusions on the mood effects of specific television programs, for instance, because they change within that time frame. We need more information to model these effects.

Once again it is important to realize that observations in natural environments, while high in realism, tend to be low in control (Kelly & McGrath, 1988; Chow, 1987). Diurnal cycles in observations can be interpreted as results of processes ruled by biologic clocks, but other non-biologic factors also influence the changes in the processes under study. Sometimes the “biologic rhythm” needs those external

influences — Zeitgebers — to recalibrate or reset (“entrain”) the process at hand (see Daan & Gwinner, 1989; Binkley, 1990). Influences that have the potential to entrain the process, and which evolve over time in a systematic way correlated to the rhythm, are especially difficult to discern. Monk et al. (1990, 1991) have studied the co-occurrence of social Zeitgebers in the diurnal rhythm of depressive subjects.

SUMMARY AND DISCUSSION

We took a long way around to present the building blocks for a new assessment strategy in psychiatry. Our starting point was that traditional assessments of schizophrenia lack the appropriate sensitivity and specificity to describe the heterogeneity of the disease adequately. We argued that — since the 1920s — “behaviorism” has created a “missing link” in psychiatric research. Often inappropriate and reductionistic indicators of psychic symptomatology were propagated. We invested primarily in so called objective indicators — often high in reliability — but with a limited validity to describe the actual phenomenology of the illness. Of course, objective indicators are important for the study of etiology, but with only poor descriptors of the actual adaptational processes that form the illness experience, the study of etiologic factors is significantly handicapped. Pain is also an important indicator in internal medicine.

The building blocks for the new assessment strategy developed in this chapter are based on three different questions. What information do we need to anchor our assessments of the symptomatology of mental illness (*content of assessment*)? What additional information do we need to understand and interpret this content (*context of assessment*)? And finally, how can we synthesize our observations into a meaningful interpretation?

We argued that the most adequate place to look for mental illness is in subjective experiences. One might even say that mental illnesses *are* subjective experiences and the adaptational processes that are reflected in them. Even objective — biological — indicators have to be validated against these. Consequently, the content of the assessments should be based primarily on these subjective experiences. Furthermore, mental illness processes occur in a context and are often best assessed through the (mal)-adaptive interactions of the subjects with their environments. Therefore, a dynamic assessment that monitors person-environment interactions is important for the evaluation of mental illness. To allow such an appraisal of contextual variability we need a dimensional — not a categorical — model of illness and to focus on symptoms — not on syndromes. The repeated assessment of emotions yields important indicators of the self-perception of the person/environment adaptation that is related to mental illness processes. “Time” and “place” are the (objective) contextual dimensions in which the emotions occur. Meaning for the subjects can be found by integrating activities and persons involved in the interpretation of the emotions within these dimensions.

There are no general rules to decide what a context means for a subject and consequently it is difficult to interpret their adaptational strategies as reflections of illness or well-being. The theoretical learning principles were discussed because they shed some light on how these processes work. In illness and well-being the learning

processes regulate how unimportant environments become meaningful for a subject. The meaning of the environment for a subject is the base for the understanding of the emotional responses and processes that form the base of the mental illness. Hermeneutics rules were finally presented. These allow interpretations of our observations of the person/environmental interactions.

The integration of these building blocks for innovative assessments in psychiatry into a useful assessment technique is the target of Part 2. In the following chapters I will present the “Experience Sampling Method” (ESM) that we have developed as a solution for the lack of validity of traditional assessment instruments in mental health.

Part 2: The Experience Sampling Method

Chapter 3: Self-Observed Mental State and “Time Sampling”¹⁶

INTRODUCTION

Current diagnostic practice in psychiatry

Over the last two decades, the development of a number of diagnostic instruments such as the DSM IV and the ICD 10, as well as various psychometric instruments to describe cognitive, affective and social processes related to psychopathology, has improved diagnostic precision in psychiatry. Researchers became aware that psychopathologic phenomena are complex and that expansion of our knowledge is only possible if scientific conclusions are integrated on an international scale. The study of psychopathology based on the current diagnostic development in psychiatry suffers from three important problems:

- *Deterioration of taxonomy*: Traditional Western medicine, from Greek antiquity to Renaissance, consisted of the study of man in his natural environment. To accelerate our acquisition of knowledge about psychopathology and to improve the efficacy of interventions, modern medicine restricted its area of activity to anatomy and physiology. At this moment, the analysis of cells, molecules, genes, and atoms determines the medical scientific domain. While this reduction to its essence has proved fruitful in different fields of medicine, we think that in psychiatry, the limited diagnostic taxonomy of modern medicine misses the essence of the illness. The phenomenology of the psychiatric illness experience — reflected by the reality and variability of the daily experience and environmental embeddedness — is very distant from its current diagnostic descriptions. This distance should be reduced.
- *Unsuitable models*: In somatic medicine, physiology — the study of healthy processes — is the basic discipline for pathology. Psychology should offer psychiatry such a basis. However, until now psychology has not been able to develop a suitable model for the healthy individual and in particular for the variations of experiences in daily life. So psychiatry does not dispose of a theoretical discipline to develop the study of abnormal consciousness in daily life or the fluctuations in time between optimal and illness experiences. A model for a “psychology of the healthy person”, while still unavailable, is currently under development in social and developmental psychology (Argyle, 1992; Csikszentmihalyi, 1990; 1993).
- *Failing methodologies*: When we use our traditional research methods, many pathologic experiences will not be discovered. Significant problems in mental health care are found in groups that live at a certain distance of the medical world (“hidden” populations; Costello, 1993): e.g. drug users, agoraphobic women, homeless

¹⁶ An abbreviated version of this chapter was originally published as: deVries, M. W. & Delespaul, Ph. A. E. G. (1993). Zelfobservatie en Intensieve Tijdsampling in de Psychiatrie. In C. A. L. Hoogduin, P. Schnabel, W. Vandereycken, K. van der Velden & F. C. Verhulst (Eds.), *Jaarboek Psychiatrie en Psychotherapie* (pp. 94-108). Meppel: Van Loghum Slaterus.

schizophrenics and eccentric marginals. The risk behavior and suffering of mental patients does not occur in the consultation room but mainly in their daily life and is often experienced in low frequencies, invisible to others ("hidden" experiences). Because of our unreliable diagnostic instruments, panic frequencies are often overestimated, whereas accurate descriptions of the experience of hallucinations and delusions are rare.

The above-mentioned problems force us to reconsider our research practice in psychiatry and to incorporate more adequate assessment strategies that include descriptions of the experience of illness (symptoms, behavior, and context). The proof of the pudding is in the eating. Our diagnostic system should become a useful tool for the selection between medical treatment options. This requires an assessment method that enables us to observe the occurrence, variability and development of symptoms and behaviors over time. Until recently, this kind of information was not available for research. We will show that a tighter description of daily life functioning including environmental adaptation yields data that is "fundamentally" different from the data obtained through traditional diagnostic methods. Interpretation and use of these methods require a shift in our scientific orientation from causal interpretations to processes and contextually oriented explanations. This way the gap between the diagnostic system and clinical decision making can be reduced.

Questions in psychiatry

Mental health workers need information which they can use for clinical decision making. What is the appropriate intervention? Should the patient be referred? Is it useful to give medication and, if so, what kind of medication? What data do we need to make these decisions? In traditional psychiatric assessment we can distinguish three different kinds of information (Delespaul, 1992):

- *diagnosis*: classical assessment strategies place a subject within a reference group. Previous knowledge about that reference group can be applied to the subject and prognostic statements about mental state, course and therapeutic success can be made. This makes diagnosis important.
- *symptom (co-)occurrence, frequency and severity*: it is often unclear what really happens when someone says he has a "panic attack", "migraine headache" or a "psychotic breakdown". It is easier to have enough information to decide on a specific diagnosis than it is to ascertain how often a specific symptom occurs or set of symptoms co-occur or how impairing the symptoms are. This information is important for deciding where to start the interventions .
- *etiology and course*: assessments of the distribution of symptoms over time, as well as descriptions of the context in which symptoms occur with higher or lower intensity, are necessary for optimizing interventions. With inadequate symptom monitoring it is difficult to know if we can find specific situations in which symptoms emerge and/or in which the subject recovers from previously emerged symptoms. We need to know how those processes dynamically evolve over time.

Methods in psychiatry

We usually get our information through clinical interviews with the patient, relevant others, such as family members, and from the referrers. The quality of the data collection can be improved by conducting the assessment interview more systematically. In this way, “white spots” in the symptom complex can be avoided and differential diagnostics can be optimized. Systematized psychiatric interviews (PSE, DIS, SCAN, SCID, CIDI) contain additional decision rules, determined by experts or based on empirical research, in order to derive syndromes from sets of symptoms. The psychiatric interview is the most prominent technique to acquire information about the patients and the way the patients and persons in the patient’s environment perceive the problems. This narrative information has direct consequences for psychotherapy.

Clinical interviews and questionnaires, the second most important assessment method in psychiatry, are much alike. In the construction of a questionnaire we go to a lot of trouble to select and formulate items appropriately. In this way we are able, for instance, to distinguish between habitual symptoms (the “trait”: “I’m just like that” - “That’s how I usually react”) and time-limited symptoms (the “state”: “Now, the last few days, I’m suffering from...”). Thanks to modern psychometric theory we can determine the location of individual items on a scale or latent construct. For example, if someone says he’s feeling exhausted, there is a specific chance that he is depressive. On the basis of statistical principles, the informative value of the items can be optimized. Questionnaires are inexpensive tools to get reliable and valid information quickly. They are especially suited for screening. Table 3.1 relates questions and methods together.

	Questions and Assessment Objectives	Methods	
One-Time Measures Repeated Measures	Syndrome	<ul style="list-style-type: none"> • Clinical interviews • Standardized questionnaires 	
	Symptom • occurrence (frequency) • severity	<ul style="list-style-type: none"> • Repeated retrospective event contingent assessments (e.g. each depressive episode in the last 5 years) • Repeated retrospective time contingent assessments (e.g. end of day ratings) 	Retrospective Assessment
	Etiology and Course Δ intensity in relation to: • environmental; and • biological changes	<ul style="list-style-type: none"> • Repeated prospective event contingent assessments (e.g. each panic attack for the next three weeks) • Repeated prospective time contingent assessments (e.g. 'Experience Sampling Method' (ESM)) 	Prospective Assessment

Table 3.1 Integrating questions and methods in psychiatry.

The first level summarizes the questions that can be responded to by means of instruments with a once-only or retrospective character, such as an interview and a questionnaire. To get a better insight into the self-image of the patient, a clinical interview can be used. For a global overview of the symptoms and for the placing of subjects in an reference group, structured interviews or psychometric questionnaires are appropriate. The second level contains retrospective, but repeated, measures. With

this kind of method, such as the “end-of-the-day” measures, the frequency and the distribution of the symptoms can be assessed when they are not evident. For example, did a panic or migraine attack occur today or not? The third level contains the instruments for repeated measures that limit the observations to the moment (prospective). These methods are very useful when the definition of the symptom, the context, the frequency and the time pattern of the symptoms are not clear.

Quality of the collected information

Both structured interviews and questionnaires are restricted in their applicability to individual cases and therefore for customizing interventions. This is a result of the probabilistic rules that are part of the psychometric theories on which the development of those instruments is based. Summary scores are computed on the answers that are collected on individual subjects. We compare the reduced summary score of individual subjects with reference material collected from previous subjects that share that same score and for which more data, for instance on course or concurrent symptomatology, is available. By extrapolating from the appropriate reference group, this procedure allows us to predict current and future behaviors, cognitions and emotions of individual subjects. However, this reductionism comes with a cost: psychiatric interviews and questionnaires lose ecologic and member validity. For this reason, they are less appropriate for assessing important questions in psychiatry: the presence of specific symptoms, the context in which they occur and their development over time.

A second drawback of traditional assessment procedures is that interviews and questionnaires are based on the assumption that the informant, the patient, the family or referrers dispose of the required information and that they can report it in a reliable way. Unfortunately, this is not the case. An example illustrates this. DSM-III-R defines a panic disorder as having at least four panic attacks — well-defined periods of intense fear or being unwell, characterized by at least four symptoms out of a list of thirteen — over the past four weeks¹⁷. Using cueing questions we ask the subject to remember the time as well as the symptoms of his last panic attacks, a virtually impossible task. As a test, let us verify how many times we took a day off in the past four weeks. Considering that a day off is often planned in advance and on meaningful dates, it is surprising that we still have to look in our diary to answer this question. It is even more difficult to answer the question: “How did you feel on those days?”, which would be relevant information for psychiatric diagnostics. Following Wheeler and Reis (1991) we summarize the problems with reporting mood and psychopathology reliably in global assessments such as interviews and questionnaires as follows:

- The presence and intensity of psychiatric symptomatology *varies over time and over situations*. Global assessments do not capture this useful variability.
- Descriptions of concrete moments from the (recent) past require an extremely difficult *cognitive process*.
 - *SELECTION*: is it clear for the patient in what feelings and behavior the diagnostician is interested? Even given with specific instructions, the patient can have

¹⁷ This pseudo-objective assessment was dropped in DSM-IV (APA, 1994).

a personal interpretation of certain concepts (migraine, panic, anxiety, depressed, ...) and as a consequence report non-target experiences;

- *MEMORY*: it may be possible for a patient to recall a certain event but, it seems, it is often impossible to remember the exact moment in time, the current context, the present mood, the antecedents and consequences;
- *AGGREGATION*: how do we summarize characteristics of recollected events? Usually this is not a mathematically sound aggregation, but an evaluation and weighting, influenced by many factors.
- Reports of events are likely to be influenced by *aspects of the context* of the moment at which it is remembered; e.g. it is easier to recall situations that occurred recently, in the same season, or in the same social context.
- *Motivational and emotional processes* also play an important role in reporting. People who feel depressive for a longer period of time will color their descriptions of more positive events from the past depressively.

In summary, reports of mental state are rarely accurate, showing quantitative and qualitative differences between the actual experiences and the report. These subjective differences are the basis for psychotherapy, but for research and diagnostics they are a source of error. For research, a detailed description of the variability and contextual relation of symptoms, mood and behavior is necessary. This is also true in behavior therapy where the current context of the symptom behavior, its antecedents and consequences should be described in detail (Delespaul & deVries, 1991).

To optimize psychiatric practice, a diagnostic instrument able to detect small changes in cognitions, moods and behaviors is necessary. The same goes more generally for psychiatric research. To realize this, unreliability should be avoided. Therefore, we decided to minimize selection bias, memory effects and incorrect aggregation by using a repeated observations strategy in which only the actual mental state is to be reported. This will be described in the next section.

PROPOSED CHANGES IN DIAGNOSTIC PRACTICE

Intensive observations in daily life

The social sciences have developed intensive observation techniques that better and more directly approach populations and describe the subjective context of syndromes. From a long tradition of observation techniques, we mention the ethnographies of Malinowski (1935), the more systematic application of time observation techniques in naturalistic field studies (Chapple, 1970; Munroe & Munroe, 1971a,b; Barker, 1978), also applied in psychiatry (Reynolds, 1965; McGuire & Polsky, 1979) and the behavior monitoring techniques (Nelson, 1977).

The methods we discuss here have further been made possible by a renewed interest in subjective experiences in psychopathology in a neo- or post-behavioristic climate (Strauss, 1989). The promise is that a growing knowledge of our patients, the occurrence of symptoms and how these phenomena affect them — not what they think caused them — will give us a better insight into the different psychiatric

syndromes. After all, people can observe and report on themselves, under realistic circumstances, about places and moments where the health care worker or researcher is not present (ecologic validity: Bronfenbrenner, 1979). These descriptions are real experiences of the subject and not, as in traditional methods, the classification of individuals within a reference group used to extrapolate to individual behaviors and symptoms. As a consequence, patients can better recognize themselves in clinical and scientific descriptions (member validity: Mehan & Wood, 1975).

Diary methods and more recently structured methods for repeated observations provide a solution to the lack of ecologic and member validity in the study of mentally disturbed behavior. In a quantitative and repeatable way, these methods give a good impression of psychopathology with the necessary situational and temporal detail. Their growing popularity in behavioral research has been demonstrated by a large number of recently published scientific articles, special issues of journals and books (Gross, 1984; Pervin, 1985; deVries, 1987, 1992; Tennen et al., 1991).

Temporal approaches to psychopathology

Historically, psychiatry has not been blind to the influence of temporal and situational factors on mental disorders. They were recognized in the late 19th century by Kahlbaum, Kraepelin and Bleuler in their detailed descriptions of schizophrenia and today are the focus of study in life history and life event research. This work has resulted in an appreciation of the long-term outcome of mental illness and a better understanding of relapse-precipitating and recovery-enhancing factors in the episodic course of the illness. Much work, however, has still to be done.

Another important influence has come from research on circadian rhythms. This research has evolved from the early work of Wada (1922), who described rhythmicity in gastric motility, to the discovery and detailed exploration of cyclical rapid eye movements in different stages of sleep (Kleitman, 1963; Kripke, 1983). These investigators have sought a comparable rhythmicity affecting cycles of arousal and behavioral activity in the waking person (Minors & Waterhouse, 1981; Reynolds, 1965). This line of research has highlighted the profound influence of physical "Zeitgebers" (daily environmental time setters (entrainment) such as light; see Binkley, 1990; Daan & Gwinner, 1989) and setting influences (social "Zeitgebers": Ehlers et al., 1988; Kupfer et al., 1990; Monk et al., 1990, 1991; Prigerson et al., 1994) on a variety of biologic and behavioral measures.

DATA COLLECTION BY INTENSIVE REPEATED OBSERVATION

Advantages of intensive field methods

The most important motive for choosing a data collection with repeated measures is a statistical one. Repeated assessments of the same phenomenon reduce the measurement errors of the summarizing "statistics". The more you measure, the more reliably you can describe the variability in the phenomenon under study. This is, however, not

an automatic result. Accuracy depends on the rationale that is used to collect the repeated observations.

What does this mean in practice? A case manager worries that an ambulatory chronic mental patient spends a lot of time alone in his flat and wonders how the patient feels, especially in the evening. He is not interested in what the subject thinks about this, but in the actual feelings and cognitions in that concrete situation, because he fears that negative feelings on those moments might drive the patient to drink. As we have already demonstrated, it is unlikely that the patient can give a reliable answer to this question in a clinical interview. The case manager chooses a repeated measures strategy and “draws a sample of moments” from the patient’s daily life. He compares assessments during moments alone at home in the evening with a random sample of the habitual moods and cognitions, determined through the analysis of non-target moments. If only a few observations are available or if the subject uses observations at other specific moments for comparison, coincidental factors, such as a nice program on TV or an argument at work, are likely to influence the results. The case manager therefore tries to maximize the observations in both situations, using a sample of observations that is as reliable as possible.

Behaviors, moods, and cognitions are variable and this variability can only be assessed by measuring repeatedly. Often we are not interested in the average behavior, the habitual mood, or the most typical cognition, but in their variability from moment to moment and from situation to situation. In contrast with traditional statistical techniques — which reject variability as annoying (error) — this variability is useful information for the clinician and offers clues for therapy.

Time and event sampling methods

Repeated observations should be systematized and the collection of data should be cued by a specific rule. Two different approaches exist. The first is based on the occurrence of a small set of preselected events and is called “event sampling”. The alternative uses a time based rule to cue the data collection and is called “time sampling”. Event and time sampling procedures differ in their applicability in research practice (Wheeler & Reis, 1991; Suls & Martin, 1993).

- *Event sampling:* In event sampling observing is done contingent on the occurrence of a target event, for instance, each time the subject meets people or each time he smokes a cigarette. In psychiatry we can instruct subjects to report each panic attack, each time an anorexic patient vomits or each outburst of anger but it is very difficult to ask subjects to report varying intensities of psychotic symptomatology. Observations contingent on events can only be used when the problem behavior is well-defined (countable and with a clear beginning and end) and restricted (without changing salient characteristics over time). Unfortunately, the use of event sampling can fool us into thinking that situations that occur in varying intensity are in fact discrete events. Furthermore, because the reports are often written down *after* the event, retrospective bias can creep in. This is always the case for the descriptions of the antecedent situations which led to the event. Those descriptions will always be colored by the knowledge of the occurrence of the event, and because

we do not know whether these antecedent situations would also have occurred if not followed by the target event, we might be on the wrong track. Therefore, event sampling is often restricted to traditional monitoring (only record the occurrence of discrete events). In this situation it is a cost-effective alternative to collect valuable information prospectively.

- *Time sampling*: In time sampling the observation is contingent on a time schedule and the subject is asked to describe the current situation or the situation in a fixed preceding or consecutive period. For instance we can collect observations each hour, or at the end of each day. Methodologically, this is a very elegant measuring strategy, because the time frame of each observation is well-defined. This allows more control in the assessment of the duration of behaviors. Moreover, while event sampling focuses on a well-defined target event, time sampling does not need this. Time sampling allows multiple targets, which mirrors the psychiatric symptoms more appropriately. Also, because in time sampling the assessment is independent of the occurrence of the events, we can collect unbiased observations of antecedents. A disadvantage of the time sampling strategy is that many relevant events can get lost, especially when they occur infrequently and/or are of short duration (Mann et al., 1991). We distinguish between two time sampling procedures:

- *FIXED* interval time sampling: Sophisticated statistical techniques, such as the time series analyses and the Markov chains, assume that the data are at *equal intermediate intervals* (fixed). This, however, makes the observations predictable for the subject. Chances are that he will change his behaviors and that his mood and thoughts will be influenced by the anticipation of the next observation. As a consequence, the sampled moments are no longer representative for the process under study. These threats to an unbiased assessment are always hidden in a fixed sampling. Assume we observe every 60 minutes at the full hour. The moment of the radio news will more frequently be sampled and as a consequence the occurrence of specific behaviors (listening to the radio), feelings (to war and violence), and possibly social interactions (the family together in one room) are not representative of the subject's living conditions.

While fixed interval strategies allow interesting statistical methodologies, they are difficult in psychiatry because conscious processes, necessary for self-observations, are only possible when the patient is awake. The result is a night break that disturbs the fixed interval structure anyway and makes the application of time related statistical techniques problematic anyway.

- *RANDOM* interval time sampling: In random time sampling procedures the monitoring of target assessments is *contingent on a random signal*. Consequently, intervals have a random length and if we ascertain that the signals come unexpectedly, the subjects cannot control or know the time at which the observation has to take place. Thus, the chance on behavioral, cognitive or emotional reactivity is decreased. In random interval time sampling a signaling device is a prerequisite which can create its own reactivity (Chapter 4).

The questions we wish to answer in psychiatry are related to behaviors but also to diffuse feelings and cognitions. Therefore, repeated measurement designs in psychiatry should favor time over event sampling and random over fixed time sampling.

Some principles of intensive field methods

Intensive field methods are applied directly in the appropriate situation, i.e. in the daily life of patients with mental problems, and not in the working office of the mental health professional. Because the behavior and the situations in which we observe are variable, the intensive field methods are repeated measures designs. We will now discuss some principles of intensive observation techniques for psychiatry.

1. *Observation of daily life*: Observations in daily life are ecologically valid. Contrary to the treatment setting, the snake, spider or rat you come across in daily life is real and can bite you. Although unlikely, the plane you are in can indeed crash and the therapist cannot intervene. Ecologically valid observations are not a panacea, however (Chow, 1987; Banaji & Crowder, 1989; Maher, 1991). The researcher does not control the occurrence of situations, behaviors, and feelings. Therefore, we do not know if the depressive patient feels better because it is evening, or because her husband is back home from work. Both situations often occur together in daily life. While the reality character of our data is high (we measure real situations, behaviors, feelings, and motivations), our potentiality to derive causal explanations is limited. To allow for this we have to manipulate the situations experimentally in a second phase in a laboratory or through a field experiment. In field research, the ecologic validity seems obvious, but is not. All kinds of situations change through the observational process. In the following we discuss how we can limit the reactivity of the observational process in daily life.
2. *Self-assessment versus other-assessment* Observations by a third person can be used to depict externally observable behavior. This is possible when the observation environment is restricted and when the observer does not disturb the interactions by his presence, e.g. the observation of family interactions during a family interview behind a one-way screen. In psychiatry, it is, however, important to know how people think and feel, not only how they behave. External observers do not have this information. It can only be obtained by "just asking" (Kelly, 1955). The shortcomings of observational methods (Harré & Secord, 1972) and the power of verbal self-reports (Simon, 1969; Ericsson & Simon, 1984; Harré & Gillett, 1994) support the use of methods that rely on subjective reports (see also Rorer, 1965; Norman, 1967; Wiggins, 1964). Given that there is no guarantee for the validity of the self-reported information, however, it is necessary to organize the data collection in a way that minimizes the risk for errors.
3. *Assessing repeatedly (sampling daily life)*: As we have already argued, assessing repeatedly and aggregating over those observations results in more stable estimators due to the statistical fact that a standard error is related to the sampling size.
 - *HOW OFTEN* do we have to observe? This is subject to the intended use of the results. If we assume that the repeated assessments are an unbiased sample of the same population of possible observations, then the larger the sample size, the smaller the error of estimation. If we want to compare assessments made in two different situations, the number of observations in each situation should be

large enough to allow a statistical discrimination (power argument). If, however, observations are used to describe a process over time, you need a sampling frequency at least twice the assumed rate of change. For example, when you expect that someone will recover from a panic attack within 5 hours, you have to sample at least twice within that period. A practical limitation is manageability. While we might assume that anger can emerge within a 5-minute talk, it is impossible to make repeated assessments over longer periods of time with twice this intensity. The combination of the rules illustrate that sampling frequency depends on the specific research questions. If we have more research questions, we use the highest sample frequency.

- *HOW* to observe? A basic rule of sampling is that each moment in the period on which we want to generalize our conclusions must have an equal probability to be selected as a observation moment. It is, however, often impossible to meet this requirement, because situations in daily life occur in clusters. If you sample someone during a period of domestic quarrel, quarreling can be overestimated when extrapolating over a year. Equally so, sampling in the summer will probably yield overestimated ratings of “outside” time. Therefore, we prefer random sampling procedures over the whole reference period.
 - *WHAT* to observe? When using field methods we are concerned to assess the phenomena under study as comprehensively as possible. Mental processes are described in terms of cognitions and emotions. The context in which they occur is defined by the time and place, the activities performed and the persons present. The environment of the moment is defined by the previous and the consequent observations. The collection of information that allows sequencing of the repeated observations over time is crucial (Bakeman & Gottman, 1986).
 - *RELATING CONTENT TO PROCEDURE?* Sampling success is related to frequency and duration of the events under study. We are better equipped to assess high frequent/long-lasting than low frequent/short-lasting behaviors. The latter requires a longer observation than the first. If a time sampling procedure is ineffective because we would report too often on non-target moments, we should consider event sampling to target the sampling process more purposefully. Time sampling, however, also generates information on non-target moments to use in comparisons, results in more precise assessments of event frequency, especially for long-lasting behaviors (with high or low frequency of occurrence), allows assessments of anticipatory moments and of variability within the events.
4. *Restricting the reactivity of the observation instrument:* Observing repeatedly in natural situations is not safe by design, since the subject, his environment and the observation instrument itself can produce reactivity. When a subject is instructed to make hourly assessments and has to check his watch constantly, he will not be able to describe his normal daily thought process. Therefore, it is often better if the subject doesn’t know when the next observation will come. An unexpected signal can “freeze” the flow of thoughts for reporting, ascertaining that normal life goes on between the observations. Furthermore, only the observer, and not the environment, has to react to the signal. We prefer a clear but unintrusive signal.

5. *Restricting time interval between the event and reporting:* The problem of global and retrospective memory is often present in one-time sampling procedures in psychiatric research (Lamiell, 1981; Willems, 1969; Stephens et al., 1984; Yarmey, 1979; Mischel, 1968; Fiske, 1971). Different experimental studies have revealed that assessments should be concrete and that a short time interval between an event and reporting of this situation is important for the quality of the observation. The sampling design can help to realize this, but the questions also have to be reformulated to apply only to a specific moment. This can be done by asking the subjects concrete questions, e.g. "How did you feel last week, when you came home tired from work on Tuesday night?", instead of "How do you usually feel when you come home from work?" Also, the time interval between the situations to be described and the report should be restricted. Memory problems already occur for intervals of 15 minutes or more. At that moment retrospective bias plays a role.

THE EXPERIENCE SAMPLING METHOD (ESM)

The Experience Sampling Method (ESM) is a field-intensive assessment instrument that uses a self-observational strategy cued by an electronic beeper. Repeated structured observations of mental state and psychopathology are collected contingent on a random time signal. The assessment questionnaires collect reports of cognitions, moods and behaviors, as well as the context in which they occur, within the natural, daily life of the research subjects. In our application of ESM we sample subjects typically ten times a day between 7:30 in the morning and 10:30 in the evening, over the course of one week.

Rationale, development and applicability

The traditional field-intensive methods of psychiatry are unstructured diaries and structured monitoring procedures. Neither method is well suited to generate valid, subjective reports on mental state or to assess experiential variability in the natural flow of life. Thanks to the modern digital technology, ESM is able to integrate the best part of both instruments. Over the years it has developed independently from different fields and variations of this family of repeated measurement strategies have been applied in different fields.

During the last 15 years, the applicability of ESM and related methods has been demonstrated in a large number of studies. The primary field of application has been in developmental psychology. Studies of normal childhood development assessed the communication between mothers and their children (Wells, 1985), puberty (Jaquish & Savin-Williams, 1981; Savin-Williams & Jaquish, 1981) and adolescence (Csikszentmihalyi & Larson, 1984; Csikszentmihalyi et al., 1993). Specific groups were also studied, such as, young delinquents (Csikszentmihalyi et al., 1977) or in US and African schools (Malik, 1981). Adults were studied at different ages (Graef, 1979; Larson et al., 1980; Zuzanek, 1980) and in specific states — e.g. Sjöberg & Magneberg (1990) described the relation between actions and emotions and Hamilton et al. (1985) mood fluctuations during the menstrual cycle.

Another important field of interest studied by these intensive sampling techniques is the effect of the cultural environment on experience. In social psychology social interactions have been studied with a related assessment strategy (Wheeler & Reis, 1991; Reis & Wheeler, 1991). Delle Fave & Massimini (1988, 1991; Delle Fave et al., 1992) used ESM methods in their research on human adaptation in different cultural niches. Hormuth (1986) applied ESM in a study of the effect of urbanization on daily life. Larson & Richards (1994) published a monograph on the lives of parents and their adolescent kids. Richards & Duckett (1994) assessed the influence of maternal employment and adolescent experience. Kubey studied the impact of TV viewing on perceived life satisfaction (Kubey & Csikszentmihalyi, 1990; Kubey, 1990).

Related techniques for the study of normal and abnormal mental state have been developed independently in psychiatry and clinical psychology. deVries and colleagues studied adaptational processes and environmental reactivity in relation to psychopathology (deVries, 1987, 1992). To date, a large range of psychiatric symptoms have been studied: bulimia (Johnson & Larson, 1982), addiction (marijuana: Larson et al., 1992; heroin: Kaplan, 1992), multiple personalities (Loewenstein et al., 1987), anxiety (deVries et al., 1987, 1990, 1992; Dijkman-Caes & deVries, 1987, 1991; Dijkman-Caes, 1993; Dijkman-Caes et al., 1993; Margraf et al., 1987; Margraf, 1990; Kenardy et al., 1988, 1989), depression (Kraan et al., 1992; Wilson et al., 1992; Merrick, 1992; Van Diest, 1992, 1993; Hurlburt, 1993); pain (Portegijs, 1988; Lousberg et al., 1993); adult autism (Hurlburt et al., 1994) and schizophrenia (deVries, 1983; deVries et al., 1986; Delespaul & deVries, 1987; deVries & Delespaul, 1989; Hurlburt et al., 1987; Hurlburt, 1990).

Supplementing the work of Klinger (Klinger et al., 1980; Klinger & Cox, 1988), Hurlburt (1990, 1993) and Emmons & King (1992) have studied the course of conscious thoughts in daily life. Wells (1992) monitored variations in self-esteem. In behavioral medicine the impact of stress and cortisol has been studied (Nicolson, 1992, van Eck & Nicolson, 1994) while the effects of hassles and small events on mental state (daily coping) was studied by Stone and colleagues (Stone & Neal, 1982, 1984a,b; Stone et al., 1985, 1987, 1993).

Finally, clinicians have used ESM techniques in mental health practice. Some authors have focused primarily on assessment using ESM (Figurski, 1992; Delle Fave & Massimini, 1992; Massimini et al., 1987; Massimini & Massimo, 1988; Donner, 1992). Other therapists developed ESM based applications for behavioral based treatments (Filstead, 1988; Taylor et al., 1990; Kenardy et al., 1992; Paty et al., 1992).

Procedure

Briefing session — Subjects who are referred for ESM are invited to a briefing interview with a research assistant. During this interview we explain the purpose of the study and we ask whether they would like to participate (informed consent). Subject characteristics are collected and cross-sectional questionnaires are filled in. When assessing cross-sectional psychopathology we also define the subject's most important complaints as he or she would formulate them in his/her own words (idiographic complaint). These will be assessed repeatedly. Because of the complexity of the sam-

pling procedure — assessing mental state repeatedly over the course of one week — we direct much attention to the development of a personal research relationship (alliance). Anticipated problems with the signaling device, a digitized wrist watch, and the questionnaire forms are discussed as concretely as possible. A try-out sampling moment is simulated in which the subjects were coached to answer the forms. We ensure that everything is clear for the subject. If in doubt, for instance when we observed that the subjects had trouble assessing the Likert-point scales (see next section), we decided to use a less abstract questionnaire (see Appendix 4) with a smaller number of items and in which the 7-point scales were substituted for “yes/no” alternatives. This was very exceptional, even with chronic schizophrenics. Another problem that occurred more often was that the subjects had trouble reading the small scale, pocket-sized booklet. When we observed this we used a booklet with a larger print. Finally, we sometimes observed during the briefing that subjects had trouble hearing the digital beep. In very exceptional situations the subjects could not be sampled. As a fall-back strategy we could decide to give the sampling device to a family member or a nurse on the ward and ask them to notify the subject at each signal. Not all problems can be taken care of, during the briefing session. We give the subjects a telephone number and encourage them to phone us when they encounter any problems. Also, we check on them repeatedly during the sampling week. Rather often we have observed that the subjects have drifted away from the original instructions but most problems can easily be solved with a “booster” instruction after one day of sampling. For the most invalidated subjects on the chronic hospital wards we first sample the nursing staff. This gives the patients a model to comply with, but has also ensured that skilled ESM users would be available on the floor continually during the sampling weeks of the chronic schizophrenics.

ESM week — For the next week, the subjects continue their “normal life”. He or she will be beeped to fill in the ES-forms randomly, 10 times a day, between 7:30 and 22:30. They will also have to answer additional questions in the morning when they wake up and in the evening when they go to sleep.

Debriefing session — After a week, a debriefing interview takes place. During this interview the research week is reviewed on the basis of the completed questionnaire. Missing beeps are discussed. Also, the forms are checked for anticipated coding problems. These can occur because the subjects are stimulated to respond to open ended questions using their own most natural descriptions. They will, for instance, respond to the question “Who are you with?” by writing down “with John”. The researcher does not know who “John” is for the subject and has to check whether he is a colleague, a friend or a family member.

Questionnaire construction

The Experience Sampling Forms (ESFs) are packaged together in small pocket-sized booklets — one for each day. They can easily be carried. The questionnaire itself is derived from the time budget diaries and is based on the most elementary diagnostic instrument in psychiatry, the mental state examination. The mental state items are

assessed on 7-point Likert scales and are presented in clusters on thoughts, moods, activity motivation, perception of social interaction and physical well-being. The questionnaire further contains a specific diagnostic module for psychopathology, characteristic for the disorder under study. Thus, with the ESFs, we can measure the relation between psychopathologic symptoms and the overall mental state. It provides a strong repeated evaluation of diagnostic criteria. The questionnaires also contain information about the context of the beep: where are you, what are you doing, who are you with, what time is it? This allows the assessment of environment-related changes in mental state. Finally, the forms also contain questions on positive or negative events that occurred between beeps, or psychotropic substances that might have been consumed. These are necessary to evaluate the effect of relevant between beep events in daily life.

Item selection — Item formulation in ESFs is different from traditional questionnaires. Traditional questionnaires are one-time assessment strategies. Therefore, reliability is maximized by measuring a concept from as many point of views as possible. The target concept (construct) lies somewhere in the middle and the score of a subject can be obtained by a combination of the items in a scale. In ESM, subjects are rated repeatedly over time and the reliability of the assessment is obtained by this repeated assessment. However, we are also interested in the moment-to-moment variability on an item and therefore, individual responses on specific items should be measured reliably (Chapter 5). The first rule is that concepts should be measured directly and not indirectly. We ask the subjects directly whether they are “anxious”, “depressed”, “irritated”, “lonely”,.... The fact that “anxiety” and “loneliness” are related to “depression” is no argument to use a depressive summary score, washing away between item differences at a specific moment in time. Our interest in variability makes non-anxious but depressive moments interesting study objects. We can always find arguments to state that two mood items are different, but we will also agree that subjects will only exceptionally rate themselves “tense” and “relaxed” at the same moment. More likely we will find moments of “irritation” without “anxiety”.

The selection of items for ESFs is primarily based on content arguments. Emotion research (Frija, 1985; Izard, 1991; Plutchik, 1993) and also focus groups (Krueger, 1988) can help in optimizing the item choices and fine-tuning the formulation. Our experience is that there is no penalty, in response time of the whole ESF or in subject drop-out, for adding even a large number of Likert type scales. If we are in doubt, we usually include both item formulations. There is a high penalty, however, for additional open ended questions, which take a lot more time to answer. We should be careful adding this kind of question.

Data reduction should be a concern in ESM. However, computing the factor structure of an ESM questionnaire is a complex achievement. The variance/co-variance matrix that is used in the factor analysis can be computed differently as a result of the complexity of the ESM data structure (repeated measures in salient situations for each subject)¹⁸. The resulting factor structures will be different from one subject to another,

¹⁸ Technically, the Q-technique (to assess variability between subjects) or the P-technique (to assess the variation of a single subject over a population of occasions) in contrast to the traditional R-

when computed on subject means by item, on normalized beep data by subject, with and without additional dummy variables, when computed for different moments of the day, or even for different situations (e.g. alone/not alone). Almost every data analysis warrants an individual factor analysis, a most confusing situation. Considering this highly unreplicable situation and except for a rather tautologic general factor "positive/negative mental state", the resulting factors will be extremely difficult to interpret. As a tool to target item selection it is of little value.

Item characteristics are important anywhere and also in ESM. Items with low variability are of little use, but should they be skipped? Anxiety, for instance, shows little variability in normal subjects, but even then it allows us to discriminate between moments with and without anxiety. It is therefore clear that the fact that data are not normally distributed is not an argument to skip items. Of course, parametric statistics will be skewed in this situation and we should rely on non-parametric alternatives. Most of the analyses in this book are based on summary scores by subject. While the beep level data are skewed, those subject level data are more normally distributed, allowing the use of parametric statistics.

Item formulation — The reference point for the answers is the moment before the beep — a fraction of a second before the beep to be exact. We are not interested in emotional reactions as a result of the signal¹⁹ but only on the emotional state at the moment of the beep. Some ESM developers have formulated all the questions in the past tense, to stress the fact that the moment just before the beep went off is referred to. I strongly disagree with this practice for two reasons: 1) past tense sentences have a more complex syntax and this tends to confuse the cueing item for the scale (compare "I'm happy" with "I felt happy (just before the beep went off)"); and 2) past sentences include not only the moment of the beep, but also an undefined period before the beep into the assessment frame. As a consequence, the reference to the actual moment is easily lost. As a rejoinder to the critics I sometimes use a past sentence in its extensive form as a reminder of the instruction, but only for the first question of the ESF questionnaire: "What was I thinking just before signal went off?". Although less strongly, I also favor the formulation of items in the "I"-form over the "You"-form, arguing that the first creates more intimacy (as in a diary). This way the subject's task is more private and focused on the moment and its current mental state ("I'm happy" versus "You are happy"). Furthermore I prefer items formulated as an assertion over items formulated as a question ("I'm happy" contrasted to "Am I happy?" or "Are You happy?").

Scaling issues — The assessment scale format is kept standard over all the items. Original attempts to use semantic differential scales were discontinued because they were too confusing for chronic mental patients. More than 50% of the patients with schizophrenia, but none of the normal controls, gave double answers on semantic

technique (assessing the variability between items over subjects) could be applied (Cattell, 1946, 1950). More recently developed approaches of relevance are the dynamic factor analysis (Immink, 1986) and of course the LISREL models (Jöreskog, 1990; Jöreskog & Sörbom, 1989).

¹⁹ Although a question on the disturbance due to the beep is added at the end of each ESF, ESM reactivity assessments are rated for emotions and activities at the end of each day, and at the end of the whole sampling week.

differential scales (deVries et al., 1986). They gave independent answers to both ends of the scale. The frequency of these uninterpretable answers, however, was low (0.3 percent). To understand what caused this difference we conducted a pilot study with normal subjects and this showed relatively low correlations (0.40 - 0.60) between the emotional cues that composed the two ends of the original semantic differential scales — now presented as Likert type scales. Considering the problems we found in normal subjects with these scales the “concrete thinking” of schizophrenic subjects may have confused them, resulting in double answers on the same semantic differential scale.

Visual analog scales (VAS), while allowing a lot of differentiation, make later data-entry a more cumbersome process. Guyatt et al. (1987) compared visual analog scales and different Likert scales and found no significant differences, concluding that there is no gain in the use of the VAS. Another alternative would have been to use yes/no scales. However, these scales are not sensitive and maximize social desirability (Wentland & Smith, 1993). We therefore settled for 7-point Likert type scales (Likert, 1932). These scale alternatives have proved to be sensitive in detecting clinical changes of at the least 0.5 in the scale (Jaeschke et al., 1989).

Some may argue that the cueing words for the scale entries (1=“not”; 4=“moderate”; 7=“very”) apply more to one item compared to the other. While the item “I feel relaxed” can use the standard labeling, “We are doing things together” is less confusing with an alternative labeling: 1=“a little” to 7=“very much”. However, changing the labeling of scale entries results in a more complex questionnaire, something we try to avoid. The instruction specifies that, when in doubt, the scales should be interpreted as 1=“the item doesn’t fit my current situation at all” to 7=“the item applies perfectly to my current situation”. I prefer using the most natural “expression” in the culture to cue items. This can, however, lead to double negations in a question (e.g. “I feel unwell” is preferred to the stronger “I feel sick” but a score of “1” = “not feeling not well” is confusing). Situations like this can be avoided using the above mentioned extensive scale interpretation (applies / does not apply).

Even when yes/no alternatives seem more appropriate for an item we still use an ordinal scale. For instance, “I hear voices” and “I’m panicking” seem to be “yes/no” kinds of questions. However, when rated on 7-point scales, the intensity of the symptoms will be assessed. There are exceptions: “I am alone” is a fact and it is pointless to assess that on a ordinal 7 point scale (this in contrast to “feeling lonely”).

Questionnaire design — Because filling in the ESFs disturbs the flow of daily life, everything in the questionnaire design should help the subjects to fill in the forms as quickly as possible, without losing reliability. Items are considered stimuli and are kept as short as possible — often not more than one cueing word. Longer descriptions take up more time and will not be read after the first beeps. Furthermore, items are organized in blocks of maximum four. This avoids errors of giving answers on the wrong scale. Finally, a large number of open ended questions should be avoided. If necessary, open ended questions should allow responses that require no thinking (for instance, use the name of persons, not their relation with the subject). Coding problems should never be delegated to the subjects. This would take too much time to answer. Furthermore, we would miss the opportunity to check on coding reliability.

Balancing positive and negative items in the presentation order is an important trick to force subjects to respond actively to individual items. Without this balancing of items subjects would respond to a general positive and negative feeling and we would miss the experiential nuances. Sometimes it is easier to find negative mood scales than positive mood scales. Nevertheless, positively formulated items can result in more variability and should be preferred (compare “I feel tense” with “I feel relaxed”).

Sampling device and sampling schedule

Chapter 4 is entirely devoted to the selection of sampling devices and schedules. In the ESM application we use in Maastricht, the sampling device is a digital watch (Seiko RC-1000) that is preprogrammed by a computer to signal 10 times each day. The sampling procedure is a stratification of one random moment within each time block of 90 minutes between 7:30 and 22:30.

Analyses of the ESM data

ESM data are complex. This complexity relates to the extensiveness of the data file as well as to the diversity of variables. Chapter 7 is entirely devoted to ESM analysis.

Instructions for analyses — There are pitfalls for analyses of ESM data that can easily lead to wrong conclusions even with these ecologically valid data:

- *INTERDEPENDENCE OF OBSERVATIONS*: Data bits in standard statistical techniques should be independent. This condition is not respected in ESM. When we compare two subjects in a traditional experiment, we can assume that the answer of one subject is completely independent of the answer of the other: the reaction of one subject to the medication in no way determines the reaction of the other subjects (except for the quality of the medication). In ESM, data is not independent and we must assume that when a subject is depressive on the third beep of the fifth day, he was probably depressive the moment before and will probably feel the same on the next moment.
- *DEPENDENCE OF SITUATIONS*: With ESM we try to describe subjects by repeated observations on randomly chosen moments in their daily life. From those observations it is wrong to conclude how subjects feel in well-defined situations. The problem is caused by the fact that the subject can select the places, persons and activities as a consequence of his feelings. The ESM observation that agoraphobics feel fine outside can be the consequence of the fact that they only go outside when they feel good, which rarely occurs. The assessment of the experiential value of an environment (protective or illness-inducing) is only possible when the occurrence of those contexts is under experimental control (randomized or at preprogrammed moments in time). This requires a “natural experiment”. For instance, the subject will be told to go outside regardless of his mood and we compare his mood at those moments with prescribed moments of being inside.

ESM study design — Analyses of ESM data are very different when studying large groups of subjects or individual cases (n=1 studies):

- “*N>1*”- STUDIES: In studies with a large number of subjects, researchers often use statistical techniques to correct for the lack of independence in the data. Strange and unexpected things can happen in this situation, leading to significant relations that are “blown up” (for instance when we disregard the fact that data comes from the same subject), or meaningful trends that are blurred (when summary scores by subject are used) (Margraf et al., 1987; Delespaul & deVries, 1987). Complex statistical techniques such as models based on Markov chains, time series analyses or multilevel analysis (Gibbons et al., 1993) can correct for this interdependency and promising results have been obtained (Allison, 1984). New cluster strategies can also be used to reconstruct diagnostic groups using daily life variables (van Meter, 1990; van Meter et al., 1992; Fink, 1986; deVries et al., 1987). The problems with analysis of ESM data are, however, caused by the nature of the data material itself. Due to the systematic intercorrelations involved they are difficult to correct, even with complex statistical techniques. Caution is warranted. Larson & Delespaul (1992) have advised that analyses can best be done using simple statistical techniques (means, standard deviations, correlations, frequencies) using data aggregated at the subjects level. Conclusions should always be cross-checked using different analyses strategies. Recently, complex random regression techniques were developed that adequately assess the variability at the beep and subject levels (Gibbons et al., 1988, 1993; Hedeker et al., 1993).
- “*N=1*”- STUDIES: In the clinical application of ESM for individual patients these problems are less important. First, the independence of observations is better ensured because all the observations belong to only one subject. Furthermore, the conclusions derived from the data are not generalized “laws”, but interpretations that target our interventions in therapy and whose validity can be assessed in practice through the evaluation of the success of a clinical intervention. If the results were caused by spurious covariation in the data, the expected effect will not be obtained. An example can illustrate this. If we assume that alone time causes depressive feelings because we observe more depressive feelings in alone situations for our patient, we can design an intervention to minimize alone time. If the depressive feelings do not diminish as a consequence, our hypothesis is not confirmed and another conclusion should be forwarded.

Questions and data in the ESM research — To minimize possible confusion we advise formulating hypotheses before starting the analyses. This should always be done, but especially in ESM research. The relation between depression and alone time can be presented in many ways. Both “Are depressive subjects alone more often?” (a question about subjects) and “When people are alone, do they feel more depressed?” (a question about moments) are valid operationalizations but they can yield different results (Larson & Delespaul, 1992). The best way to gain insight into ESM is to illustrate how ESM-variables supplement traditional psychiatric assessments. Examples are primarily drawn from schizophrenics, anxiety and normal control subjects. They illustrate the questions that can be answered with ESM:

1. *RAW DATA (IRREGULAR LINES)*: Figure 3.1. is an example of the daily variability of mood (“tension” and “depression”) and psychopathology (“anxiety” and “hearing voices”) over a period of six days of a young male schizophrenic patient. The data

illustrates variability with patterns of synchronization (items showing the same pattern) and desynchronization (items diverging for periods of time). The availability of these dynamic patterns of observations of daily variability is quite new and we hope that they can be used therapeutically as well as for the development of new theories of psychopathology.

Figure 3.1 Irregular patterns of synchronized and desynchronized experiences in mood (tension and depression) and psychopathology (anxiety and "hearing voices") of a schizophrenic patient (Delespaul et al., 1994).

2. **TIME ALLOCATION:** Time budget patterns are frequency distributions of observation codes (presented in pie or bar graphs) that can be translated into the amount of time subjects spend in a well-defined situation. Time budgets reflect the frequency with which subjects use places, engage in social interaction and perform activities. Time budgets are very sensitive indicators of global daily life adaptation. Differences were found in the time budgets from different cultures (Szalai et al., 1972; Robinson, 1987) and give a good indication of illness related behavioral changes between as well as within diagnostic groups (deVries et al., 1988; Figure 3.2) and within an individual in different phases of illness (Chapter 13, Case 1).

Figure 3.2 Time allocation for social interaction compared across diagnostic groups (chronic schizophrenics (n=33), depressive (n=40), anxiety (n=50) and pain patients (n=35) compared to normal subjects (n=44) (deVries et al., 1993).

Time budgets are not restricted to contextual codes. Figure 3.3, for instance, shows how much time schizophrenic patients are ill based on two psychotic symptoms.

Figure 3.3 Visual and auditory hallucinations in 50 chronic schizophrenic patients. Pie (score 1): more patients had no auditory compared to visual hallucinations. Bars (scores 2-7): the severity is higher for auditory over visual hallucinations (Delespaul & deVries, 1994).

Finally, time use can be dynamically patterned over time. The following example depicts the change in place from one moment to the next in a group of anxious subjects with comorbid depression. We were able to demonstrate that the transition patterns were more restricted in this group, compared with anxious subjects without comorbid depression (Figure 3.4). Data such as these may serve as an ethogram or sociogram for research purposes that may illuminate hidden aspects of social life and individual adaptation. Finally, time budget graphs can also be drawn for interacting groups of individuals such as families. In this case, the amount of time the family actually spends interacting together as well as the quality and mutual balance of this interaction can be depicted.

Figure 3.4 Ethogram: time use changes from one moment to the next for anxious subject with comorbid depression (deVries, Delespaul & Dijkman-Caes, 1987).

3. *COMPARISONS OF MENTAL STATE*: The ESM reports constitute repeated measurements of mental state. They can give an indication of the frequency and duration as well as the intensity and variability of psychopathology. ESM data therefore have diagnostic value. Table 3.2 shows that, despite the difference between ESM data and traditional diagnostic material, the self-observations from the repeated measu-

rements come up to the expectations of traditional diagnostic techniques. However, given that the ESM assessments are based on aggregated answers on single item experiential descriptors assessed prospectively in daily life, these observations can be considered very powerful descriptors of psychopathology. ESM data are therefore especially useful for evaluation purposes.

Mental state: $\bar{X}$ (s.d.)	Schizophrenia	Anxiety	Pain	Normal	$F_{(3, \dots)}$
Thoughts	4.51 (1.45)	4.51 (1.67)	4.94 (1.38)	5.23 (1.16)	28.28 ***
Mood	4.78 (1.36)	5.01 (1.11)	5.86 (0.82)	6.27 (0.67)	223.91 ***
Individual complaint	3.81 (2.16)	2.75 (1.78)	3.24 (1.67)	1.54 (0.84)	152.96 ***
Activity motivation	4.19 (1.10)	4.76 (1.21)	4.96 (1.11)	5.37 (0.89)	89.05 ***

Table 3.2 Aspects of mental state (weighted: 1=negative, 7=positive feelings) over different pathologic groups and normal controls (deVries et al., 1988) (***) $p < 0.001$.

The power of ESM as a cross-sectional assessment instrument can further be illustrated in a sample of 50 chronic schizophrenic divided in subsamples based on age (Figure 3.5). Sampled aggregated feelings of anxiety and loneliness grow with age, while satisfaction has a curvilinear relation with age.

Figure 3.5 Sampled mental state by age in schizophrenia (n=50) (Delespaul & deVries, 1993).

4. **MOOD IN CONTEXT:** The most unique contribution that the ESM makes to psychiatric research is the possibility of locating pathologic phenomena or mental states in naturally occurring contexts. Connecting mental state to context is possible with ESM because both are assessed concurrently. In the psychopathologic literature we know syndromes for which variability in intensity is explained by changes in context: e.g. places and persons in agoraphobia and time of day in depression. It is unclear if other "sensitive environments" exist, for instance in other syndromes. In ESM we describe the context of the beep by means of 4 variables: physical place (where?), social environment (with whom?), activity (what?), and time (when?). In Figure 3.6 we present the results of a comparative study of "feeling happy" during different activities. Social activities and active leisure rate highest, followed by work and study, while "watching TV" and "doing nothing" is the least concurrent with happy feelings. We compared these observations in different samples (schizophrenia, depression, anxiety, pain and normal controls) and found the same pattern in each diagnostic group. These replicated patterns of activity related moods were

found irrespective of the existing overall differences in the level of “feeling happy” for the different groups.

Figure 3.6 Mental state (“feeling happy”) by concurrent activities over all subjects and by diagnostic group (subject level analysis) (deVries et al., 1993).

A restriction of these analyses is that the context sensitivities are individually determined and often disappear by averaging over subjects. Mood in context is complex, e.g. how do people feel when they are in a “not alone” situation at one observation moment and “alone” in the next? In the example below (Figure 3.7) we select the “alone”/“not alone” situation within the sample of chronic schizophrenics. These subjects were as a group significantly more anxious when alone. However, 40% of the subjects who showed environmental reactivity had the inverse relation in their individual data: they were more anxious when not alone.

	#	%
Alone more anxiety	21	60
Context independent anxiety	4	-
Alone less anxiety	14	40
No variability on anxiety	11	-

Overall context-relatedness of anxiety and social situations in schizophrenia:
 $F_{(40,1)}=5.44 *$ (more anxious when alone)

Figure 3.7 “Anxiety” in social context (“alone” / “not alone”) broken down by subjects for 50 (subset of 20 in graph) chronic schizophrenic subjects (Delespaul & deVries, 1994).

5. *MOOD PATTERNS*: A last question that can be derived from the ESM data and will be discussed here is the evolution of mental state over time by changing contexts. The example we present here are so-called recovery curves. Recovery curves describe the process over time of recovery from a level of intensive psychopathology (episode) to the subject’s individual baseline. In the left-hand side of Figure 3.8 we can see how the recovery from anxious and depressive moments is different for normal individuals — depressive feelings show a much slower recovery. The

right-hand side of the figure shows the recovery from an anxious episode of anxious subjects with differential intensity of comorbid depression — rated cross-sectionally on the Zung. Compared to the less depressed anxiety subjects, the group with more depression recovers more slowly from a high anxiety episode — expressed in time after 3 hours in the first group and only after more than 5 hours in the latter group. We are convinced that recovery curves are promising as an outcome indicator in therapy evaluation research and a good indicator for coping mechanisms.

Figure 3.8 Recovery from anxious and depressive episodes in normal subjects (left) and recovery from anxious episodes in anxious subjects with differential intensity levels of comorbid depression (right) (deVries et al., 1987).

DISCUSSION: UNDERSTANDING ESM DATA

How can we assess variability between and within individuals in psychiatry and how can we ascertain that this assessment is meaningful for research and clinical practice? Although ESM is a powerful assessment strategy, it is not a panacea. Different aspects related to psychiatric assessment makes ESM less suitable in some situations. Assessment costs are only one of these aspects. In this discussion we will position ESM among other assessment strategies in psychiatry and present some recommendations as to the situations in which its use is the most appropriate.

Questions in psychiatry

In the introduction to this chapter we have presented a rudimentary taxonomy of various relevant assessment questions in psychiatry. These can be defined on four different levels: syndrome, symptom, etiology and course. Those questions are primarily applied to characterize individuals, but can also be applied to moments within the life of subjects. Understandingly, the specific time window becomes important. In psychiatric illness several time windows are relevant and without being exhaustive — intermediate levels are possible — we can discern lifetime development, current epi-

sode, current phase, and actual mental state. Operationalizations of specific questions for the lifetime characteristics of illness (the subject level) and the assessment of current mental state (the moment) are summarized in Table 3.3. The information that we need to answer these questions are difficult to generate with traditional assessment strategies such as clinical interviews and standardized questionnaires. Therefore, we need a broader spectrum of assessment alternatives.

	Lifetime	Moment
Syndrome	<ul style="list-style-type: none">• <i>trait</i> assessment: placing subjects within a diagnostic group	<ul style="list-style-type: none">• <i>state</i> assessment: actual combination of symptoms in relation to diagnosis.
Symptom (occurrence & severity)	<ul style="list-style-type: none">• assessing salient symptoms within syndromes collapsing for a specific subject over lifetime: comorbidity	<ul style="list-style-type: none">• assessing salient symptoms within syndrome for a specific subject at a specific moment in time: symptom co-occurrence and synchronicity (co-varying severity)
Etiology (causes of syndrome emergence / symptom co-occurrence)	<ul style="list-style-type: none">• genetic predisposition, life events, personality structure, social economic status,...	<ul style="list-style-type: none">• assessing the causal factors for syndrome emergence and symptom occurrence within the more restricted time frame. Assessment of sensitive environments, differential coping processes, effects of hassles, etc.
Course (describing / predicting)	<ul style="list-style-type: none">• expected patterns of syndrome evolution over time: full recovery, episodic, chronifying.	<ul style="list-style-type: none">• expected patterns of symptom severity over small time windows: day patterns, relapse and recovery from episodes of high symptom intensity

Table 3.3 Overview of assessment problems in psychiatry related to lifetime prevalence and momentary state.

A taxonomy of methods in psychiatry

Irrespective of the content of the assessment, three classification principles are being introduced to describe the differences in the nature of assessment procedures in psychiatry (see Figure 3.9). Not all possible alternatives are equally relevant.

Figure 3.9 Overview of assessment procedures in psychiatry using one or more observations, available or new information, and referring to single moments or previous periods.

- The first classification procedure refers to the number of measurement moments. The possible options are a one-time assessment — also called “cross-sectional” — and repeated measurements. In the latter situation the number of assessments can be extended respective to the research requirements;
- The second classification dimension refers to the time period to which the data refer: we can use available data relating to a period from the past (retrospective) or new information collected over a future time period (prospective);
- Finally, the third dimension refers to the time interval between the events and the reports. The options form a continuum starting at extremely remote periods in time and ending at a restriction to the moment itself. In the remote assessment option we can have extended assessment periods (one minute, one day, one month).

The different assessment strategies have a different sensitivity for assessing variability. Figure 3.10 illustrates this. The sensitivity for change is highest when using a prospective data collection strategy, assessing repeatedly and reducing the interval between the target event and the report.

Figure 3.10 Differential sensitivity for the assessment of variability: highest sensitivity for assessing change is found when using a prospective data collection strategy, assessing repeatedly and reducing the interval between the target event and the report.

Which change is relevant?

When assessing change in mental state we should always reflect on which change is relevant. When a schizophrenic withdraws to his bedroom or complains of lack of motivation this may be a prodromal sign of the illness, signaling an emerging relapse, but can also be a benign symptom of normal daily life variability. Not all changes are equally relevant. Therefore we will present a model describing different phases of changes. Our goal is to generate better assessments of mental state. Our method of

choice is the Experience Sampling Method — a technique that enables us to describe changes in mental state over time and within a natural context in a sensitive and valid way. The reasons why ESM is a sensitive assessment strategy are not straightforward. Figure 3.11 is an attempt to illustrate the complexity of mental state changes. It shows the potentiality of ESM but also its limitations and should be considered a rationale for selecting appropriate assessment techniques for change in psychiatry.

Figure 3.11 Working model of ESM data.

Mental states — irrespective of their biologic and/or environmental cause — originate somewhere in the human organism. This human organism is a bio-psycho-social system that pursues an internal and external homeostasis, actively and passively monitoring and adapting to internal and external influences. Due to this multilevel adaptational process, measurement of the present mental state of the organism is possible at different levels: the bio-physiologic level; the level of momentary emotional recognition; the behavioral level, the level of retrospective self-perception and the level of global life adaptation. Understanding the meaning of those different levels can help us to develop and select specific diagnostic tools and target the assessment strategies for our specific research and clinical needs.

Immediate emotional reaction — We assume that an internal or external event, that disturbs a subject's homeostasis, causes an almost immediate physiologic and emotional reaction²⁰. For the purpose of this description we will follow the path of emotions. We are convinced that these immediate emotional reactions (IERs) are different from a traditional concept of emotional states. While IERs are perceived adequately by the subjects, they are usually not remembered. The limitation resembles the working mechanism of short term memory. The recall of IERs has an implicit capacity limitation. Immediate reporting, before the information is lost, is therefore crucial.

The most important threat to the reliable reporting of this "immediate emotion" is the mental processing of the experiences. To avoid possible disruption of the reports the assessment strategy should be designed in such a way that we avoid the subject interpreting his mental state. The subject should stick to the moment and only to the moment, when reporting his experience. This moment is not a characterized moment,

²⁰ It is not important to evaluate which of both changes comes first. The model assumes that change is a disturbance in the systemic equilibrium and will, therefore, be scattered in the different fields of the system. Moreover, due to the hysteresis of both — physiologic and emotional — processes the question of which comes first will probably never be answered reliably.

moment, when reporting his experience. This moment is not a characterized moment, but a specific instant — now — experienced as a “gestalt”. It is not the moment when “I am alone and feel depressed as I always do in these circumstances” but rather the decontexted instant with a specific time, place, activity, persons, motivation and concurrent mental state without a (personal) theory that interrelates and interprets these elements. The possible hypotheses and interpretations the subject might have about the concurrent characteristics of the moment are hidden — not activated into his or her awareness.

This process is entirely different from the one in behavioral monitoring. For instance, in a study of the link between loneliness and depression using behavioral monitoring we ask subjects repeatedly to rate depression and count the number of persons in the environment. However, by restricting the assessment to the research question itself, the research hypothesis is imposed upon the observed situation and aspects of the retrospective self-assessment, especially the hypotheses that the subject or the researcher has about the symptoms, will influence the answers.

On the other hand, when we describe the moment as it is, with no explicit reference to relating characteristics, we generate a situation that is comparable to “free association”. The subject will formulate his assessments in response to the global “stimulus-moment” in a pre-conscious/ pre-interpretative way. All answers will be related to that moment and not to previously given answers or assumptions. Each item characterizing the moment, is rated independently, with reference to the moment itself. The subject is not invited as in a diary to reflect upon his life, mental state and activities, but, in contrast, dissuaded to create his/her interpretative “gestalt” of the moment. Such an assessment procedure minimizes and avoids non-target influences such as “social desirability” or assumptions about the relation between characteristics of the moment and the mental state, to slip into the data. The measurements that are obtained, therefore, are more valid reflections of the actual and momentary mental state. If we are interested in emotional reactions to events or, more generally, in the present mental state, an assessment strategy such as the Experience Sampling Method is a good choice — probably even the best possible option. To be sure, however, that what we assess is relevant for the study of psychopathology, we must know the relation between this “immediate emotional reaction” and bio-physiological change reactions to events as well as the subject’s evolving behaviors and his or her self-narrative. What is the relevance of reliable assessments of moment-to-moment changes in mental state, when these are unrelated to induced changes in neuropsychological functioning or influences of psychotropic substances. Equally so, what is the use of these assessments if the subject is still unable to perform specific activities or continues to complain of depression or a marginal quality of life? Therefore, assessment strategies in psychiatry should be multi-method and a in-depth study of their interrelations should be central in our future research.

Behavioral change — What subjects do is partially and indirectly a reflection of how they feel. How can emotions change behavior? The “operant learning paradigm” explains how behavioral change occurs through its consequences. The momentary emotions are the best possible assessment of the subjective positive or negative experien-

ce²¹ and act as reinforcers. Therefore, emotions influence the frequency of expressed behaviors. Because IERs are not remembered, this learning process is unconscious. Although the subjects are aware of their own behaviors, they usually do not know why they make subtle frequency changes in their daily life routines. The schizophrenic has no conscious rationale for his withdrawal, nor does he know which behavioral strategies are effective coping options. Nevertheless it might be possible that his behavioral repertoire has changed in a meaningful direction — ruled by his cognitions and emotions. Because the subject is not aware of the actual working mechanisms he is not able to change his behavior purposefully and cope adequately. Behaviors are a reflection of the daily life adaptation. Assessments of behavioral changes are possible through systematic repeated observations over time. The functional analysis of those data yields useful information for the planning of psychiatric interventions.

Intriguingly, the “operant learning paradigm” should enable predictions of behavioral change through the observations of elucidated emotions. If this paradigm works as predicted, subjects in a steady-state situation would be observed in those environments in which they feel best. Empirical data show that this is not the case. Behaviorists have tried to explain this counter-intuitive phenomenon by challenging our assessment of “reward” and “punishment”. The “law of effect” states that rewarded behaviors will increase in frequency. Equally so, consequences of behaviors that increase in frequency over time are defined as “rewarding”. This is circular reasoning. Relabeling the consequences from reward to punishment based on observations of changes in behavioral frequencies is unsatisfactory. When we compare those relabeled consequences with assessments by the subject of his like and dislike of the current context — directly assessed as IERs — we see that “rewarded” behaviors do not always change in frequency. Moreover, such circular reasoning does not allow us to predict the behavioral change that we need for planning treatment interventions.

Nevertheless, the observation of behavioral changes is important for treatment planning. The combination of both behavioral observations and assessments of IERs are important to target interventions and to optimize the selection of living environments for subjects suffering from mental illness.

Subjective self-awareness — Does it make sense that the subject’s self-awareness (his narrative or personal hypothesis of the functional relations between emotions, behaviors and context) is different from the behavioral observation? Equally so, does it make sense that direct assessments of momentary emotions are divergent from the

²¹ A reward is always what a subject — here and now — perceives as rewarding. Assuming that the direct assessment of momentary emotions is genuine and can be reliably and validly recorded, the momentary emotion of an optimal feeling — “liking” — should also be a “rewarding” experience for that subject. Consequently, we expect that what a subject describes as a pleasant situation will be pursued. Also, the likelihood of engaging in those positive mood situations will increase. We must avoid becoming trapped in an etymologic discussion. The behavioristic notion of “reward” is different from its common sense meaning. The behavioristic operationalization is a crude approximation of the subject’s “like and dislike” perception in a situation where no adequate assessment is available. We cannot merely ask the subjects if they like their living environment from moment-to-moment and if we did we probably will not trust their responses. We believe, however, that the assessment of the IER — what the subjects like, not in an abstract/generalized way, but varying from moment to moment — is a valid and direct assessment of this condition. Its variability also fits the observation that no consequence has general reward value.

subject's narratives? To understand this, we postulate that a separate reality exists, which we call the "subjective self-awareness" (SSA) and which can be understood as personality, social desirability, or the narratives or the personal hypotheses the subject has about his behaviors and emotions. This picture does not always fit his actual mental state or current behavioral repertoire. Moreover, this subjective mental state, while definitely divergent from current emotions and behaviors, may even be unrelated to aggregated emotions and behaviors of an individual. On the other hand, the narrative is the premium reality for assessing mental illness in a number of situations and the general appreciation that the subject has of his life in general ("quality of life").

Subjects and their behaviors are not totally ruled by emotions. Most of the human behaviors are intentional. They originate from awareness. In the selection of behavioral options the subject tends to reduce the dissonance between his behaviors and the integrated picture he has developed of himself (Festinger, 1957). The subject will choose those behavioral options that confirm his personality or his or her current mental state. Contrary to behavioral changes ruled by the operant learning paradigm, this process is intentional and will only match the mental state of the current period when this has graduated into the subject's awareness. The behaviors, therefore, are not only explained by the momentary emotions, but also by a feedback mechanism from the level of retrospective SSA. Because the subject builds his self-awareness partly upon the observation of his own behavior, this process consolidates the subject's existing self-perception and opposes the percolation of the information of momentary emotional states through behavior where it can be picked up and incorporated into the subject's integrated view of himself. Because the phrasing of those relations creates changes the narrative consolidates the subject's "personality" (Nuttin, 1974).

The changes in behavior reflect globally how a subject responds to periods of emotional turmoil or mental illness. Therefore, behavioral observations are more sensitive to changes than retrospective self-assessments. These are usually assessed through the traditional instruments to assess mental states: questionnaires and clinical interviews. The retrospective self-assessments are slowly changing processes. Assessments of retrospective narratives would be less prominent in psychiatric diagnosis if those self-conceptualizations did not influence behavior and global mood through a feedback loop. The illness narratives can invalidate us. If not, we would reject the "stories" the subjects tell as an epiphenomenon and expect that the observation of non-fitting purposeful behaviors would automatically correct them. Unfortunately this does not happen and what is worst, the narratives create behavioral impairments (feedback procedure) that can even be unrelated to actual cognitions and emotions (IERS). As a consequence, psychotherapy on this general level is useful. In psychotherapy the power of narratives is applied to induce change in behaviors, cognitions and emotions ("Le récit qu'on fait de soi nous modifie" — Lacan, 1966).

To assess IERS the momentary assessments are the method of choice. The intermediate level of behavior can be a cost-effective approach to assess emotional changes but is an indirect assessment with limited validity. It is not sensitive to contextual influences and moment-to-moment variability. Finally, the level of self-awareness is the field of

choice to assess the more fundamental narrative level that is often the target of psychotherapeutic change.

Demographic change — The last and least sensitive reflection of mental state changes is the level of “global adaptation”, reflected in the subject’s demographic situation. Illness processes can lead to structural changes in the subject’s living situation. In the most extreme situations illness can generate a demographic drift.

While the effect of illness at such a level is a slowly evolving change process, the assessment of demographic variability can be a cost-effective instrument to assess illness-related changes in epidemiologic research. As an instrument in treatment effect research, however, it is totally inappropriate.

Integration: the time process of change

The whole model of change presented in this discussion is a working hypothesis that can be used for the selection of appropriate assessment strategies. It is highly hypothetical to make predictions about the interval length related to the different phases of the model. We assume that large intra-, or situational, as well as inter-individual differences exist. In a general, tentative and averaged way, the IER lasts for approximately 15 minutes. After longer intervals, the adequacy — reliability and validity — of the report of the immediate reaction is lost. Behavioral changes can be observed within a day. Global daily life routines such as doing household chores, relating to others and getting out of bed can already be observed some hours after the mental state changes. Changes in the subject’s self-perception graduate to the level of awareness over a period between one week to a month — sometimes even longer. Therefore, the retrospective instruments traditionally used in psychiatric diagnosis are not sensitive for the present state assessment. Finally, the globalized adaptation reflected in the demographic characteristics can only be observed and checked after a year or more. Therefore, the assessment of changes in chronic illness will not be measured sensitively when using parameters such as work and living situation. Individualizing interventions should rely on the most sensitive and nuanced observations.

The development of hierarchies among cognition, affect and behavior is not new²². In the “high involvement hierarchy” — developed primarily in advertising — cognitions about an object are followed by the formation of affect toward the object, which then results in behavioral intention or expression. In low involvement hierarchy mere exposure to an object creates low-level cognitions but beliefs, attitudes and affects are only formed after experiencing the situation (Fishbein & Ajzen, 1975; Ray, 1982). In an important article Feldman & Lynch (1988) warned that the interplay between the structure of mental state genesis and the retrieval process in assessment can lead to “self-generated validity”. They argue: “By shaping the nature of the computational and retrieval processes by which answers to survey questions are generated, the measurement context can increase and decrease revealed correlations among those constructs” (p. 424). Hettema’s “Open System Adaptational Model of Personality” describes the relation between behavior and behavioral characteristics and personality

²² Post (1992) even presented a paralleling model of stress-related neural mechanisms.

(Hettema, 1979). He argues that personality traits do not predict behaviors deterministically (see also Mischel, 1968) but that the behaviors should be understood as emerging within an open system in which, among other influences, environmental cues and personality traits play an important role. Later, he presented a study that confirms that traits are constructs of the specific assessment mode and that rating scales constitute the appropriate way to measure them. Self-rating scales are particularly suited to reflect the intentional aspect of individual behavior including the goals of the behavior as well as the means to reach those goals. Our working model of ESM data, although different from the models discussed by Feldman & Lynch and Hettema, bears on the same type of relations: change assessments evolve with differential temporality for different targets. As Campbell & Fiske (1959) have done previously, they warn us that sometimes the type of assessment procedure — not their content — affects the validity ratings and therefore the use of scaling instruments. We should, therefore, be cautious when planning an assessment strategy and carefully select the research method in relation to the research target.

Chapter 4: Technical Note: Devices and Sampling Procedures ²³

In this chapter we will discuss technical issues related to the application of the Experience Sampling Method (ESM) both in clinical practice and for research purposes. The aim is to show how rather straightforward and seemingly uncomplicated choices between technical alternatives and procedures may affect the quality of the observational data that will be collected. ESM links experience rating to environmental cues. Due to the attractiveness of the research questions that can be assessed with this methodology we can easily overlook the basic technical requirements of such sampling strategies in natural environments. This chapter highlights the possible consequences of these choices for the quality of the observational data.

CHOICE OF THE SAMPLING DEVICE

A sampling device that meets all the needs of researchers does not exist. The researcher who plans to buy or hire signaling devices faces a difficult investment decision. The starting point of the decision making process should be a common sense analysis of the purpose of the study and the nature of the target population. A pilot run with different devices before making hardware investments is recommended.

General considerations

The shape, size and use of the ESM device should create minimal disturbance for the subject and his environment. The device should allow a flexible applicability as well as full control over the sampling process by the researcher. Ideal devices should be:

- *Non-reactive*— The goal of ESM studies is to collect data with ecological validity. The ideal sampling device should minimize the reactive effect that the data collection process has and lend itself to the feeling that nothing unusual is happening in daily life of both subjects and the persons in their environment. Reactivity can best be minimized by small, reliable and inexpensive devices that emit unpredictable signals and are fully operational within a range of environmental constraints.
- *SIZE*: small devices create less reactivity. Large devices require constant attention and can not easily be carried by the subject which results in misplacing and forgetting of the device and therefore data loss. Large devices also increase the chances of unwanted reactions of the environment. Such cumbersome beeping devices are no real problem when the sampling environment is relatively small or when the technological intrusiveness of the device is exploited to enhance

²³ Originally published as: Delespaul, Ph. A. E. G. (1992). Technical Note: Devices and Time Sampling Procedures. In M. W. deVries (Eds.), *The Experience of Psychopathology* Cambridge, UK: Cambridge University Press.

compliance or reporting accuracy and induce behavioral change, for instance, in therapy (Nelson, 1977). Considerations are many, e.g.; a men's wrist-watch may attract the attention of others when carried by a women, while they, on the other hand, may hide larger devices in their handbag. In contrast, when sampling white collar workers at their office we can consider beeping through the electronic mail system of the desktop computer; a large device is no problem.

- **RELIABILITY:** The researcher should be convinced that the device will signal the subject in the way intended by the research protocol. The device should not be an encumbrance or fragile. If the subject has to keep the unit operational by repeated manipulations (e.g. changing batteries) this will have an untoward effect upon the subject's daily life and the quality of the observations. Many high-tech beepers available on the market are not appropriate for use in ESM studies, because the conditions of daily life — such as extremes of humidity and temperature — may compromise the operation of the equipment.
- **EXPENSE:** Inexpensive devices or “cheap” looking ones are best. An expensive device tends to create over-concern in the subjects. Cheaper devices also relax the researcher because they are less likely to jeopardize the research budget when lost, stolen, smashed or crashed.
- **UNPREDICTABILITY:** The random or unpredictable nature of the beep can avoid anticipatory behaviors and thoughts and therefore minimize reactivity. Random signal generation also has repercussions for analysis (see “sampling process”).
- **SUBJECT CONTROLLABILITY:** A beeper may be too loud for people working in a library but no problem in a market place. Accordingly, a volume control for the beeping signal is useful. On the other hand, pre-programmed devices should have a limit in subject control. Particularly, the subjects should not be allowed to inspect or change the beeping schedules.
- **Flexible** — Because we cannot anticipate all sampling conditions of future research we should use flexible devices and avoid those that are restricted to specific sampling environments or sampling schedules. Programmable devices — downloaded from computers — are best suited in most circumstances.
- **Verifiable** — Researchers should generally have access to: the time on which the beep was emitted (schedule); and the time the subject answered the beep. When analyzing experience patterns over time we need to know when the subject actually received the signal (not when he should have received it) and how much time it took to respond to it. The delay between beep and response may distort the quality of the reports (Ericsson & Simon, 1984). Therefore, “real-time” devices with pre-programmed (or stored) sampling schedules are best. In this situation the “clock” of the researcher and the clocks of the subjects should be synchronized (Chapter 6B). However, in studies where only reported frequencies of activities and experiences are the object of research, knowledge of the actual signal time is not crucial.

Sampling devices

Early devices — Early mechanical devices that were primarily used in observation studies, signaled subjects at quasi-random intervals. The first generation of compu-

terized pocket instruments were developed for the “sampling” of a representative set of behaviors and thoughts. They do not need to link the beeps to a real time clock. The random timer used by Hurlburt (1990) is a good example. A creative alternative is the use of a standard digital watch (Brandstätter, 1983, 1991). These methods, however, do not meet ESM requirements because often the beep can be anticipated by the subject. Furthermore, the actual beep moment is not recorded which makes it difficult to check if the subject actually responded directly after the beep.

Real-time devices — Most current ESM devices use the signal generating techniques of modern micro-technology. They can be summarized in three groups: radio-transmission based systems; pre-programmed systems; and small scale computer systems.

- Csikszentmihalyi & Larson (1984) and deVries et al. (1984) have used “doctor pagers” linked to a telephone network. A host computer with a modem generates the beeps and phones the pagers. This allows the use of flexible time schedules, limited only by the options available on the controlling software on the host PC. A diversity of paging devices are currently offered by telephone and electronic companies, but most of these devices give overloud signals, that disturb the environment. One interesting exception is a vibrating buzzer. This signaling method is probably the best option to meet the non-intrusive intentions of ESM: only the subject and not his environment will be aware of the signal. However, pagers are expensive, usually have to be rented, and their operation requires a computer as well as repeated telephone connections which can increase research costs. Furthermore, signals may be lost because of the restricted working radius of the transmitter: the presence of buildings and hills in the environment may obstruct the radio signals. Another important handicap is that one can not be sure that the telephone connection was made at a specified time. Sometimes you are only able to dial into a switchboard that takes care of your message as soon as a radio channel becomes available. Also, telephone lines may be overcrowded and the numbers may need to be redialed repeatedly leading to significant delays or signal loss. We have found that — when sampling more subjects simultaneously over conventional lines — the technical limits of the local telephone switchboard may be reached. Radio transmitted systems have, however, important advantages. They are specially designed to signal people discretely. Moreover, additional signals may be requested simultaneously by an external observer. If something interesting happens related to the target event he can request an additional report by dialing the pager’s number.
- Delespaul & deVries (1987) and deVries et al. (1988) used SEIKO RC-1000 and RC-4000 wrist terminals and Alliger & Williams (1993) CASIO BGP-20 MultiPlanner wrist-watches²⁴. These watches do not generate a time-series themselves but execute schedules prepared by PCs. When the current time matches a pre-programmed moment, the watch will beep and, if necessary, provide information on display. In situations where the activity patterns of the subjects are known, you can intensify

²⁴ Other researchers have used pocket databanks as a signaling device (e.g. CASIO DBC-60). While this is less flexible than a watch — you can forget to carry it with you more easily — it is an alternative to be considered.

the sampling in moments that are more interesting for the research questions, for instance, during working hours in a study of work related stress.

Because those watches are commercial products targeted for a larger market they are relatively inexpensive. The fact that the device is also a watch, solves the problem of the synchronization between the researcher's clock that generates the beeps and the subject's clock and maximizes the likelihood that it will always be carried by the subject. A variety of sampling schedules can be executed, limited only by the memory of the watch and program options on the time-schedule-generating PC. The device is a men's wrist-watch (normal size), therefore the portability is optimal, while the minor aesthetic inconveniences for women can be dealt with using a pocket version of the SEIKO RC-4000 (the RC-4400). The sound, a digital beeping signal, is often heard in modern society and causes little commotion in the environment. For people living or working in noisy places or for elderly with hearing problems the beep may be insufficiently loud. The SEIKO watches are also sensitive to humidity which may affect the reliability.

- Recently, laptop, hand-held or pocket computers that can be used for sampling as well as data recording have become commercially available. These more complex devices allow a new way of data gathering, for instance, branching to additional questions or dynamically intensifying the sampling frequency when specific events are detected (Van Egeren & Madarasmi, 1988). Also, they minimize the cumbersome data-input work and offer a better reliability control — we can record when the subjects actually answered and how long that took. Original instruments with a full keyboards were too large and bulky to carry (e.g. the Epson FX-20; Cambridge Z-80). Later devices often had interfaces that were too complex to handle for most subjects (e.g. Psion Organizers 2 and 3, Sharp Wizards and IQ-8300 and OZ-9500/9600 series, Casio Boss, HP-95LX, Atari Portfolio). Recently, new instruments have been developed such as the Apple Newton and the Sharp "Message Pad" that allow a computerized paper and pencil analog. Now, well-designed user-friendly software can avoid most of the problems related to the original set of instruments, but such devices remain expensive, rather bulky and need additional manipulations such as changing batteries or repeated cumbersome procedures to download data due to the memory limitations or for back-up safety.

An interesting development is the combination of the ESM and the sampling of biophysiological data (e.g. Thakor et al., 1989, Meldrum, 1988a,b; Nicolson, 1992) for which new hardware developments are planned. The "Journal of Ambulatory Monitoring" is devoted to following these developments. The portability, price and standardization of those devices is still a problem. We believe, however, that medical applications using ambulatory monitoring techniques combining experiential and contextual information with physiologic data will be an important future trend.

Summary

Selecting a sampling device is not easy. It will depend on the focus of the study, the available hardware at a specific time, and on the research budget. Table 4.1 summarizes the elements that can be weighted. The assessments are indicative. We evaluate

if the devices are reactive (induce change in the observed subject through its use); linked to real time; flexible in the different applications and finally allow control over the beep and response times by the researcher.

Devices	Non reactive	Real time	Flexible	Allow control	Description of relevant elements to be weighted
Egg-timers; random beepers;...	-	-	-	-	<ul style="list-style-type: none"> • cheap; • if the observer knows the time of beep we can not use it for self-observations; • only when no real-time interests exists: we do not need to know the exact time of observation;
Radio transmission devices	-/+ noise	+	+	-	<ul style="list-style-type: none"> • for a small number of subjects (expensive); • in a restricted area; e.g. "doctor pagers" on hospital grounds (building adapted for radio transmission ?); • sampling (time budgets) / not time linked modeling; • concurrent external observation (remote signaling).
Pre-programmed devices	+	+	+	+	<ul style="list-style-type: none"> • for large samples (cheap); • sampling (time budgets) and time linked modeling; • when control of real time is needed (e.g. interactions in family systems); • for total free mobility (long distance,...);
Laptop computers	-/+ size	+	+	+	<ul style="list-style-type: none"> • for small number of subjects (cost-factor); • branched questioning (on computed events); • concurrent signal monitoring (EEG, temp,...) • direct data-input (cost-effective) • not user friendly (cumbersome keyboard)
Palmtop computers	+	+	+	+	<ul style="list-style-type: none"> see "laptop computers" • pen based systems offer a true digital paper & pencil analog (user friendly interface)

Table 4.1 Summary table for device selection (- : problematic; + : covered problem).

CHOICE OF THE SAMPLING PROCESS

The sampling process depends upon the purpose of the study. If the intention is to compute stable assessments of the frequency and duration of events it is essential that all moments within the target period have an equal chance of being selected at a beep. Representativeness may only be realized with *random* sampling schemes. If the assessment of an underlying time dimension for, say cortisol (Nicolson, 1992), is the object of study it is important to sample with *fixed* interval lengths. Once the distance between beeps is selected, randomization around those moments is possible, thus still meeting the criterion of unpredictability for the subject.

Sampling frequency

Sampling frequency, signal density and the between-beep-time-interval are interchangeable notions that characterize the beep schedule. Theoretical and pragmatic considerations related to different options will be discussed.

Theoretical approach to sampling frequency — For an appropriate choice of the sampling frequency two questions are important:

- *HOW MANY* observations are needed? The answer to this question depends upon the frequency and duration of the target event. Fewer beeps are required for the assess-

ment of highly frequent, long lasting events, in contrast to rarely occurring, short events. The precision in estimating event-duration is limited by the chosen sampling interval, e.g.: sampling every 90 minutes gives duration assessments in multiples of that interval (shorter than 90, 180 or 270 minutes).

- What is the *MINIMAL BETWEEN EVENT INTERVAL* of interest? Interval length depends on the answer to the following two questions:
 - a) What periodicity is expected in the phenomenon under study? Ideal, theory-based sampling intervals must be half the period-length of the sampled process, e.g. when we expect a 1-hour cycle we should sample every 30 minutes.
 - b) What is the expected time frame of an event-effect? When we intend to describe how the subject recovers from a target event, say a panic attack, we have to sample in a frequency corresponding to that expected duration. We should, for example choose intervals which are half the recovery time of anxiety.

Practical factors in selecting sampling frequency — ESM is a very demanding assessment method. It is often impossible to sample subjects in the ideal frequency range because it would exhaust them or have a drastic impact on the sampling environment. A sampling procedure should, therefore, seek a compromise between compliance and sampling requirements. In Maastricht, we use beep intervals of 90 minutes producing 10 beeps per day between 7:30 and 22:30. This interval was initially selected based on the basic rest/activation-cycle of human behavior (Minors & Waterhouse, 1981; Daan & Gwinner, 1989), but pragmatic considerations weight more heavily.

Our “rule of thumb” for the sampling density is based on 10 years of experience with a large number of populations. Assessing subjects for one week makes 10 beeps each day an acceptable rate, giving 60-70 observations for each subject. When sampling is carried out for shorter periods of time, higher sampling frequencies may be possible. If the research topic requires sampling over longer periods (e.g. 3 weeks) we advise not to cross a 6 beeps/day threshold. Pilot studies showed that 6 beeps/day were possible for 10 weeks. When it is only feasible to sample less than 4 times a day, randomization becomes meaningless and within day or contextual variability becomes difficult to detect. Of course, compliance may be influenced by other factors such as the use longer or shorter questionnaires — do not expect too much from this — and by the use of incentives or the additional recording of physiological parameters.

Parameters for the sampling process

An unlimited number of sampling schedules can be defined. All these different schedules have a specific influence upon the nature of the data and we therefore strongly recommend — for the purpose of replicability — to describe the schedules that are used by reporting the schedule type, the between beep interval, the sampling window and specifying the additional sampling characteristics. We also advise to pursue conformity with schedules that have been used in earlier ESM and avoid exotic variations.

- *type*: Schedule types will be discussed at full length in the next section. We distinguish two: fixed or random sampling schemes. In fixed schedules the time interval between two consecutive beeps is always the same; in random interval schedules the time span between beeps varies according to a random rule.

- *interval*: the average period between two beeps. Intervals are of fixed length in fixed schedules but not in random schedules and it is, therefore, appropriate to provide reports of the algorithm (e.g.: linear random or random normal deviates — see later), the mean and standard deviation of the intervals, e.g.: 90 ± 30 minutes.
- *sampling period*: period over which the sampling took place. E.g., a day; a week; a month. If the period is shorter than 1 week, the report should contain a mention of the starting weekday. When sampling over longer periods, the starting day, season or relation to the therapeutic and other processes should be reported.
- *period restrictions*: Sampling of experience is necessarily restricted to waking hours: for instance, between 7:30 - 22:30 (weekdays) and 9:00 - 23:30 (week-ends) for adolescents or between 12:00 - 15:00 hours for heroin addicts and prostitutes (Kaplan, 1992). It is also possible to alternate day periods with frequent, infrequent or no sampling. Examples are an intensive “saturated sampling” of late afternoon (kids come home from school); or sampling mental patients with 30-minute intervals during industrial therapy activities while using 2-hour intervals the rest of the day. We have used such strategies with good result.

The researcher should be aware that changes in the systematic sampling structure may lead to complications for the analyses of the data and the interpretation of the results. ESM data are very complex and we strongly recommend the use of standardized sampling schemes. With the availability of user friendly software that offer a large number of alternatives on the host PC that is used to download the schedules, or when “smart” devices are used that can generate their own sampling schemes, it is increasingly tempting to generate different samplings for each subject or switch rationales dynamically from moment-to-moment in relation to occurring events. Such strategies can be defended methodologically — maximizing the randomization effect or optimize data-collection in salient periods — but for now it is best to restrict the number of sampling schemes used until improved statistical capacities for complex time related data analyses become available.

Variation in types of sampling series

We differentiate two basic types of sampling schemes: fixed and random schedules.

Fixed schedules — In fixed schedules beep-moments have a regular periodicity. As a consequence, the different moments in the target period do not have the same probability to be selected as a beep. Probabilities range from 100% after the fixed period of time to 0% between the beeps. Fixed schemes are advisable whenever the use of a traditional time related statistical analyses — time series analysis, Markov chains — are planned. Fixed schedules, however, result in poor computations of time budgets. Assessments of events occurring in smaller intervals are also lost. As a consequence, modeling of experimental processes is restricted by the sampling interval and — because two points always define a line — three observations and thus periods of two intervals. In practice, fixed intervals are usually too long for most processes. Intervals of 90 minutes can only describe processes over 3 hours while the effect of a hassle or the recovery from an anxious moment may take less time.

Most importantly, with fixed schedules the periodicity in the signaling pattern is easily recognized by the subjects (and their environment), resulting in anticipatory behaviors, thoughts and emotions. Although reactivity can not be avoided in this type of sampling, small random variations around a fixed point can minimize the reactivity.

Figure 4.1a Fixed interval time sampling schedule: standard fixed interval time-series.

Figure 4.1b Fixed interval time sampling schedule: fixed interval with random normal deviate-variations.

Figure 4.1c Fixed interval time sampling schedule: fixed interval with variations of fixed probabilities.

Figure 4.1 shows the standard fixed schedules. Reactivity may be minimized by disguising the fixed time-series, starting each day with a different (random) starting time and using fixed intervals with unexpected length (4.1a). For example, periods of 52 minutes starting Monday at 8:23 and Tuesday at 9:07. Figures 4.1b and 4.1c are modifications of the basic fixed schemes. Here, the actual beep will occur at a random moment within a period around the fixed-scheduled-beep-time. The algorithm used

for the randomization in picture 4.1b is based on random numbers drawn from the standard normal distribution. When using this procedure the probability that a beep will occur at a certain moment will be highest when the moment is nearby the fixed-scheme-beep-time. In picture 4.1c, however, the probability for the selection of a moment is equal within the sampling range. Outside this range the probability is zero.

When a schedule contains periods in which, due to the algorithm used, some moments can never be selected, or are selected with a different probability, the computation of time budgets becomes very difficult and sometimes impossible. Because, however, the fixed periodicity allows the use of the more sophisticated statistical techniques, this kind of schedules are relevant, especially when we intend to model processes over time. Recent statistical techniques such as multilevel random regression techniques allow the modeling of data in real time. If expertise and software are available to apply these techniques we advice to choose random schedules.

Figure 4.2a Random interval time sampling schedule: standard random interval time-series.

Figure 4.2b Random interval time sampling schedule: stratified random interval time-series.

Random schedules — Random schemes are sampling schedules in which each possible moment of the target period has the same probability of being selected as a beep moment. This is a prerequisite for computing time budgets. Randomization techniques further reduce the likelihood of anticipation and lead to less unwanted effects — changing of behaviors, anticipatory thoughts. Figure 4.2 presents two schedules. In the truly random schedule (a) a random number of beeps will be selected each day. The intervals between beeps may vary widely. We do not advise this schedule because the long between-beep interval may demotivate the subjects. The stratified

random scheme (b) generates one or more beeps within each time block of the target sample period. Each beep has the same probability of being selected. Each time block will also have the same number of beeps. With the stratification shorter sampling periods can be sufficient to represent the whole period.

When sampling experience, particularly thought content, it is necessary to use schedules which signal the subject with a range of uncertainty about the beep — approximately 20 minutes will suffice. In our experience, this can be achieved by introducing randomization in fixed schedules. However, even then anticipatory behavioral change will occur. We therefore conclude that stratified random series are best if the sampling period is rather short (e.g. 1 week, 10 times per day), true random schemes can be advised for longer sampling periods (one or two months with six beeps per day) while fixed interval samplings — with slight random variations — are possible in high density assessments over very short periods (one day 40 times).

Summary

Selecting a sampling procedure is independent of the available hardware and cost issues and depends entirely on the research goals. A number of options have been discussed. The pros and cons of the sampling schedules are summarized in Table 4.2.

Description	Fixed series	Random series
• Reduced reactivity no anticipating behavior or thoughts (subject and/or environment)	only in mixed series	+
• Stable interval length requirement of high-tech statistical model-building	+	only in long series
• Assessment of time budgets	-	+
• Assessment of event duration	only for long events (> interval)	only for the most frequent events
• Assessment of time patterns (e.g. response decay)	only for long time patterns (> 2 x interval)	only in long series

Table 4.2 Overview of the effects of fixed and random ESM sampling schemes.

Chapter 5: Reliability and validity ²⁵

In response to dissatisfaction with traditional retrospective, cross-sectional approaches to collect information — which may provide a limited and biased view of a person's functioning — new measurement techniques have recently been developed. While there is no single nomenclature describing these new approaches, the term Experience Sampling Method (ESM) has been applied to the sampling of representative moments in course of a subject's life. ESMs are assessment strategies of momentary phenomena, linking subjective experience to descriptions of context and optionally, physiological indicators (see deVries, 1992). Other related sampling techniques of moments do not include all the types of information or the specific sampling procedures that define the ESM. The reliability and validity issues discussed in this chapter apply to this whole family of assessment instruments. We will use "ESM" for the sake of brevity. The observational procedures of this family of new assessment instruments use a different approach to the data collection in comparison with previous cross-sectional and retrospective methodologies. First, the approach shifts the site of measurement from the investigator's location, for instance, a medical office or laboratory, to the *subject's natural everyday environment*. Second, in contrast to retrospective assessments, the approach is focused on *immediate* activities, thoughts and emotions — what is occurring at the moment as opposed to a summary of previous activities and mental states. Third, the prior two characteristics are usually combined with *multiple assessments*, because investigators utilizing the approach are usually interested in dynamic fluctuations and interactions of the phenomena they are studying.

We believe that ESM offers exciting new ways of understanding human behavior. Providing a justification for use of the methodology and reviewing the encouraging results of studies that have employed the method (see, for example, deVries, 1987, 1992) is not, however, the focus of this chapter. The purpose of this chapter is to discuss issues of the reliability and validity of ESM methods, because only when the methods have been shown to have adequate reliability and validity will they achieve widespread acceptance. An introductory chapter on ESM psychometrics is necessary because the application of traditional reliability and validity measures — at least the ones common to behavioral and medical science researchers — appears to be problematic for much of the data collected in ESM studies. Two other publications discuss specific ESM aspects of reliability and validity not covered here. In Chapter 4 I discussed reliability of ESM assessment strategies as far as it can be optimized through specific choices of devices and sampling procedures. In Chapter 7 I will address validity pitfalls in the formulation and analysis of ESM research questions.

²⁵ This chapter is written in collaboration with Arthur Stone and has the following reference: Delespaul, Ph. A. E. G., & Stone, A. A. (1993). *Psychometric considerations of sampling moments* (unpublished manuscript). Maastricht: Rijksuniversiteit Limburg.

Researchers in the ESM field have acknowledged the inadequacies of traditional definitions of reliability and validity, and sometimes reject these statistics. They often simply accept ESM data as reliable and valid, since ESM data is considered to be a “gold standard” criterion (e.g. Lousberg et al., 1993; Delespaul & Reis, Chapter 6D). It is our contention that the situation is more complex than this and that researchers using these methods need to pay more attention to these reliability and validity issues. We argue three points in this paper: (1) that the content and procedures of the information being collected with ESM methods influence the appropriateness of assessing reliability and validity; (2) that there are operations on the raw ESM data, for example, ways of combining basic level ESM reports, for which traditional reliability concepts are applicable; and, (3) that traditional approaches to the assessment of reliability/validity are required for certain types of ESM data and for variables that results from operations on ESM data.

We first present a very brief overview of reliability and validity and how they are typically applied in the behavioral sciences. Several ways of classifying ESM data useful for determining when and how reliability and validity methods should be applied are then discussed. The last section of the chapter describes how traditional reliability and validity concepts apply to ESM data.

TRADITIONAL RELIABILITY AND VALIDITY

Several excellent texts are available that present the concepts of reliability and validity (e.g. Anastasi, 1976; Cronbach, 1990; Suen, 1990; Peck & Shapiro, 1990).

Reliability — Reliability is a measure of reproducibility and/or consistency of the obtained assessment statistics. When a test is reliable, we refer to several different, but related, properties of the test. We may be referring to the test’s temporal stability; that a person’s score on an IQ test taken 1 month ago will be very similar to the same test taken today. Test-retest reliability is the typical metric used to index temporal reliability. Alternatively, we may speak of a test’s internal consistency; that the items making up the test measure the same construct. A typical measure of this internal consistency is coefficient alpha. Both of these constructs address reliability, but note that an internally consistent test does not have to possess the property of temporal stability. Finally, another type of reliability refers to how consistently raters score a particular task, for instance, coding of classroom behaviors. This is called interrater reliability and a statistic commonly used to measure it is the kappa statistic. Sometimes, specific properties are pursued in the development of an assessment method. When measuring stable characteristics (traits) of an individual, such as an IQ, temporal stability is crucial. However, an instrument aimed at measuring intervention effects, should be sensitive for the momentary state of the individuals. In the first situation test-retest reliability should be high, in the latter the interrater reliability is appropriate. Kraemer (1992) stresses the importance of acknowledging the reference period for the evaluation of test reliability.

Validity — Validity refers to the degree that a test measures what it is intended to measure. There are several kinds of validity and they may be ordered, to some degree,

in terms of the complexity of each criteria. Face validity refers to the degree to which a test's items "appear" to measure what the test is attempting to measure. This is a low level form of validity since face validity items may, for example, be interpreted by the people taking the test in ways that do not address the intended concept. At the other end of the spectrum is construct validity; it has the most demanding set of criteria. To achieve construct validity the test must correlate positively and with predicted magnitudes to related concepts and negatively with dissimilar concepts. These associations form what has been termed a nomological network and can provide strong evidence that the test is, indeed, measuring what the developers say it measures.

While it is generally possible to demonstrate the reliability of an assessment method to a large degree, validity can only be partially approached. Validity is subject to the purpose of the use of the test in a specific situation and therefore should always be re-evaluated. Furthermore, the association between reliability and validity is a lopsided one. A measure that is reliable may or may not be valid; however, a valid measure must possess some degree of reliability since an entirely unreliable measure cannot be valid²⁶. We will see that reliability and validity have interesting associations to ESM.

CLASSIFICATION OF ESM DATA

In reviewing reports of ESM data, we have identified issues that relate to psychometric procedures. There are certainly other ways of thinking about ESM data, but we have found these useful for determining which reliability and validity concepts should be applied. We refer to the basic unit of data (the "moment") as the "beep" level of data and the first three options are target to that level. The fourth option combines information over "beeps" — the "subject" level.

- *Raw data* — Raw data collected in ESM assessments and the problems related to aspect of self-monitoring in ESM methods. Many studies collect information reports of the *SUBJECTIVE* states of individuals, for instance, reports of mood, thought experience, perceived stress, and psychiatric symptoms such as hallucinations and anxiety. These data have few, if any, external referents for confirming their accuracy. In these same studies, other data is collected that may be called *OBJECTIVE*. The characteristic that qualifies reports for this label is that there are observable referents for the reports. Examples of such data include the activity that a person is involved in at the moment of the beep, whether or not they are in the presence of other people, and the type of place the respondent is in (home, work, car, etc.).
- *Post-hoc coding* — This issue represents an operation that has been imposed on the beep level data. ESM data is sometimes obtained in free-format (for example, asking subjects to indicate where they are, with who, what they are doing and

²⁶ This argument is traditional and logically sound, but not always true in practice. In a monitoring study of the effect of neuroleptics on psychopathology we found valid assessments — the differential effect of consecutive medication regimes could be evaluated — with interrater reliabilities of 0.20. The low reliability was compensated by repeated measures (Delespaul & Bak, in preparation).

what they are thinking at the moment of the beep in ESM). These qualitative reports are subsequently coded according to guides developed by the researcher.

- *Scaling procedures* — Investigators often ask subjects to complete a set of questions about a conceptual area. In the case of mood measurement, many adjectives with dichotomous or multilevel response scales are presented to subjects. Responses to multiple questions are then combined, either empirically (e.g. with factor analysis by Kraan et al., 1992; Kaplan et al., 1992; Dijkman, 1993) or on rationale grounds, into one or more scales (for example, positive and negative mood scales).
- *Aggregation of moments* — The last issue refers to how a researcher analyzes and reports his data, combining information of different “moments”. Analyses of ESM data are often done using this raw data material aggregated over typical situations or over time. For instance, if weekly patterns of some dependent variable are of interest, then an investigator might take the average of all beeps for each day of the week and analyze this summary variable. If diurnal patterns are of interest, an investigator could take the average of all beeps over all days for certain hours of the day (say, 9:00 through 11:00, 11:00 through 13:00, etc.) and use this as the basic unit of analysis. In a study of loneliness an investigator can aggregate over the “alone”- and the “not-alone”-beeps. While by aggregating beep level data in this way we lose information, there is nothing intrinsically wrong with the data analytic strategy as long as proper controls are instituted to insure that missing data, for example, does not confound the aggregates (see Larson & Delespaul, 1992; Jaccard & Wan, 1993).

These four ways of discussing ESM studies are not often found in the discussion of psychometric properties, but we will argue in the next section that they do have implications for reliability and validity. We should note that these qualities have at least partially been discussed elsewhere in other contexts. Stone (1981) discussed the objectivity and subjectivity of life event measures, and concluded that this classification scheme affected the conclusions that were possible from these studies. Aggregation of data has also been discussed in terms of reliability by Epstein (1979, 1980) and others, and certainly standard texts discuss scale construction techniques. Post-hoc coding of data and scale construction techniques are topics that have received considerable attention in behavioral observation studies and some of the conclusions from that literature are presented below. Despite this prior work, we feel that these issues deserve explicit attention in the ESM field.

RELIABILITY AND VALIDITY IN ESM STUDIES

Raw data

Whether or not self-reported data is verifiable, has implications for the evaluation of both reliability and validity. Data that are objective — the time, the activity, the person and place at the beep-moment — may be verified by either others with knowledge of a subject’s activities or by the investigator. In that case, a *consensual validity*,

namely, an agreement among several observers, may theoretically be achieved with sufficient expenditure of effort and resources. If all observers agree that the subject was with another person at the moment of the beep, then validity is assumed. Otherwise, the report is suspect in terms of its validity. Consensual validity is equivalent to interrater reliability: once a reasonable level of consensual validity is achieved, a report is reliable by definition. Thus, the distinction between reliability and validity blurs with objective ESM data. Other formulations of reliability — internal consistency and test-retest — do not appear to be relevant in the evaluation of objective ESM data. Since there is only a single indicator for the occurrence of an event or a salient aspect of it, there is no opportunity to evaluate internal consistency. Likewise, the concept of temporal stability does not apply to the sampling of moments, since the moment itself is the reference and variability over time is assumed. We note that this is also the case in the psychometric evaluation of behavioral observation data (Silva, 1993; Suen, 1989).

Subjective ESM measures, such as thoughts, experienced mood and mental states should be evaluated differentially. Unlike the aforementioned objective data, there are no external referents for these states. There are observables that could be associated with subjective reports, for example, distinctive facial expressions correlate with certain affects (Ekman, 1972; Izard, 1991) and eye and other body movements may correlate with hallucinations (Paul et al., 1987). However, those indicators are not considered very valid and the criterion for their validation is the assessment by the subject itself. Sometimes scaling techniques can be applied to combine multiple indicators of a subjective construct. For those, reliability can be assessed (see later). Usually, however, no multiple indicators are available: the item “I feel guilty right now” assesses momentary guilt, while “I feel anxious” assesses anxiety. These items have maximized face validity, but researchers should be aware that normal daily life vocabulary does not always match our theoretical constructs. A high score on “I feel depressed” does not reflect a clinical depression, and high “anxiety” is not a panic attack. Things get even worse with items reflecting a possible pathological or social non-desirable state. Subjects could be reluctant to give high scores resulting in invalid and skewed response distributions. To avoid this, a weakened formulation of the item can be chosen (not “depressed” but “sad”; not “angry” but “irritated”) or the item can be cued using a positive formulation (“I feel cheerful” to reflect (not) “depressed” or “relaxed” for (no) “tension”). In these situations reliability tests can be performed through a study that combines both formulations, herewith creating a multiple criterion situation. Furthermore, irrespective of the within subject variability, overall, depressive subjects should feel depressed more often than normal controls; a higher level of stress measured through an end-of-day measure should also be found in the ESM reports for that day; and assessments made in a “party”-situation should reflect higher affect than those made at school. This kind of multimethod/multisituation comparative studies can be used to assess ESM validity. Once validity is demonstrated, reliability can be assumed.

In summary, different validity and reliability standards should be applied to raw ESM data depending on the degree of subjectivity inherent in those data. Some investigators may feel that the raw data of moment assessments are more “accurate”

than data collected in other ways, and hence the rules for data coding are different. The fact that the data have been obtained at specific moments minimizes invalidation due to retrospective recall (Wheeler & Reis, 1991), but does not alter the need to assess reliability and validity.

Post-hoc coding

The application of reliability principles to coding of ESM data after they have been collected is not complicated. The coding procedure can be replicated by two raters and basically, interrater reliability can be assessed to be sure that the coding scheme devised for the data is adequately applied and that assessors are not drifting in their use of the rules over time. Sometimes, researchers assume that it is easier, more reliable and valid to allow the subjects to code themselves by presenting them with a restricted set of options. What the subject says about a specific moment is assumed to be accurate by definition. This procedure might solve some coding problems but it is impossible to control if subjects use the scheme in the way the researcher intended it or that they drift in their interpretation over time. Getting control over the subject's perception of the questions is very difficult (Sudman & Bradburn, 1982). Therefore we advise using open questions and code them afterwards.

Momentary assessments are embedded in the natural flow of the subjects daily lives and some familiarity with the subject and his living environment can optimize the coding task. When coding is done by the research staff, the information collected in the briefing/debriefing meetings is crucial to make adequate assessments. Codings are usually straight translations (places, persons and activities in a list) but sometimes they are globalized clinical assessments (e.g. the level of thought disturbance in thoughts). Irrespectively, all coding should be checked on validity. Coding recorded this way must be considered as raw data material and checked upon considering the comments described in the previous section.

Scaling

The main point considering reliability and validity of subjective ESM measures is that unlike the objective data discussed above, an investigator may apply a multiple indicator approach to assess subjective constructs such as depressive, psychotic or panic moments. Several items on a ESM assessment may be used to assess a specific subjective state, and when this occurs internal consistency reliability should be computed as they would with any multiple indicator application. The argument for using traditional internal consistency reliability methods and traditional validity methods for scales derived from ESM data is, in essence, the same as that mentioned for post-hoc coding. Scales derived from beep assessed data should not be immune to the same psychometric scrutiny given to scales formed from any other source of data. This applies to scales based on both subjective and objective data. However, test-retest reliability is not a required step. Just as with reliability, the assessment of validity with subjective ESM data reverts to standards used in the development of any psychological test, and therefore construct validity is desired. Again, even if the reliability /validity of

objective ESM data has been established, the reliability and validity of a scale formed from these data are not automatically conferred to the scale.

Scale construction is more difficult in the ESM because latent structures can be divergent for the different intended applications of the data: scales measuring properties of subjects might be different from those measuring the same properties at the moment-level. Considering the wide range of scale applications, each use should be validated.

(Temporal) aggregation

Although high interrater reliability confirms validity for objective beep data, new reliability issues are faced when investigators aggregate over many observations. While there may be excellent reasons for aggregating, the procedure does not imply that the aggregate score (usually an average of the points) is reliable. Considering that one of the reasons for using an ESM data collection strategy in the first place was to capture temporal variability, it is likely that an aggregate will not be very reliable. The level of reliability depends, to a large degree, on the number of observations included in the aggregate as Epstein (1979) has shown. In ESM situations, the occurring situations are not under experimental control. They just happen in the subject's life. Increasing the number of observations of infrequently occurring events may be difficult and sometimes even impossible. Other sampling strategies (for example event sampling) might be more appropriate. Furthermore, the number of observations are not balanced over subjects and over situations nor are the occurrences of the events under experimental control. As a consequence validity is threatened: the aggregate over a specific naturally occurring situation (e.g. mood in "alone"-situations as they occur on normal days) can be divergent from aggregates over experimentally controlled situations (mood in randomly — subject independently — occurring "alone"-periods). The possible misunderstanding of the nature of ESM data, as shown here, is probably the most important validity threat.

It is reasonable to compute traditional reliability statistics, such as test-retest coefficients, to confirm that an aggregate is stable. In the case of diurnal cycles, an investigator with multiple days of sampling might separately compute hourly aggregates on the even and the odd days. High test-retest reliability suggests a stable phenomenon, whereas low reliability coefficients suggest that other factors (such as contextual reactivity, unmeasured other influences or error) are affecting the data. Because unreliable aggregates will not correlate well with other variables, leading to non significant findings, it is essential that an investigator evaluates the stability of temporal aggregates.

SUMMARY AND DISCUSSION

We initially stated that ESM data had been considered by some investigators to fall out of the realm of traditional psychometric analyses as these procedures did not appear applicable to such data. Our discussion of the ESM data has shown that this conclusion is true for a certain class of data — namely, objective contextual descriptions, which

only require interrater reliability. Subjective ESM data have no external reference and therefore only validity assessments can ascertain reliability. Derived variables using post-hoc coding and scaling on the raw data material, as well as aggregation over moments, are all subject to typical and rather traditional psychometric analyses.

Data description	Reliability issues	Validity issues
<div><div>• Raw data</div><div><div>• objective</div><div>• subjective</div></div></div>	<div>interrater (external reference available)</div> <div>no check (no external reference available)</div>	<div>face</div> <div>consensual</div> <div>face</div> <div>through correlations of (dis)similar items</div> <div>through behavioral correlates</div> <div>through aggregation</div>
<div>• Post-hoc coding</div>	<div>interrater</div>	<div>see "raw data"</div> <div>construct (with other repeated measurement assessment strategies)</div>
<div>• Scaling</div>	<div>see "raw data"</div> <div>internal consistency</div>	<div>construct (with other repeated measurement assessment strategies and subject to its use)</div>
<div>• (Temporal) aggregation</div>	<div>see "raw data" subject of n</div> <div>(# observations/category)</div> <div>internal consistency</div> <div>test-retest (with stable interval)</div>	<div>construct (with other ESM -methods /cross-sectional instruments)</div>

Table 5.1 Summary table of reliability and validity strategies in ESM studies.

ESM is not a structured test, but a complex set of procedures applied with a divergent sets of items, using different beeping devices and sampling schemes and with different purpose in a large number of populations. Therefore, no general statements of reliability and validity can be made. From the different published reports on ESM, however, we can conclude that the method has been applied successfully in a large number of situations leading to interesting new conclusions, but also replicating a number of traditional findings. Most of those publications have demonstrated some aspects of reliability and validity. An extensive demonstration of reliability and validity aspects in all aspects of ESM following the lines presented above, has never been done. Such attempt would be of marginal use because such an extensive demonstration in one study would not generate another studies by definition. In Chapter 6 I will present some studies that we have conducted exploring ESM reliability and validity. We will also refer to the original publications (deVries, 1987, 1992; Annals of Behavioral Medicine, 1993) to avoid the false belief that our studies can prove validity and reliability of ESM strategies once and for all.

Chapter 6: Empirical Studies in ESM Reliability and Validity

ESM is assumed to be an assessment strategy that is sensitive to changes in daily life experience, activity patterns and context. To assess this assertion we need to know if ESM measures are reliable and valid reflections of daily life experiences and activity patterns. Considering the wealth of possible ESM designs and applications, it is impossible to draw definite and generalized conclusions about the psychometric merits of ESM (see also Hormuth, 1986; Csikszentmihalyi & Larson, 1987; Alliger & Williams, 1993; Delespaul & Stone, 1994). To accommodate for this complexity, we will often rely on meta-analytic techniques to assess ESM reliability and validity (see Petitti, 1994; Wolf, 1986; Hedges & Olkin, 1985; Rosenthal, 1984; Hunter et al., 1982; Glass et al., 1981). Meta-analytic statistical techniques are developed to integrate results from different studies that are usually implemented with different operationalizations and different sample sizes. Weighting techniques are central in that situation. This is not the case in meta-analyses of ESM data. Here, our aim is to give a general appreciation of the trends in a set of statistical tests that are computed on the same sample over different operationalizations of a specific concept. Such global appreciations are not possible by correcting p-values for multiple tests because this would only ascertain the significance of the individual tests and would leave us with a set of results that are not integrated. We, in contrast, need a summary of the tests — we need to know if ESM is reliable and valid in chronic mental patients — and this is achieved by testing whether the proportion of significant individual statistical tests is significantly higher than the p-value of the individual tests.

The different sections of this chapter illustrate empirically the ESM reliability and validity. Over the different sections we will discuss the following *reliability* issues:

- A. Test/retest reliability: ESM stability over time. A prerequisite for the use of ESM variables in evaluation research is a proof of ESM variable stability over time, when no intervention or event occurs in the interval between the assessments;
- B. Reliability decay by ESM response delay: the ESM reliability time-window. In ESM we assume that longer intervals between event and report cause a decay in report accuracy and as a consequence represent a threat to ESM reliability.

In the next three sections we evaluate the *validity* of ESM by comparing ESM results, with results obtained using a different type of assessment strategy:

- C. Comparing ESM time sampling and a traditional diary approach for the study of daily life activities;
- D. Comparing ESM time sampling and RIR event sampling for the assessment of social interactions;
- E. Comparing ESM time sampling and retrospective assessments of time budgets and experience in context.

In the last section (section F), we will discuss the possible implications of the development of ESM assessments for the implementation of psychiatric evaluative research. We will argue that ESM promises to yield sensitive measures of change that can be applied as outcome measures in clinical trials.

A: STABILITY OF ESM OVER TIME (TEST/RETEST RELIABILITY)

Subjects

In this section we will present data on the stability ESM variables in a sample of 15 chronic mental patients assessed twice over a period without specific intervention. Due to missing or incomplete data in one of the sampling periods, subjects had to be skipped, resulting in different n's for each analysis (range between 11 and 13). Most of the subjects were schizophrenics, currently under treatment in ambulatory or clinical facilities. The data were collected in a study that explored the use of the ESM method in rehabilitation practice in a population of chronic mental subjects. Our intention was to learn about salient ESM variables that could be useful for clinicians in their decision-making for interventions. The results of the study have been used to develop the feedback-protocol of ESM (see Part 4 and Appendix 7).

Method

The chronic mental patients were measured repeatedly over a period in which no specific interventions was provided. To ascertain whether systematic changes had affected the subject's mental state during the interval we tested the difference between both target periods at the group level. We expected no difference in the pre and post conditions. For the assessment of ESM-variable-reliability we computed the Pearson product-moment correlation coefficient for the means and standard deviations of the sampled experiences by subject. The same procedure was used for the time use indicators. To summarize the results we used a meta-analytic technique. Comparison for the test/retest reliability of sampled experiences and activities assessed using ESM was provided by a cross-sectional assessments of mental state (Brief Psychiatric Rating Scale/BPRS — Overall & Klett, 1964; Lukoff et al., 1986; Ventura et al., 1993) and social adaptation (Social Functioning Checklist/SFCL — Appendix 5).

Results

Assessment of the test/retest interval

The test/retest period is 6 months (181.45 days) with a standard deviation of almost 3 months (85.41 days). There was no difference in the number of days the subjects participated in the ESM over the two sampling periods ($t_{(10)}=0.90$ n.s.) and drop-out — resulting in a smaller number of days in the study — was not a subject characteristic ($r=-0.13$; n.s.). The number of beeps a subject responds to, was related over time ($r=0.95$; $p<0.01$). However, the number of beeps responded to in the second period was significantly lower in the second period (decrease of approximately 5 beeps from 42.5 to 37.6; $t_{(10)}=-3.84$; $p<0.005$).

No significant changes in averaged moment-to-moment experience occurred over the interval between the assessment periods (moods — $z=0.559$, n.s. — and psychopathology — $z=0.00$; n.s.). The same trend was found for moment-to-moment variability by subject (moods — $z=0.839$, n.s. — and psychopathology — $z=0.00$; n.s.). A marginal significant trend was found for the item “cheerful” which was significantly different over time both for means ($t_{(10)}=-2.10$, $p<0.10$) and variability ($t_{(10)}=-2.59$, $p<0.05$). Six months later the subjects were more cheerful and less variable for that emotion from moment to moment. We also compared the differences in time use patterns for our subjects between both periods. No significant changes were found between the assessment periods. Overall, the subjects as a group spent equal amounts of time with different categories of persons ($z=0.00$; n.s.), at different places ($z=0.00$; n.s.) and during different activities ($z=0.00$; n.s.).

Summarizing, no systematic changes in sampled mental state and activity patterns occurred between both periods. Considering the fact that this is a chronic mentally ill population with no specific systematic treatment this was also expected.

Test/retest reliability

The position of subjects relative to each other based on summaries of sampled experience are stable over time (r 's range between 0.60 and 0.90 — Table 6A.1) The extent of the correlations is impressive considering that the subjects are chronic mental patients, primarily schizophrenics, who are currently under treatment. Overall, the number of significantly reliable items is satisfactory, both for the mean levels of sampled experience ($z=3.39$; $p<0.001$) and for the variability within subjects from moment to moment ($z=3.00$; $p=0.001$). The standard deviations by subject, however, were less stable for the psychopathologic items ($z=1.33$; $p=0.09$). In contrast to the high test/retest stability of sampled experience, the time use indicators are not reliable ($z=0.07$; n.s.). This, however, can be caused by the fact that the time interval between the assessment periods results in observations made in two different seasons in a very homogeneous population.

Sampled experience	$r \bar{x}$	$r_{s.d.}$	Time use patterns	$r \bar{x}$
Cheerful	0.68 *	0.86 **	Persons	[-0.12 ...0.28] n.s.
Satisfied	0.67 n.s.	0.83 *	Places	[-0.30 ...0.22] n.s.
Anxiety	0.81 **	0.94 **	Activities	[-0.28 ...0.20] n.s.
Demotivation	0.86 **	0.65 *	• Meals	0.63 *
Lonely	0.76 **	0.87 **		
Suspicion	0.73 *	0.89 **		
Derealization	0.75 *	0.80 **		
Auditory hallucinations	0.83 **	0.30 n.s.		
Visual hallucinations	0.79 *	0.63 n.s.		
Fear of losing control	0.61 n.s.	0.60 n.s.		
Meta-analysis	8/10	7/10	Meta-analysis	1/18
	$z=3.39$ ***	$z=3.00$ ***		$z=0.07$ n.s.

• $p<0.10$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$.

Table 6A.1 Test-retest reliability scores (n=11) after 6 months.

Comparing sampled prospective and cross-sectional reliability

We compared test/retest stability of aggregated sampled moment-to-moment experience with conventional cross-sectional assessments of mental state and social adaptation. None of the individual BPRS item-scores ($z=-1.11$; n.s.) or the summated BPRS-score for illness severity ($t_{(9)}=0.57$; n.s.) changed over the course of the assessment period. This confirms the overall status quo over time in this sample of chronic mental patients. For social adaptation (SFCL) we found no significant changes over the assessment interval for individual items ($z=-0.51$; n.s.) and summated SFCL-items reflecting general adaptation ($t_{(9)}=-0.43$; n.s.). Here also, no systematic trends were found.

This is a good situation to assess test/retest stability. Both psychopathologic assessments ($z=2.75$; $p=0.003$) and social functioning ($z=1.86$; $p=0.03$) have test-retest stability over the 6-month period. However, the cross-sectional assessments of moods (anxiety, depression, guilt) — which were the most stable items in sampled moment-to-moment experience — show the lowest test/retest correlations of all the assessments in the cross-sectional instrument we used.

Mental state (BPRS)	r	Mental state (BPRS) cont'd	r
• Somatic concern	0.48 n.s.	• Motor retardation	0.64 n.s.
• Anxiety	-0.23 n.s.	• Blunted affect	0.68 **
• Depression	0.46 n.s.	• Tension	0.40 n.s.
• Guilt	0.34 n.s.	• Mannerism/posturing	0.68 *
• Hostility	0.84 **	• Uncooperativeness	-
• Suspiciousness	0.84 **	• Emotional withdrawal	0.44 n.s.
• Unusual thought content	0.74 **	• Suicidality	0.56 n.s.
• Grandiosity	0.99 **	• Self-neglect	0.66 n.s.
• Hallucinations	0.76 *	• Bizarre behavior	0.43 n.s.
• Disorientation	-	• Elevated mood	1.00 **
• Conceptual disorganization	0.13 n.s.	• Motor hyperactivity	0.59 n.s.
• Excitement	0.95 **	• Distractibility	-
BPRS (illness severity)	0.48 n.s.		
Meta-analysis	9/24 $z=2.75$ ***		

* $p<0.10$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$.

Table 6A.2a Test-retest stability of BPRS cross-sectional scores ($n=11$) after 6 months.

Social adaptation (SFCL)	r
• General functioning	0.60 n.s.
• Work & productivity	0.85 **
• Independent living	0.80 **
• Direct social network	0.84 ***
• Remote social network	0.62 n.s.
SFCL (general adaptation)	0.78 **
Meta-analysis	3/5 $z=1.86$ *

* $p<0.10$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$.

Table 6A.2b Test-retest stability of SFCL cross-sectional scores ($n=11$) after 6 months.

Conclusions

A reliability study was performed on a small sample of chronic mental patients, primarily schizophrenic subjects, a population that is usually considered too chaotic to generate reliable self-assessments (Birchwood et al., 1994). The self-reports of those subjects are notably unreliable (Frith, 1992). This study shows this is not the case.

Even in this extremely invalidated population, ESM based characteristics of daily life experience are reliable assessments that are stable over a long period of time. We ascertained this reliability in a homogeneous small sample — all chronic schizophrenic subjects — over a long period of time — 6 months.⁴ Compared to daily life experience, the assessments of time use patterns with ESM were not reliable. This was caused by the fact that the study was performed in a homogeneous population. Everyone is “alone” and “at home” (hospital ward) most of the time. The other time budget categories have a low frequency, making their assessment using a small number of observations less reliable (5% is only one observation for a subject with 20 valid reports).

Mental state	% significant	Social adaptation	% significant
Moment-to-moment moods	80.0%	Time use patterns	5.6%
Cross-sectional assessment	37.5%	Cross-sectional assessment	60.0%
	$z=2.99$ ***		$z=1.83$ *

* $p<0.10$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$.

Table 6A.3 Reliability meta-analysis: prospective ESM compared to retrospective questionnaires.

We compared the test/retest reliability of sampled experience and cross-sectional assessments using traditional techniques such as one-time questionnaires and clinical interviews (Table 6A.3). The proportion of significant correlations were compared in both situations. We found a difference favoring the ESM ratings ($z=2.99$; $p=0.001$). The difference was most striking for the assessment of moods, which were rather unreliable in the BPRS. From a reliability point of view, cross-sectional assessments are adequate for the measurement of core symptomatology, but not for mood. ESM assessments, on the other hand, give reliable assessments of both mental and mood states. In contrast, ESM time use patterns were not very stable over time. Here, global adaptation, as measured by a cross-sectional questionnaire, outperformed ESM ratings ($z=1.83$; $p=0.03$). In defense of the ESM time budget ratings we will argue that social adaptation and time use are different topics and that ESM outperforms retrospective time use patterns (see section 6E).

In conclusion, we demonstrated that ESM mental state rates are reliable over time. We can now be more confident about relating ESM assessed changes in a clinical trial to the intervention itself, and not to other, uncontrolled, random events.

B: DEFINITION OF THE "ESM RELIABILITY TIME-WINDOW"

An important rationale for the use of ESM kind of methods is that we believe that the reliability of the assessment of an event can be enhanced by reducing the time between the observed event and the report. Daily assessments are better than weekly ones, event assessments are better than end-of-day ones. How soon after the event do we have to report? Is there a gain in moving from end-of-day reports to reporting within the hour? Should we reduce the interval even more? In ESM we assess moments and reduce the interval between the event and the report to a realistic maximum. Figure 6B.1 illustrates the different tasks involved in filling in the responses on the ESM-forms. Subjects need time to interrupt their activities. They might, for instance, be driving a car and have to park it first. Next, they have to use the booklets and a pencil. Finally, they need time to fill out the questionnaires. Those activities and the administration typically take 2 to 5 minutes. Because we have preprogrammed the watches we know when the beep went off. If the subject refers to the watch to check the time, the beep and response times will be synchronized. However, subjects can use a different time reference such as a church tower, a kitchen clock or the radio. By experience we know that compared to the sampling device this can give an error of up to 10 minutes before and after the beeping time. We, therefore, concluded that realistically the interval between the event and the report can be reduced to a maximum of 5 minutes before and 15 minutes after the signal.

In this study we will analyze whether reducing the response time interval to such a small period, makes a difference.

Figure 6B.1 Time reference of the ESM response sequence.

Subjects

The subjects used in this analysis are the same sample as in Part 3 (Chapters 9 to 11) of this study. A description of the population and subpopulation is given in Chapter 9.

Method

In this study we will compare the response quality of the answers given within the minus 5 / +15 minute interval (“IN”-beeps), with answers formulated up to 1 hour after the signal (“OUT”-beeps). Based on theory (Chapters 2 and 3) we operationalized the expected distinctions between both situations. We assume that the intensity of mental states will be different in a direction related to social desirability in the more distant answers. Moreover, a longer interval will lead to more aggregation and confounding caused by the subject’s self-narrative (trait related influences), resulting in less adequate assessments of the momentary emotions (current mental state). As a consequence we will observe reduced variabilities in the later answers. For this study we assessed the following null hypotheses²⁷:

- [1] H_0 OUT-beep answers show more socially desirable emotions;
- [2] H_0 OUT-beep answers have smaller variability.

Analysis will be performed with meta-analytic techniques testing the significance of the amount of significant tests.

Results

Tables 6B.1a and b summarize the analyses comparing the means (hypothesis 1) and standard deviations (hypothesis 2) of the moment-to-moment experience computed by subject for the IN- and OUT-beeps.

The first hypothesis can be accepted if the *mean* scores of the positive mood items (cheerful, satisfied and motivated) are higher (more socially desirable) in the remote condition (OUT-beeps), compared to the adjacent responses (IN-beeps). The inverse relation should be found in the socially less desirable items (anxiety and loneliness). None of those differences were significant at the 0.05-level. Contrary to our expectation “anxiety” and “motivated” were more socially desirable in the IN-condition. Inconclusive results were also found for the psychopathologic items. Meta-analytic tests confirm this. None of the computed proportions were significant. Therefore, the first hypothesis could not be supported. The reliable responses should not be restricted to the 15 minutes interval to protect ourselves from social desirability.

Mood		Paired t-Test	Psychopathology		Paired t-Test
Cheerful		$t_{(112)} = 1.43 \wedge$	Suspicion		$t_{(124)} = -0.14 \text{ n.s.}$
Satisfied		$t_{(62)} = 1.77 \cdot$	Derealization		$t_{(115)} = 0.49 \text{ n.s.}$
Anxiety		$t_{(120)} = 0.48 \text{ n.s.}$	Auditory hallucinations		$t_{(71)} = -0.69 \text{ n.s.}$
Motivation		$t_{(124)} = -1.66 \cdot$	Visual hallucinations		$t_{(61)} = -2.20 \cdot$
Lonely		$t_{(123)} = -0.81 \text{ n.s.}$	Fear to lose control		$t_{(71)} = 1.51 \wedge$
Meta-Analyses	$H_0: p=0.50$	$z = 0.32 \text{ n.s.}$	Meta-Analyses	$H_0: p=0.50$	$z = 0.32 \text{ n.s.}$
	$H_0: p=0.05$	$z = -0.51 \text{ n.s.}$		$H_0: p=0.05$	$z = 0.72 \text{ n.s.}$

$\wedge p<0.20$; $\cdot p<0.10$; $\cdot\cdot p<0.05$; $\cdot\cdot\cdot p<0.01$; $\cdot\cdot\cdot\cdot p<0.001$.

Table 6B.1a “Time-window”-effect: drift (average) in social desirable direction.

The second hypothesis assumes that the variability of the answers — a measure of sensitivity — will be higher in the IN-beeps compared to the OUT-beeps. Summing up

²⁷ A third hypothesis (H_0 : OUT-beep answers show more socially acceptable activities) was not assessed because we could not generate a reliable social desirability rating of the activity codes.

over mood and psychopathology items, 2 out of the 10 relations were significant at the 5% level. This is not higher than could be expected by chance ($z=1.01$; n.s.). When we include trends, 7 out of 10 tests were significant at $p<0.20$ ($z=2.25$; $p=0.01$). Furthermore, all relations were in the expected direction, which is once again higher than could be expected by chance ($z=2.58$; $p<0.005$). As a consequence, the second hypothesis is supported. We should restrict the answers to the 15 minute interval because later the quality of the answers change.

Mood		Paired t-Test	Psychopathology		Paired t-Test
Cheerful		$t_{(63)} = -1.90^*$	Suspicion		$t_{(71)} = -1.44^*$
Satisfied		$t_{(34)} = -2.51^*$	Derealization		$t_{(64)} = -1.75^*$
Anxiety		$t_{(72)} = -0.32$ n.s.	Auditory hallucinations		$t_{(62)} = -0.02$ n.s.
Motivation		$t_{(72)} = -0.72$ n.s.	Visual hallucinations		$t_{(32)} = -1.55^*$
Lonely		$t_{(71)} = -2.19^*$	Fear to lose control		$t_{(41)} = -1.61^*$
Meta-Analyses	$H_0: p=0.50$ $H_0: p=0.05$	$z = 2.58^{***}$ $z = 1.33^*$	Meta-Analyses	$H_0: p=0.50$ $H_0: p=0.05$	$z = 1.63^*$ $z = -0.51$ n.s.

$^* p<0.20$; $^* p<0.10$; $^* p<0.05$; $^{**} p<0.01$; $^{***} p<0.001$.

Table 6B.1b "Time-window"-effect: aggregation leading to reduced variability (standard deviation).

Discussion

By intuition we accept that reports on site and at the exact moment of the beep are more accurate than reports that are collected more distant from the target event. This is why eye-witnesses and direct observation techniques have more credibility. Our current analysis shows that the hypothesis of drift towards more social desirable answers was not upheld. However, beeps that were answered outside of the predefined time interval — 5 minutes before and 15 minutes after the signal — had less variability. The systematic reduction of the variability in the more remote answers confirms our hypothesis of qualitative differences in the reports produced ad hoc or remotely. Consequently we advise to restrict the beeps that are selected for analyses to those falling within the interval: the "reliability time-window". Beeps reported within this reliability time-window are accepted, those outside are rejected. Using this criterion we rejected between 3.8% (depressive subjects) and 9.1% (chronic mental patients) of the beeps (see Appendix 2).

The relevant time-window varies for different ESM-items. We expect the time-window for thoughts to be small (within 5 minutes) while for activities time-windows up to 30 minutes are allowed. Because we are primarily interested in moods and mental state, we will use the -5/+15 minute time-window throughout the whole study.

We have argued in the introduction to this section that the interval of maximum 5 minutes before and 15 minutes after the beep-moment is the best possible due to restrictions imposed by the watch signaling technology (see Figure 6B.1). New computerized devices (palmtop computers) that substitute a paper and pencil procedure for direct data input, will give us real time records of response times. Answers before the beep, caused by retrospective guesses of the beep time will never occur and more generally we will not have to account for the annoying disturbances caused by non-synchronized beeping devices or subjects faking. Restrictions of beeps to a "reliability window" of 10 or even 5 minutes will be possible if studies show this is necessary.

C: COMPARING ESM WITH DIARY ASSESSMENTS ²⁸

In this study we compare Experience Sampling (ESM) ratings of time allocation, with the same variables collected using more traditional diary approaches. In addition, we compare ESM ratings of mental state with questionnaire mood assessments.

Diaries are the most important unstructured prospective data collection strategy in psychiatry. They are primarily used in behavioral assessments, to collect data on the ecology of the symptom behavior. Diaries are simple instruments that generate a wealth of data cost-effectively. They can be used for the first inspection of problem behaviors. Unfortunately, diaries are difficult to analyze and often miss the information that is needed to plan change strategies effectively. Only the data bits that seem important to the subject are collected. Therefore, diaries are often substituted for structured logs after the initial phase of assessment (Barker, 1985).

The study was conducted as a “classroom” exercise on ecologically valid assessment strategies in a “social psychiatry” course almost 10 years ago by Chantal Dijkman. It has never been published. We will compare time-budgets generated by their traditional assessment tool, the diary, with the time budgets computed using ESM. We will also compare mood assessment generated by a questionnaire — its traditional cross-sectional assessment tool — with aggregated moment-to-moment ESM-based moods.

Subjects

The subjects who participated in this study were 10 medical students. They attended a social psychiatric training course. As a homework assignment on ecologically valid assessment strategies, they were asked to participate in the present study. Because the study was conducted during their internship, they all had rather homogeneous activity schedules over the study period.

Methods

Experience Sampling — In this study we used ESM to assess the mental state and context concurrently in daily life. Subjects were signaled randomly, using a computer-linked pager, 10 times a day between 7:30 and 22:30, to fill in a small questionnaire. The sampling forms contained open-ended questions about current thoughts, activities and social and physical context. These were coded by CD using a preliminary version of the codebook (see Table 6C.2). The forms also contained 7-point Likert-type scales, assessing current mental state.

Diaries — In this study we used two different diary types:

- *UNSTRUCTURED DIARY* — The unstructured diary mirrors traditional diaries. Subjects were instructed to describe the current day each evening. No further instructions

²⁸ This study was conducted by C.I.M. Dijkman-Caes and was published internally under the title “Resultaten Vergelijking ESM en Dagboeken (n=10)” (1987).

were given as to the content of the reports. The diary reports were agenda type forms with the day of the week, as well as the date, at the top, and empty entry lines. The instructions mentioned that the cueing time labels were not very important. The contextual data were coded using the ESM codebook.

Monday		Date: .../.../...
9.00		
		
		
10.00		
		
		
		
...		22.00
...		
		

Figure 6C.1 Report sheet of the unstructured diary (at the end of the day).

- STRUCTURED DIARY** — The structured diary is based on the assessment forms that were used in the international time budget studies (Szalai et al., 1972). The instructions were to fill in the forms continually and to monitor each activity with at least one entry each hour from 12:00 to midnight. Sleeping periods could be described globally. Each activity was to be described — one activity on each line — and recorded with a start and end-time, the place where it occurred, and the persons involved. In addition the activities were also assessed by the subjects as to whether they were freely chosen or planned (both rated on no/yes scales) and as to the motivation (on a 10-point scale from 1= I do not like this; to 10 = I like this very much). Because a lot of activities happen concurrently, there was the possibility of adding a second activity (e.g. visiting family and watching TV). Finally, the last two questions were of less relevance for our study: did you use TV or telephone in this activity (check if yes)?

Monday		Date: .../.../...			(from 12:00 to 24:00)						
Time	What?	Start	End	Where	Who	free	plan.	motive	What2?	TV	Tel.
12:00		12:00									
13:00											
14:00											
.....											

Figure 6C.2 Report sheet of the structured diary (hourly reports).

The contextual data were coded using the ESM codebook. In the analyses we will correct the time budgets generated by the structured diaries for the observations that were collected over sleeping periods (which are unavailable in ESM).

Mood questionnaire — During the debriefing session, the subjects were asked to rate their mood over the previous week on a 31-item questionnaire structured as 5-point Likert scales. The one-time ratings on the Mood Questionnaire (MQ) were compared to

the moment-to-moment mood ratings assessed using the ESM, as well as the diary mood assessments.

In the previous week, did you feel ...

active	• relaxed	• cheerful	• anxious
exhausted	irritated	content	washed out
• self-assured	fully awake	unhappy	gloomy
sleepy	down	clear	• angry
• lonely	full of energy	helpless	weary
vivacious	sad	uncertain	dull
tired	desperate	lighthearted	worn out
listless	almost falling asleep	unworthy	

Table 6C.1 Mood Questionnaire (* refers to the six items of the ESM mood summary score).

Design — Each subject collected data over 3 days for each assessment strategy. The ESM and diary assessments were collected on different days. The first day was always an ESM day, followed by a diary day, repeating this pattern three times. Half of the subjects completed a structured diary assessment during the three diary days, the other subjects the unstructured one. MQs were collected during the debriefing.

Figure 6C.3 Study design.

All 10 subjects finished the ESM assessments. One subject did not comply with the unstructured diary instructions and was lost, resulting in 5 sets of three structured diaries and 4 sets of three unstructured ones. Two subjects did not fill in the MQ.

Due to the items that were included in the different assessment techniques, we were able to compare ESM moment-to-moment mood assessments with the MQ. We computed a moment-to-moment mood summary score on the ESM data and compared the means for subjects with the MQ and a ESM matching subset of items from the MQ (“cheerful”, “self-assured”, “lonely”, “relaxed”, “anxious”, and “angry”). The scales for negative emotions were flipped before computing the summary scores.

Time budgets were primarily compared between the ESM and the structured diaries.

Results

Time budgets

Social context — In comparing the time budgets for persons in the diary methods and ESM, the first yield lower ratings for “alone” and “family” as well as for “strangers”, contrasted to higher assessments for “friends” and “colleagues/ acquaintances”.

Activities — Both the diary approaches and the ESM were equally able to assess activities. Ratings between the two were extremely similar. Marginal differences were

found for “doing nothing” which was never assessed in the structured diary. Also, “care for other” — primarily therapeutic work for these medical students — was rated almost twice as often in diaries compared to ESM. Oddone et al. (1993) conducted a time-budgeting study in a similar population and also found important differences between the ratings by the subjects and actual time sampling data.

Places — The very least that can be said about the time budgets of places assessed using ESM and diary approaches is that they differ. Patterns are difficult to describe. ESM usually takes a mid-position between both diary assessments. The most likely conclusion is that — because the time budgets were from different subjects — the divergences were caused by structural differences between the living situation of the subgroups of subjects that were assigned the two different diary instruments.

		Unstructured diary	Structured diary	ESM
<i>Who?</i>	• alone	14 %	14 %	33 %
	• family	5 %	25 %	24 %
	• friends	35 %	34 %	13 %
	• colleagues/ acquaintances	46 %	28 %	26 %
	• strangers	0 %	0 %	4 %
	Missing %	43 %	31 %	18 %
<i>What?</i>	• nothing	2 %	0 %	5 %
	• self-care	15 %	15 %	7 %
	• care for others	10 %	13 %	5 %
	• work	29 %	24 %	33 %
	• leisure	33 %	42 %	32 %
	• transport	9 %	4 %	8 %
	• other	1 %	2 %	10 %
	Missing %	13 %	3 %	16 %
<i>Where?</i>	• home	26 %	51 %	36 %
	• network	16 %	4 %	10 %
	• work	31 %	23 %	36 %
	• public places	12 %	9 %	6 %
	• transport	11 %	4 %	8 %
	• other	5 %	9 %	4 %
	Missing %	15 %	3 %	16 %
	n	188	226	300

Table 6C.2 Time budget analysis: comparison between ESM and structured (hourly reports) and unstructured diaries (end of day reports).

Assessing missing data — The missing data analysis shows an intriguing pattern. For the ESM missing data were equally distributed over the categories and are caused by signals that were not responded to by the subjects. Only marginal differences exist between the missing data for persons (18%), places (16%) or activities (16%). In the unstructured diaries it is impossible control for missing data, because there are no structured sampling cues to respond to. The lower figures for missing data for “what?” and “where?” in these diaries are therefore irrelevant. Some data bits, especially for the social context, were difficult to code (43% missing). In the structured diary responses were requested hourly. Although cued for explicitly, still 31% of the reports had no reference to the persons involved, much higher than the missing data for activities (3%) and places (3%).

When we need extensive contextual information to assess, for instance, the functional similarities of behaviors, the diaries are a less adequate choice, considering that information on persons is often missed. For activities and places there is a slight preference for the structured diary, but this is due to the aggregation of the informa-

tion over a whole hour which results in a condensed report. Consequently, the representativeness of the categories may not be satisfying.

Mood states

The aggregated mood score over the sampling week assessed retrospectively with the MQ the subset of matching ESM mood scores within the MQ, and the averaged moment-to-moment ESM mood assessments, did not differ very much between subjects. The difference was less than 0.10 on a 7-point Likert scale.

Interestingly, the summary scores of the two item sets of the MQ — the whole 31-item version and the ESM subset of 6 items — were significantly correlated ($r=0.72$). In contrast, the one-time cross-sectional MQ was only poorly correlated with the aggregated mood scores of the moment-to-moment ESM ($r=0.26$). When only matching ESM items are accounted for, the correlation between the end of the week mood scores and the aggregated moment-to-moment ESM ratings was almost significant ($r=0.61$). This is fair considering the small number of subjects.

Concluding, when controlled for specific items, aggregated ESM ratings match the cross-sectional retrospective ratings. For the assessment of global mood states ESM is too expensive. Therefore, in a steady-state situation, ESM is only advised to assess environmental or other kinds of variabilities.

Discussion

What did subjects say about the different assessment strategies? As a general assessment, the methods were considered important intrusions in the subject's daily lives and assaults to their privacy. ESM was considered to be the most intrusive strategy. It was easier not to fill in the forms for some hours or to make up answers for *the missing parts at the end of the day in the diaries* — even in the structured one which was rated equally burdensome as ESM. However, ESM was the most “fun” to do. The beep signals and the structured assessment forms resulted in an appreciation of ESM as a disruptive but rewarding method.

When comparing the time budgets generated by diaries and ESM it is important to realize that only 10 subjects were studied with ESM and only 4 with the unstructured and 5 with the structured diary for 3 different days for each assessment instrument. This is too low to yield stable and unbiased results and allow optimal comparisons. Considering that assessments of the time budgets on activities was the most similar, we can assume that diary methods are a kind of event sampling of activities. Subjects write down what they have been doing the whole day, where they went and, to a much lesser extent, whom they met. The omissions probably reflect what the subject found less important. The differences are relevant between modalities (activities, places and persons) but also within a modality. For instance, the difference in time budgets between diaries and ESM for alone and family compared to friends and colleagues might reflect the fact that it is more important for those subjects to report on friends and work mates above family. The presence of family members is taken for granted and its report is omitted.

We had no diary ratings of moods. As a consequence, no diary / ESM comparisons can be discussed. We did have, however, a retrospective rating of the average mood states of the subjects over the previous week, assessed during the debriefing interview. The subject averages, based on sampled experience in daily life, matched the one-time questionnaire ratings, but not significantly due to the small number of subjects.

In the field of assessment instruments, different strategies yield similar as well as dissimilar results. The interpretation of those patterns can be helpful in understanding the subject's experience and self-narratives. In part we can interpret the differences as validity challenges. However, much more interesting is the picturing of the multifaceted world of daily life experience that can be targeted using the different assessment strategies. The comparison of diaries, a cross-sectional assessment of mood and experience in daily life, shows part of this complexity. Moment-to-moment samples of mood assessments are similar to one-time cross-sectional assessments but the latter miss the possible differentiation between moments (within subjects). Similarly, diaries — especially when structured — are adequate for rating activities, but not for persons involved. However, because ESM targets this assessment, we could never be sure if the subjects would have rated the social context as relevant if they had more freedom to leave it blank. Once again, these comparisons show the importance of targeting assessment strategies to specific needs.

D: ESM TIME AND RIR EVENT SAMPLING IN SOCIAL INTERACTIONS ²⁹**Introduction**

In this section we will explore the validity of the Experience Sampling Method (ESM) — a time sampling strategy — and the Rochester Interaction Record (RIR) — an event sampling strategy — for the study of social interaction.

Event and time sampling: competing assessment strategies

The difference between event and time sampling is based on the rule that is used to trigger the observations. In event sampling, the occurrence of an event triggers the observation, while in time sampling observations are contingent on a time-based rule. Event and time sampling strategies have different strengths and weaknesses. Intuitively we assume that event sampling strategies are the best choice when we assess targeted discrete events that are short-lasting or have low frequencies of occurrence, and when we are interested in concurrent or consequent moods and situations. When choosing an event sampling strategy, we assume that what happened during the possibly long periods of non-targeted time is irrelevant for the topic under study. Of course, it is possible to assess what happened before the event occurred using event sampling techniques, but this will be done retrospectively with the observer's knowledge of the imminent emergence of the target event. This occurs at the cost of validity.

Time sampling is the ideal choice when the target of the observation is a non-discrete event, a more diffuse set of events, or when we are interested not only in concurrent and consequent moods and situations, but also in their antecedents. Most emotions and psychopathologic symptoms are not discrete events. They may seem discrete, as a consequence of the assessment strategy, especially in event sampling. Moreover, in the assessment of emotions, it is crucial to broaden our scope of observational strategies, from the extreme situations in which subjects are totally aware of the contingencies ruling their emotions into the more diffuse situations where contextual determinants are not consciously activated when formulating an assessment.

In this study we will explore the relative merits of time and event sampling techniques for the assessment of social interaction. Self-observations of social interactions — clearly a discrete event — were assessed concurrently using both a time and event sampling strategy. The study of social interactions is complex. We will restrict ourselves to the assessment of a small number of social interaction parameters, i.e.:

- the “time budget” of social interaction;
- the differential emotional states while in interaction;
- the differential time use patterns while in interaction;
- the assessment of subject characteristics of social interaction.

²⁹ This article was written collaboratively by Ph.A.E.G. Delespaul; H.T. Reis and M.W. deVries, entitled “Assessing social interaction using time and event sampling techniques”, with special thanks to Sandra Gransier for participation in the data-collection.

Subjects

The subjects for this study were 42 students (23 females and 19 males) at the Rijksuniversiteit Limburg, a medium sized public university in the city of Maastricht in the southern part of The Netherlands. The women were 1.4 years younger than the men (mean age 20.26 (1.21) and 21.63 (1.61), significant at the $p < 0.001$ level). Demographically, the subjects form a homogeneous group. With the exception of two, all students were single and lived in student lodgings ($n=37$), alone ($n=1$) or with family ($n=4$). They all held scholarships.

Method

Subjects were recruited for a study of daily experiences. They were solicited through notices in the university newspaper and announcements in classes. They all received 25 guilders (approximately \$15) for their participation. They were told that they would be required to complete the Experience Sampling Method (ESM) protocol as well as the Rochester Interaction Record (RIR) for 1 week.

Experience Sampling Method (ESM)

The Experience Sampling Method is a random time sampling technique to collect repeated self-assessments in the natural flow of experiences, environments and activities of the subject's life (deVries, 1987, 1992). ESM is a signal-contingent method for assessing daily experiences (Wheeler & Reis, 1992). Subjects wear a portable, pre-programmed digital watch (SEIKO RC-1000) for 1 week (7 days). Each time the watch delivers a signal, they are required to complete a structured diary record — the ESM-form (ESF) — describing their feelings, thoughts and activities at that moment (see below). In this study, a fixed block of 10 pre-programmed signals per day was used (Delespaul, 1992). The signals were programmed to beep between 7:30 and 22:30 resulting ten 90-minute intervals. Within each interval, the timing of signals was random, subject to the constraint that no two signals could occur within 15 minutes of each other. A different random schedule was used for each day.

The ESM protocol makes some concessions to the subjects' personal needs and convenience. Subjects are permitted to turn the beeper off when sleeping, and when they do not wish to be disturbed. This means that most subjects' data contained fewer than the potential maximum of 70 data records. In the present study we averaged 46.5 valid responses per week, which is consistent with response rates found in prior ESM studies (Appendix 1).

ESM procedure — A short training session (briefing) was held 1 day before the subjects were to begin ESM recording. During that meeting, technical details about the watches were explained fully. In addition the importance of completing the record immediately after each beep was stressed. Delays in answering can be assessed because we ask the subjects to report the current time when they have completed the record. This time is compared with the programmed beeper schedule. Reports with

time discrepancies that are greater than 5 minutes before or 15 minutes after were considered invalid. 4% of the beeps were deleted. In the briefing session, the subjects responded to a “try-out” beep, during which a brief orientation to the meaning and content of the terms and concepts on the record form was provided. After the last day of recording, subjects were interviewed fully, to be sure that they complied with the intended procedures (debriefing session). We require at least 33% of the beeps to be valid to use the subject’s data. Therefore, two subjects (4.8%) were discarded.

ESM-form (ESF) — The ESM measures were modified from several existing modules, designed to examine different constructs. All items were worded in Dutch. For convenience we present the English translation. The first question probes the subject to describe his thoughts and to rate the pleasantness, clarity, as well as any difficulties concentrating on a 1 (not at all) to 7 (very much) scale. Subjects also rated a set of items probing the subject’s current mood on similar scales. The phrase “I feel...” was followed by several terms: cheerful, uncertain, lonely, relaxed, anxious, satisfied/content, irritated, sad, guilty, safe/secured, valued and rejected. In addition, scales with a more physical connotation were rated: hunger, tiredness and not “lekker”, a term with no exact subjective equivalent in English. “Lekker” primarily means “tasting good” but as an emotion refers to feeling good, with physical connotations.

Subjects were also asked to describe the place they were in, if they were alone at the moment of the beep and, if not, who they were with. In the case of social interaction — not being alone — an additional set of items concerned the nature of the social interaction and also had to be rated. Nine scales were included: intimacy, self-disclosure, other-disclosure, disagreement/ conflict, giving help or support to others, receiving help or support from others, enjoying pleasant company, doing something together, and whether they would rather be alone. The same 1-7 scale was used.

The final set of items concerned their current activities. They were phrased in short sentences: “I would rather be doing something else”; “I am being active”; “It is difficult”; “The activity is challenging”; “I do it well” (I’m managing it well); “I received help”; “I wanted help/support”; “I received the help I needed”. Again, assessments were made using 1-7 Likert-type rating scales.

Rochester Interaction Record (RIR)

The RIR is an event-contingent, fixed format diary procedure, in which subjects keep detailed records of their social activity within a period of usually 2 weeks. Immediately after, or soon after, every social interaction lasting 10 minutes or longer, subjects complete a diary record describing that interaction along several standard dimensions, including variables that are of particular interest for that study. From these records, objective summary indices are computed, which together paint a vivid portrait of the nature and extent of each individual’s social activity. In one representative study the following variables are assessed: number of interactions, average length of interactions, number of different partners, distribution of interaction across different partners and mean level of experience during interaction. Each index is computed at several levels of composition: across all interactions; within same-sex; opposite-sex; mixed-sex; and larger group interactions; and with specific partners, such as best-friends;

spouses and confidants. We refer to Reis & Wheeler (1992) for a more extensive description of the RIR protocol.

RIR procedure — During the same briefing session in which the ESM was introduced to the subjects, the RIR was also presented. Both methods were used for the same period of 7 days. Periods were targeted so as not to include special events such as major holidays or examinations.

RIR form — The RIR forms are organized by day in small sets that can be carried easily. Each page allows space for description of one social interaction. Those are defined by date, starting time, duration and persons involved. The most important three persons are listed by name and sex. In larger groups the number of men, women and children are counted. The nature of the interaction is rated on 7-point scales from “very little” to “a great deal”. The scales include ratings of: intimacy; self-disclosure; disclosure by others; quality (pleasant); satisfaction (compared to expectation); subject helping others; others helping subject; degree of conflict; and degree of closeness. The next question is an open-ended one about the content of the interaction. This was coded by the researchers afterwards. Finally a number of Likert-type, 1 - 7 scales about emotional experience during interaction were added. We asked: “Did you feel...: happy; sad; frustrated; rejected; valued; warm; hurt; anxious; stimulated; tense; productive; and accepted”.

Overlap between the time and event sampling procedures

Both the ESM and RIR can assess objective and subjective parameters of social interaction. Objective parameters are derived from questions such as “when did the interaction occur?”, “for how long?”, “where?”, “with whom?”, “to what purpose” and “which activities were involved?”. Subjective parameters are related to the concurrent emotional states. We will describe the relative strengths and weaknesses of both instruments for those assessments.

Additional scales (one-time measures)

During the debriefing sessions, subjects were asked to complete some cross-sectional questionnaires. They are (listed along with α reliability coefficients): loneliness (UCLA Loneliness scale ($\alpha=0.89$); Russell et al., 1980); psychologic well-being (Differential Personality Questionnaire, ($\alpha=0.89$); Tellegen, 1982); satisfaction with life (Diener's Satisfaction with Life Scale ($\alpha=0.74$); Diener et al., 1985) and social support (Cohen-Hoberman Interpersonal Support Evaluation List (Cohen & Hoberman, 1983): tangible support, $\alpha=0.49$; appraisal support, $\alpha=0.67$; esteem support, $\alpha=0.71$; and belonging support, $\alpha=0.80$). A closeness support scale, developed by Reis (1989) to be similar in format to the above, was also used ($\alpha=0.66$). We also included a measure of positive and negative mood. For each mood, the six most representative adjectives from Diener and Emmons' review (Diener & Emmons, 1984) were chosen (positive mood: happy, joyful, pleased, enjoying yourself, glad, and delighted ($\alpha=0.86$); negative mood: angry, anxious or worried, frustrated, sad, depressed or

blue, and unhappy ($\alpha=0.85$)). These measures were always completed after the RIR and ESM. In most cases, the questionnaires were completed in the interviewer’s office, but in a few instances they were completed at the subject’s home.

The RIR and cross-sectional questionnaires have been standardized on American students. Since in most situations no Dutch translations were available, we used the classic cross-cultural technique of translation and back-translation. We, however, found this method not totally adequate. The first problem was the difference in the daily life routines of students in both countries, making some items irrelevant. The second problem was caused by subtle differences in the American and Dutch emotional vocabulary, which were not picked up using the up and back translation technique. We therefore finalized the questionnaire using students of the University of Maastricht and expert panels of persons who were native speakers in one of the languages and who had lived in both cultures for lengthy periods of time.

Results

Response frequency and overlap of time and event sampling

Each of the 42 subjects could theoretically respond to 70 beeps (7 days and 10 beeps by day) for the ESM, a total of 2940. In response to the beeps, 33.1% of the ESFs were left blank, primarily because subjects were still sleeping in the morning (76% and 58% missing for the first two beeps of the day). Another 79 ESFs (4.0%) had to be dropped because the response-time fell out of the ESM reliability window (5 minutes before and 15 minutes after the beep time). Although the frequency of those “invalid” responses was low, the frequency of late responses was twice as high during social interaction ($\chi^2_{(2)}=64.83$, $p<0.0001$). Finally, the data of two subjects were omitted totally because they failed to respond reliably to more than 1/3 of the emitted beeps. The resulting number of ESFs is 1867 (63.5%).

From a total of 1554 RIRs, collected from 42 subjects, we had to delete 71 (4.6%; $\bar{x}$ by subject 1.69; $s.d.=3.06$) because of missing start- and/or duration times which made matching with the ESFs impossible. The 42 RIRs of the two subjects who did not meet the 1/3 reliability criterion in ESM were also discarded. This resulted in 1441 RIRs for the analyses comparing ESM and RIR. However, because the RIR-method makes assessments when the interaction has ended it can also be used for the assessment of interactions during sleeping hours. In contrast, ESM requires assessments at the moment itself for which the subject has to be awake. Due to technical limitations of the sampling device we sample between 7:30 and 22:30. A total of 102 (7.1%) of the 1441 remaining RIRs fell out of this ESM sampling window (night assessments). The resulting 1339 RIRs can be used in the comparisons with the ESM data.

Figure 6D.1 Schematic presentation of ESM/RIR matching analysis.

Figure 6D.1 presents the analytic procedure to define matching ESM and RIR observations. Different situations can be found: alone situations in ESM with no RIR report; alone situations in ESM with RIR report; not alone situations in ESM with no matching RIR; RIRs with no matching ESM; and RIRs with matching ESM. There were 680 ESM reports that matched the available RIR observations, 36.4% of the whole set of ESM observations. The number of RIR reports that overlap ESM reports is 527 (39.4%). The largest proportion of RIRs have only one matching ESM report (74.6%). The frequencies quickly diminish for RIR reports with 2 (22.4%), 3 (2.5%) and 4 (0.6%) matching ESM reports. We found no RIRs with more than 4 matching ESM reports (Figure 6D.2).

Figure 6D.2 Frequency distribution of the number of matching ESM reports/RIR report.

The frequencies of ESM and RIR observations by subject are summarized in Figure 6D.3. Because sometimes more than one ESM report matches a specific RIR report, the frequency distributions of matching ESMs and RIRs are slightly different. No additional subjects were lost. All subjects have at least one match which allowed comparisons. In a specific analysis we will restrict ourselves to those interactions with more than one matching ESM record ($n=134$): 15% of the subjects have one such RIR, 15% have 2; 32.5%, 3; 15%, 4; 15%, 5; 5%, 6 and 2.5% had 11.

Figure 6D.3 Frequency distribution of the number of matching ESM and RIR reports/subject.

The “time budgets” of social interaction

Assessing social interaction — We can compute the proportion of time that a subject spends in interaction using the ESM and the RIRs. The computational strategy is slightly different in the two situations. In ESM we compute the proportion of time spent in not-alone situations within the ESM sampling window (between 7:30 and 22:30) and this figure is a reflection of the time spent in interaction (e.g. 50% of the time). Using the RIRs we can compute the total amount of time spent in interaction. To make the data comparable between ESM and RIR, we restrict the RIR observation to the ESM sampling window. The proportion of time spent in interaction is computed by summarizing the RIR time in interaction and dividing this by the total assessment time (in minutes: 7 days multiplied by 15 hours of 60 minutes). The results of the analysis are shown in Figure 6D.4.

Figure 6D.4 Comparing RIR and ESM assessments of social interaction (with 5% error interval).

Compared to ESM ratings, the RIRs provided lower estimates of social interaction time, a difference of 20% (35% compared to 55%; $z=1.78$; $p<0.075$). The variability (s.d.) between subjects of the individual estimates of alone time was lower for RIR than for ESM (9.63% and 15.44%; pairwise t-test for “alone time” $t_{(39)}=2.46$; $p<0.02$). Part of this difference can be accounted for by the RIR instructions, stipulating that only interactions lasting for more than 10 minutes should be reported. In this study the difference is 20%, or 3 hours of the day between 7:30 and 22:30. Because each unreported interaction lasts a maximum of 10 minutes, each day at least 18 interactions are not reported in the RIR assessments — 126 interactions in a whole week. The mean number of interactions in waking as well as night hours reported by each subject was 33.48 (s.d.=11.98). Thus, only slightly more than one fifth of the full set of ESM interactions were reported by the RIR. Part of the difference also can be accounted for by missing or rejected data in ESM. The difference between ESM and RIR ratings of alone time were especially high during the first two beeps of the day. Students often sleep late in the morning and only filled in ESM ratings when they were awake. Because more social moments occur when awake and more subjects were alone when asleep, the differential missing data pattern results in lower estimates of

overall alone time. By correcting the data using conditional probabilities we estimated the underestimation of alone time in ESM to be 3.54%. A correction in the opposite direction can be considered in view of the higher number of invalid, social ESM responses when the elapsing time between event and report is large. This effect on overall alone time ratings is assumed to be marginal: about 0.5%. The corrected figure for ESM alone time is 48% (45% + 3.5% minus 0.5%). The proportion of time in interaction, computed by ESM and RIR, is significantly related for the 40 subjects ($r=0.445$; $p<0.01$). The resulting error of estimation³⁰ for the assessments of alone time is 8.62%. However, because the RIR and ESM assessments are validated against each other and no “gold standard” is available, the real error of estimation might be lower.

Match/mismatch analysis — In Figure 6D.4 we saw that ESM and RIR gave significantly different assessments of social interaction. Here we evaluate whether subjects give matching answers — concordant reports of alone and not-alone at the same moment both in ESM and RIR — on the same, specific moments for both assessment strategies. Overall, the subjects described themselves in interaction using RIR but not according to ESM on 14% of the moments. The range between subjects was from 0% — a perfect match — to 50%. When we select those RIRs containing mismatches (ESM alone reports) with at least two ESM reports (40% of the subsample), 62% of the mismatched ESM reports were in the last matching ESM (see Figure 6D.1).

ESM (rows) x RIR (columns)		Alone	Not-alone ³¹
Alone	$\bar{X}$ (s.d.)	—	14.3% (11.51%)
	min.-max.		0.0- 50.0
	test		$t_{(39)}=7.86^{***}$
Not-alone	$\bar{X}$ (s.d.)	37.3% (20.10%)	
	min.-max.	3.2 - 84.6	—
	test	$t_{(39)}=8.59^{***}$	

*** $p<0.001$

Table 6D.1 Conditional probability of matches and mismatches of social interaction in RIR and ESM.

In 38% of the cases, the error occurred in a previous report. The higher frequency of mismatched interactions in the most recent observation is intriguing. Because RIRs are reported after the interaction, we expected the highest accuracy in the most proximate reports. An explanation for this unexpected result is difficult to find. The RIR instructions allow not to report alone-periods of less than 10 minutes — embedded within a longer lasting interaction. There is, however, no reason to assume that those short interruptions would occur more frequently in the last part of the interaction.

The RIR is an event sampling procedure. Therefore, information is available for the target event, but not for the rest of the time. To reduce the burden for the subjects, only interactions lasting more than 10 minutes are reported. As a consequence, a subject might still be with others in a period not covered by a RIR report. In 37% of the observations of being “alone” according to the RIR assessments subjects were not alone according to ESM. The range for individual subjects was between 3 and 85%. As

³⁰ $\sigma_{\text{est}} = \sigma_y \sqrt{1 - r^2_{xy}}$.

³¹ Due to the instruction for the RIR — assess only when interactions last for more than 10 minutes — not-alone RIR assessment are not always not-alone moments. Likewise, short interruptions of the interaction can lead to an ESM-alone ratings while the RIR interaction still lasts.

might be expected, these figures are higher than the “in interaction” RIR report with concurrent alone ESM report (14.3%).

Assessing experience during social interaction

The study of social interaction is a study of interaction frequency, interaction characteristics and, finally, personality characteristics that show themselves in social interaction. In this section we will discuss the use of ESMs and RIRs for the study of emotional interaction characteristics from a between and within subject point of view.

Subject characteristics — Table 6D.2 presents a review of subjective means of interaction characteristics and emotions using ESM and RIR ratings in matched social interactions. The assessments made by ESM and RIR are highly correlated *BETWEEN SUBJECTS*. The correlation is higher for the interaction related self-assessments (intimacy, self-disclosure, etc.) and lower for general emotional states (happy, sad, anxious). Because all the between-subject correlation coefficients are highly significant, we can conclude that both ESM and RIR agree in the assessment of moods during social interaction.

	RIR $\bar{X}$ (s.d.)	ESM $\bar{X}$ (s.d.)	Pearson r	Paired t-test
Intimacy	4.47 (0.78)	4.68 (0.90)	0.72 ***	2.16 *
Self-disclosure	4.41 (0.94)	4.34 (1.15)	0.87 ***	-0.73 n.s.
Other-disclosure	4.44 (0.98)	4.31 (1.19)	0.92 ***	-1.77 *
Conflict	1.51 (0.45)	1.34 (0.32)	0.57 ***	2.86 **
I helped others	3.71 (1.00)	3.36 (1.09)	0.87 ***	4.02 ***
Others helped me	3.75 (1.01)	3.33 (1.03)	0.88 ***	5.08 ***
Happy	4.79 (0.55)	5.06 (0.53)	0.51 **	-3.11 **
Sad (depressed)	1.53 (0.44)	1.55 (0.42)	0.56 ***	-0.31 n.s.
Anxious	1.34 (0.42)	1.18 (0.23)	0.59 ***	3.10 **
Irritated	1.56 (0.47)	1.54 (0.37)	0.34 *	0.21 n.s.
Rejected	1.30 (0.44)	1.19 (0.23)	0.54 ***	1.81 *
Valued	5.00 (0.62)	5.14 (0.63)	0.75 ***	-2.07 *
Satisfied	5.06 (0.58)	5.02 (0.54)	0.53 ***	0.39 n.s.

* p<0.10; * p<0.05; ** p<0.01; *** p<0.001.

Table 6D.2 Validity of experience during social interactions: sensitivity *between* subjects.

In contrast, the assessments of social interaction with both techniques produced differences in overall subject characteristics. In an analysis of matched interactions significant differences were found for intimacy, conflict, help (self and other), happiness, anxiety and feeling valued.

Moment characteristics — We have compared RIR and ESM-based subject characteristics of social interaction. We will now explore the sensitivity of both assessment strategies to assess the variability of emotions and emotions *WITHIN SUBJECTS* for different interactions. For each matching RIR and ESM interaction report we have one RIR response and one or more ESM emotional ratings. Rank correlations (Spearman rho’s³²) by subject between RIR and averaged ESMs by interaction reflect whether the emotional characteristics of specific interactions were rated equally by both instru-

³² Computations using the Kendal tau give comparable results.

ments. The first column of Table 6D.3 contains the means³³ and standard deviations of the different correlations by subjects. Significance is computed using two strategies. First, we computed a t-test on the rho mean with a null-hypothesis of no correlation (rho=0.00). All t-tests — except “anxiety” and “feeling satisfied” — were significant. The second strategy uses a meta-analytical technique (Rosenthal, 1978, 1979). We tested the significance of the proportion of significant individual rho’s (p<5%). This proportion of significant tests is evaluated against the null-hypothesis of 5% random spurious significance. Although more conservative, this strategy confirmed the Spearman rho based findings except for “feeling rejected” that was no longer significant.

	Spearman rho	t ₍₃₈₎ on rho	Meta-analysis	
	$\bar{X}$ (s.d.)	H ₀ : rho>0.00	P _{subjects} (p<0.05) ³⁴	t ₍₃₈₎ on P _{subjects}
Intimacy	0.57 (0.32)	6.30 ***	67.50%	5.81 ***
Self-disclosure	0.49 (0.34)	5.00 ***	47.50%	4.32 ***
Other-disclosure	0.51 (0.35)	5.42 ***	60.00%	5.25 ***
Conflict	0.41 (0.37)	3.47 ***	35.00%	3.35 **
I helped others	0.48 (0.33)	4.01 ***	52.50%	4.69 ***
Others helped me	0.45 (0.31)	4.62 ***	40.00%	3.75 ***
Happy	0.29 (0.31)	2.05 *	22.50%	2.27 *
Sad (depressed)	0.30 (0.32)	2.19 *	22.50%	2.27 *
Anxious	0.18 (0.32)	0.59 n.s.	12.50%	1.19 n.s.
Irritated	0.32 (0.37)	4.01 ***	22.50%	2.27 *
Rejected	0.16 (0.43)	3.06 **	12.50%	1.19 n.s.
Valued	0.33 (0.34)	2.23 *	37.50%	3.55 ***
Satisfied	0.25 (0.27)	1.73 n.s.	12.50%	1.19 n.s.

* p<0.10; * p<0.05; ** p<0.01; *** p<0.001

Table 6D.3 Validity of experience during social interactions: sensitivity *within* subject by interaction.

Compared to the between-subject analysis, the within-subject analysis (interaction characteristics) assesses ESM and RIR concurrent validity more stringently. While the results of the previous analysis could be caused by subject response styles — both the ESM and RIR were rated by the same subjects using the same 7-point Likert scales — this is ruled out in the present within subject analysis. Here, correlations are computed within-subject over matching interactions for each subject (14 on average). The ESM/RIR pattern of correlations found between subjects was replicated within subjects. Higher correlations were found for the interaction assessments, while lower correlations were found for moods. Although still significant, the within-subject correlations are lower than the between-subject ones. Those differences, however, are not a threat to validity, because we do not expect that the aggregated RIR based emotional assessments totally match the ESM ratings at specific moments within the interaction.

Discussion

The observed differences between the patterns of correlations allow us to engage in a preliminary discussion on which mood is characteristic of the subject, the interaction or the moment within the interaction. We assume that items that have a high correla-

³³ Because of the problems relating to averaging correlations we have rerun everything with Fisher z-transformed correlations which resulted in the same trends. Contrary to the generally accepted belief, using the average of Fisher z-transformed correlation coefficients is not a valid procedure (Koele, 1984; Silver & Dunlap, 1987) not even for product moment correlations.

³⁴ This test uses a meta-analytic technique. We counted the number of significant within-subject correlations (rho’s) in the sample and checked the resulting proportion against chance (5%).

tion between subjects for ESM and RIR assess a subject trait. Their assessment — even in daily life — disregard contextual cues. In contrast, items that have high within subject correlations for ESM- and RIR-assessments of the same interaction, reflect characteristics of those interactions — possibly the persons involved in it. Finally, items that have a low between subject and within subject correlation assess features of that change from moment to moment within the interaction. Of course, we should always consider the fact that in the latter situation the items might not reliably be assessed. Because the ratings for social interaction and mood assessments behave differently we will discuss both clusters separately:

For the assessments of *social interactions*, the highest between-subject correlations were found for “disclosure” and “help” (both self and others), the lowest for “intimacy” and “conflict” (Table 6D.2). Assessing the moment, the highest within-subject correlations were found for “intimacy”, the lowest for “conflict” (Table 6D.3). The largest differences in pattern were found for “intimacy” (more characteristic of the interaction) and for “others helped me” (more a subject characteristic). Finally, “conflict” had the lowest correlation in both assessments. This may indicate that this mental state is not characterized by subject or interaction-related features and can vary moment to moment within an interaction.

In the assessment of *moods*, feeling “valued”, “anxious” and “sad” have the highest between-subject correlations, while feeling “rejected”, “satisfied”, “happy” and “irritated” have the lowest. At the within-subject level, feeling “valued” or “sad” characterized the interaction best, while “anxiety” and “rejection” were the worst. The latter mood states also have the highest differences in ranking. “Anxiety” and “rejection” seem more characteristic of the subject than of the interaction, but for “rejection” the correlation is low in both situations. “Irritation” ranks much higher as a within-subject characteristic than as a between-subject one. It is therefore more reflective of the interaction. Overall, feeling “happy”, “satisfied” and “rejected” have low correlations, both between- and within-subject, and might be the most moment-sensitive items.

Social interaction	between	within	rank Δ	Mood	between	within	rank Δ
Intimacy	5	1	4	Happy	6	4	2
Self-disclosure	3	3	0	Sad	3	3	0
Other disclosure	1	2	-1	Anxious	2	6	-4
Conflict	6	6	0	Irritated	7	2	5
I helped others	4	4	0	Rejected	4	7	-3
Others helped me	2	5	-3	Valued	1	1	0
				Satisfied	5	5	0

Table 6D.4 Experience during social interactions: comparing ranks of correlations to assess mood states as characteristics of the subject (between), the interaction, (within) or the moment (none of both).

Which source of information influences the RIR aggregated mood states the most? Theoretically, we have operationalized a number of different alternatives:

- RIR assessments of a social interaction are a reflection of the mean mood — as rated using ESM — over the whole interaction ($ESM_{\bar{x}}$: optimal aggregation);
- RIR assessments are a reflection of the last (ESM_{last} : recency effect) or the first mood of the social interaction (ESM_{first} : primacy effect).
- RIR assessments are a reflection of the most extreme mood states within the social interaction ($ESM_{max/min}$: maximum or minimum ESM-rated mental state).

We used the Spearman rho rank correlation coefficient to assess the relative strength of the relation³⁵. Analyses are based on the meta-analytic technique that assesses the proportion of positive ($\alpha=0.50$) and significant ($\alpha=0.05$) correlations (Table 6D.5).

Irrespective of the actual assessment procedure, the most significant relations of ESM and RIRs were found for the first and maximum ESM rating, with the mean over the ESM ratings being a close third. The minimum and last ESM ratings, although still significantly related to the RIR assessment, are less important for the retrospective evaluation of the social interaction. There is some contamination in this analysis, however. To maximize the discrimination we should exclude those pairs of observations from the analysis which are based on the same observations. This is difficult to avoid due to the small number of ESM observations for each RIR. A prospective study, in which the occurrence of an event triggers a more intensive time sampling schedule, would be more appropriate for this kind of assessment. Conditional probability schemes could then be used to assess the relative importance of the different data bits.

rho (RIR x ESM ...)	α	ESM $\bar{x}$	ESM last	ESM first	ESM max	ESM min
Intimacy	.50	88 (2.90) **	84 (2.56) **	84 (2.56) **	84 (2.56) **	84 (2.56) **
(n=25/19)	.05	42 (2.70) **	47 (2.97) **	32 (2.12) *	42 (2.70) **	42 (2.70) **
Self-disclosure	.50	78 (2.12) *	74 (1.82) *	78 (2.12) *	78 (2.12) *	74 (1.82) *
(n=27/21)	.05	33 (2.33) **	29 (2.04) *	43 (2.88) **	29 (2.04) *	33 (2.33) *
Other-disclosure	.50	74 (1.82) *	70 (1.53) *	74 (1.82) *	74 (1.82) *	70 (1.53) n.s.
(n=27/20)	.05	40 (2.65) **	30 (2.08) *	45 (2.92) **	35 (2.37) **	35 (2.37) **
I helped others	.50	86 (1.37) *	90 (2.87) **	86 (1.37) *	90 (2.87) **	86 (1.37) *
(n=21/16)	.05	7 (0.15) n.s.	13 (0.75) n.s.	13 (0.75) n.s.	25 (1.58) *	7 (0.15) n.s.
Others helped me	.50	72 (1.53) *	72 (1.53) *	83 (2.12) *	83 (2.12) *	78 (1.73) *
(n=18/15)	.05	13 (0.79) n.s.	13 (0.79) n.s.	13 (0.79) n.s.	20 (1.24) n.s.	20 (1.24) n.s.
TOTAL	.50	80 (4.77) ***	78 (4.47) ***	81 (4.92) ***	81 (5.07) ***	78 (4.47) ***
(n=118/91)	.05	29 (4.67) ***	27 (4.40) ***	31 (4.95) ***	31 (4.95) ***	29 (4.67) ***
rho (RIR x ESM ...)	α	ESM $\bar{x}$	ESM last	ESM first	ESM max	ESM min
Happy	.50	75 (1.79) *	71 (1.48) *	79 (2.11) *	75 (1.79) *	71 (1.48) *
(n=24/22)	.05	18 (1.37) *	14 (0.99) n.s.	18 (1.37) *	9 (0.52) n.s.	14 (0.99) n.s.
Valued	.50	76 (1.79) *	76 (1.79) *	71 (1.42) *	62 (0.78) n.s.	76 (1.79) *
(n=21/19)	.05	16 (1.09) n.s.	5 (0.04) n.s.	11 (0.64) n.s.	11 (0.64) n.s.	11 (0.64) n.s.
Satisfied	.50	55 (0.39) n.s.	55 (0.39) n.s.	66 (1.20) n.s.	52 (0.13) n.s.	62 (0.93) n.s.
(n=29/24)	.05	21 (1.64) *	25 (1.94) *	25 (1.94) *	33 (2.49) **	21 (1.64) *
TOTAL	.50	68 (2.17) *	66 (2.00) *	72 (2.69) **	62 (1.49) *	69 (2.34) **
(n=74/65)	.05	18 (3.50) ***	15 (2.97) **	18 (3.50) ***	18 (3.50) ***	15 (2.97) **

* $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ (italic figures are best fit options.)

Table 6D.5 Meta-analysis of *positive* ($\alpha=0.50$) and *significant* correlations ($\alpha=0.05$) between RIR and ESM indices of concurrently measured interactions within subjects: proportion of correlations that fit the criterion; z-value between parentheses; and significance level.

Assessing activities during social interaction

We now turn to the assessment of activities during social interactions. These assessments can be done both with the ESM and RIR techniques. However, the RIR was not developed for that purpose and assessments of activities will be less fine-grained than

³⁵ Differential summary scores of ESFs within a specific interaction are only possible when at least two ESFs were available. Rank correlations are only meaningful when at least two interactions — preferably much more — are available. Six out of 40 subjects were excluded. Additional subjects were lost when items have no variance for them. Finally, to avoid capitalizing on chance, we only withheld items for which data of more than 10 subjects was available. Conflict (6); sadness (10); anxiety (2); irritation (7); and rejection (3) were excluded.

in the ESM time sampling. In the RIRs the subjects rate the activity that most appropriately characterizes the interaction in retrospect. In the ESM, the subjects only report their current activities and in repeated ESFs within one RIR period different descriptions of activities are possible.

Activity time budgets — In Table 6D.6 we present the activity time budgets as they were computed using ESM and RIR assessments at matching (social) moments (680 ESFs and 527 RIRs). The number of codes used for the ESM and RIR assessments were almost the same. The code “work related transport” was never rated in ESM, while the codes “personal hygiene” and “personal mental” were never rated in the RIRs. The latter observation is understandable because those are low frequency private — non-social — activities. On the other hand, it is not true that the event sampling procedure missed low frequency activities. Only some low frequency activities were missed. The time budgets computed by ESM and RIR were significantly different ($\chi^2_{(33)}=281.10$; $p<0.0001$). A closer look at the significant differences shows a number of trends. Unfocused activities are underrated in the RIRs. We found significantly higher ESM frequencies of “doing nothing”, “taking breaks at work”, and “personal mental activities”. Differences in the same direction are found for leisure activities such as “watching TV” or “reading”. In contrast, higher frequencies of social interaction reports were found in the RIRs, for instance, “conversations”, “meals”, and “cafe and pub visits”. This is true, even though the ESFs used in this comparison are restricted to observations during interactions and they match the RIRs.

What	ESM	RIR	z	p(z)	What (continued)	ESM	RIR	z	p(z)
Nothing	1.70	0.20	2.44	**	Hobbies, art...	4.02	3.28	0.65	n.s.
Sleep, nap, rest	0.15	0.20	-0.20	n.s.	Movie, theater...	1.08	0.61	0.83	n.s.
Work	1.85	2.25	-0.47	n.s.	TV, radio...	9.58	7.17	1.44	•
Volunteer work	2.32	3.07	-0.79	n.s.	Read book, ...	1.24	0.20	1.94	*
School, study	9.89	10.66	-0.42	n.s.	Related transport	1.24	0.41	1.48	•
Study, irregular	8.50	6.56	1.22	n.s.	Other leisure	1.55	0.61	1.46	•
Work & study breaks	1.08	0.20	1.75	*	Conversation	17.31	19.67	-1.02	n.s.
Related transport	—	0.41	-1.63	•	Physical interaction	0.15	0.20	-0.20	n.s.
Other work, study	1.39	0.82	0.90	n.s.	Visits	1.24	4.92	-3.71	***
Household chores	8.66	4.71	2.58	**	Café, pub visits	0.62	2.46	-2.60	**
Shopping & services	1.85	1.84	0.01	n.s.	Parties	0.62	2.25	-2.39	**
Personal hygiene	2.16	—	3.27	***	Related transport	1.39	1.43	-0.06	n.s.
Related transport	0.31	0.20	0.34	n.s.	Other social	0.93	1.02	-0.17	n.s.
Other maintenance	1.08	0.20	1.75	*	Regular meals	8.66	13.32	-2.52	**
Personal mental	0.62	—	1.74	*	Special meals	0.62	1.43	-1.39	•
Active sports	1.08	1.02	0.09	n.s.	Snacks	3.55	3.69	-0.12	n.s.
Taking a walk	1.70	1.43	0.36	n.s.	Other transport	1.85	3.48	-1.72	*

• $p<0.10$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$.

Table 6D.6 Time budgets of activities in matching RIRs and ESM reports.

Match/mismatch analysis — The proportion of matching activities between ESM and RIR is 43.9%, much better than chance considering the 34 alternative codes. Higher matches were found for those activities that characterize social interactions. For instance, structured work and study appointments resulted in high match-frequencies

in both directions, given a RIR or ESM work report. On the other hand, unstructured work-related activities such as work and study-breaks in ESM are rather seldomly linked to RIR work breaks — but more to the general work category. On the other hand, the RIR interaction characterized as a work break was also detected by the ESM. The same trend is found for household chores and focused leisure activities.

% What	RIR/ ESM	% in group	ESM/ RIR	% in group	% What (Continued)	RIR/ ESM	% in group	ESM/ RIR	% in group
Nothing	0.0	0.0	0.0	0.0	Hobbies, art,...	56.0	4.0	70.0	5.0
Sleep, nap, rest	—.-	—.-	—.-	—.-	Movie, theater,...	16.7	16.7	33.3	0.0
Work	75.0	0.0	75.0	8.3	TV, radio,...	48.3	0.0	63.6	2.3
Volunteer work	85.7	0.0	66.7	0.0	Read book,...	14.3	0.0	100.0	—.-
School, study	82.5	11.1	76.5	8.8	Related transport	12.5	50.0	33.3	0.0
Study, irregular	56.0	6.0	59.6	17.0	Other leisure	0.0	40.0	0.0	60.0
Work/study break	16.7	50.0	100.0	—.-	Conversation	40.2	37.5	37.4	5.2
Related transport	22.2	33.3	100.0	—.-	Physical interact.	0.0	0.0	0.0	0.0
Other work, study	28.6	0.0	50.0	25.0	Visits	12.5	37.5	3.7	29.6
Household chores	38.9	0.0	55.2	0.0	Café, pub visits	75.0	0.0	16.7	33.3
Shopping/ services	58.3	0.0	58.3	8.3	Parties	100.0	—.-	23.1	53.9
Personal hygiene	0.0	0.0	—.-	—.-	Related transport	0.0	75.0	0.0	50.0
Related transport	0.0	100.0	0.0	100.0	Other social	0.0	66.7	0.0	40.0
Other maintenance	14.3	0.0	100.0	—.-	Regular meals	63.5	3.9	40.2	1.2
Personal mental	0.0	0.0	—.-	—.-	Special meals	25.0	0.0	11.1	33.3
Active sports	57.1	0.0	50.0	50.0	Snacks	27.3	4.6	31.6	0.0
Taking a walk	63.6	0.0	63.6	9.1	Other transport	25.0	8.3	13.6	22.7

italic: the matching code is the highest scoring category

Table 6D.7 Activity-codes in ESM-interactions and RIRs: matching analysis by conditional probabilities. “In group” category refers to conglomerate scores except for transport .

Poor matches were found for the social interaction categories such as visits to friends and relatives, going to a pub, parties, etc. A possible explanation is that the social interactions are characterized by those activities (as recorded in the RIR), while the content of the interaction might easily be “conversation”, “eating or drinking” and even “studying” (which will be the ESM report).

Assessing personality traits in time and event sampling

In order to compare RIR and ESM in predicting independent cross-sectional measures of psychologic well-being and social support — as between subject characteristics —we first selected the rating scales that were identical (or functionally equivalent) on both protocols. Eight mood scales fulfilled this criterion, four positive (happy, valued, productive, and satisfied) and four negative (anxious, irritated, sad/depressed, and rejected). The positive adjectives were highly correlated with one another, as were the negative adjectives, as has also been shown in many other studies. Therefore, to simplify presentation, we combined the individual ratings into positive and negative summary indices. Not surprisingly, RIR and ESM ratings of positive and negative affect in daily life were correlated significantly. For positive adjectives, the RIR correlated $r = 0.65$ ($p < 0.001$) with ESM ratings during social interaction, and $r = 0.47$ ($p < 0.01$) with non-social ESM ratings. For negative adjectives, the corresponding correlations

were 0.62 and 0.63, respectively, both $p < 0.001$. Correlations of these daily life indices with cross-sectional well-being and social support are displayed in Table 6D.8.

Measures of psychologic health		RIR	ESM social	ESM non-social	ESM total
• Positive mood (Diener & Emmons)	• positive items	0.45 **	0.26	0.18	0.27
	• negative items	-0.37 *	-0.18	-0.26	-0.25
• Negative mood (Diener & Emmons)	• positive items	-0.18	-0.07	-0.13	-0.16
	• negative items	0.60 **	0.28	-0.57 ***	0.52 ***
• Life satisfaction (Diener)	• positive items	0.22	0.12	0.29	0.26
	• negative items	-0.46 **	-0.24	-0.41 **	-0.37 *
• Loneliness (UCLA Loneliness)	• positive items	0.19	-0.06	-0.26	-0.19
	• negative items	0.40 **	0.10	0.28	0.19
• Psychological well-being (Diff. Person. Q.)	• positive items	0.38 *	0.18	0.24	0.26
	• negative items	-0.58 ***	-0.29	-0.43 **	-0.41 **

Measures of social support		RIR	ESM social	ESM non-social	ESM total
• Tangible (Cohen-Hoberman)	• positive items	-0.08	-0.07	-0.12	-0.13
	• negative items	0.13	0.12	0.14	0.15
• Appraisal (Cohen-Hoberman)	• positive items	-0.29	-0.10	-0.08	-0.11
	• negative items	0.41 **	0.25	0.22	0.29
• Belonging (Cohen-Hoberman)	• positive items	-0.09	0.11	0.25	0.21
	• negative items	0.39 *	-0.11	0.30	0.26
• Esteem (Cohen-Hoberman)	• positive items	-0.39 *	-0.05	-0.02	-0.04
	• negative items	0.38 *	0.12	0.25	0.21
• Closeness (Reis)	• positive items	0.00	-0.08	-0.18	-0.16
	• negative items	0.31 *	0.25	0.22	0.26

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Table 6D.8 Correlations between daily life ratings (RIR and ESM) and cross-sectional measures of well-being and social support.

We next compared RIR and ESM ratings of social interaction³⁶. We aggregated the ratings on the basis of prior factor analyses with these measures in the RIR. This indicated that four scales could be combined into a single composite, labeled promotive interaction: “I helped other”, “Other helped me”, “I self-disclosed”, and “Other self-disclosed”. In addition, for the RIR only, ratings of pleasantness and satisfaction were combined into a single index of “Quality”. Other scales for RIR as well as for ESM were not combined.

		Positive mood	Negative mood	Life sat- isfaction	Loneliness	Psychologic well-being
• Intimacy	RIR	0.19	-0.16	0.08	-0.23	0.44
	ESM	-0.10	0.11	-0.22	-0.09	0.14
• Promotive Interaction	RIR	0.25	-0.06	-0.03	-0.11	0.24
	ESM	0.23	0.02	-0.12	0.21	0.24
• Conflict	RIR	-0.13	0.24	-0.24	0.25	-0.29
	ESM	-0.08	0.19	-0.22	0.06	-0.17
• Quality	RIR	0.41 **	-0.04	0.25	-0.20	0.26
• Closeness	RIR	0.56 ***	-0.33 *	0.41 **	-0.41 **	0.60 ***
• Rather be alone	ESM	-0.12	0.08	0.01	0.09	-0.13
• Pleasant company	ESM	0.31 *	-0.19	-0.01	-0.05	0.28
• Interest	ESM	0.18	-0.31 *	0.01	-0.51 ***	0.35 *

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 6D.9a Correlations between daily life ratings and cross-sectional well-being.

³⁶ There are no ESM non-social data for the social interaction item since subjects completed these scales only if interacting when they were beeped.

As might be expected, the RIR and ESM indices of social interaction correlated highly. These correlations were: intimacy, $r=0.65$; promotive interaction, $r=0.84$; and conflict, $r=0.56$ (all p 's < 0.001). Correlations of these measures, as well as the remaining RIR and ESM indices, with health and social support are displayed in Tables 6D.9a and b.

		Tangible	Appraisal	Belonging	Esteem	Closeness
• Intimacy	RIR	-0.08	-0.35 *	-0.19	-0.38 *	-0.33 *
	ESM	0.08	-0.21	-0.06	-0.16	-0.37 *
• Promotive Interaction	RIR	-0.13	-0.27	-0.22	-0.14	-0.09
	ESM	-0.23	-0.29	-0.29	-0.17	-0.20
• Conflict	RIR	0.08	0.10	0.11	0.14	0.21
	ESM	-0.06	0.04	0.03	0.05	0.06
• Quality	RIR	0.02	-0.36 *	-0.05	-0.30	-0.15
• Closeness	RIR	-0.07	-0.49 ***	-0.37 **	-0.51 ***	-0.42 **
• Rather be alone	ESM	-0.01	0.07	0.18	0.11	0.37 *
• Pleasant company	ESM	0.11	-0.34 *	-0.06	-0.26	-0.36 *
• Interest	ESM	-0.21	-0.40 **	-0.36 *	-0.18	-0.11

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 6D.9b Correlations between daily life ratings and cross-sectional social support.

Discussion

In this study we conducted an analysis of the concurrent assessment of social interaction using a time (ESM) and event (RIR) sampling procedure. The data were originally collected to evaluate the RIR validity. The subjects were university students.

RIR and ESM assessments of social interactions were related. The RIR assessment, however, seriously underestimated the time subjects spent in interaction. This holds both for the total amount of time in interaction, as well as for the number of interactions. Moreover, the RIR assessments did not reflect the changing composition of the interacting group over time — significant levels of alone time were found within the interaction.

The RIR and ESM assessments of mood and social interactions were significantly related, both between subjects and within subject. However, RIR and ESM also resulted in significantly different assessments of experience during interaction. We assume that because ESM ratings are totally moment-bound, the accuracy of the momentary descriptions, mood assessments as well as social context and activities, will be higher in ESM compared to RIR. Although the RIR methodology reduces the time between the event and the report, and is an improvement on the traditional cross-sectional assessment strategies, it is still possible that the subject has to infer the mood state, combining the range of mood swings into one aggregated measure which is assumed to be representative for the whole interaction period. This would result in less variability and more socially desirable responses in the RIR. Analyses were inconclusive, however. The RIR and ESM differences in mean experience showed less socially desirable answers for "help" (both "I helped..." and "Others helped") but more for five other scales. Differences in variability were also inconclusive. We can conclude that both instruments are sensitive and valid to assess the subject characteristics of social interactions. The differences between both instruments can be related to the differential

sensitivity of specific items as descriptors of the subjects' mental state, the persons involved in the interaction or varying moments within the interaction.

The analysis of time use during interactions favors the ESM ratings, because specific low frequency activities are described in ESM but missed in the RIR assessments. Those observations match the observational model that was developed by Delespaul (Chapter 3) to describe the process of awareness of information and aggregation in self-reports. The model predicts that the self-selection of information will cause spontaneous migration from less socially acceptable or communicable activities to more public ones in the more remote or retrospective assessments. While the interval between the observed event and the observational recording is kept small in the RIR, it is longer than in ESM, probably causing the difference in Time Budgets.

In a final analysis we compared RIR and ESM to assess subject traits. RIR assessments are more successful than ESM assessments in generating reports that are similar to the cross-sectional assessments of traits. For both instruments, correlations with the social support items are weaker than with the well-being measures. Low correlations were also found between trait aspects of well-being and social support, and specific RIR and ESM items, with the exception of "closeness" both as trait and state and "interest" as a state. The difference is in line with the expectations of the observational model. RIR assessments are more influenced by the subject's traits. The ESM assessments give a better description of moment-to-moment variability. The low correlations between states and traits are indicative of sensitivity to moment-to-moment variability. If researchers are interested in correlates of trait measures, the RIR may be preferable.

In summary, the ESM and RIR each have their own strengths and weaknesses. We would argue that for the optimal monitoring of salient events, a combination of time and event sampling should be attempted. In this method, a standard ESM sampling procedure should be used as the basic sampling scheme. This allows optimal assessment of time budget and antecedents of the events. An event sampling procedure should be used when the researcher wants to avoid missing events. Ideally, a concurrent event sampling should trigger a computer generated intensive sampling scheme that is superposed over the basic and less frequent ESM time sampling scheme. The salient characteristics of the event monitoring are start- and end-time as well as the event type when more than one event is targeted. Event modalities, however, should be monitored using the superposed intensive time sampling scheme.

The technical aspects of this kind of sampling procedure, which necessarily relies on computerized sampling devices, might scare off subjects and lead to higher drop out rates. They nevertheless allow specific controls over the reliability and validity of the data, for which no alternative exists.

E: COMPARING ESM WITH RETROSPECTIVE ASSESSMENTS ³⁷

In the next section we will assess the validity of ESM for the assessment of daily life activity schedules and contextual reactivity of mood and psychopathology to different environments. We compare ESM data of chronic mental patients with retrospective data rated by those patients and their case managers.

Subjects

The subjects were 20 chronic mental patients who are currently living in a residential treatment setting. 60% were men and 40% women, with mean age of 43 years (range from 22 to 68). 65% had schizophrenia, the other subjects had chronic depression or a personality diagnosis. The "case managers" who rated the chronic mental patients were nurses. We selected those individuals who were in charge of the subject's individual treatment programs and who were on shift the largest part of the ESM week. They are assumed to know the most and should be able to make the most reliable ratings.

Method

After the ESM-assessment by the patient, we asked the patients and their case managers to assess the patient's time budgets for "persons", "places" and "activities", as well as their mood (anxiety and depression), psychopathology and activity motivation in those different situations over the previous week.³⁸

Rating time use

The assessment of time budgets is based on a technique that was first used by Delespaul & Kerpentier (1990). The subjects have to rate the relative amount of time a patient spent in different conditions on a computer screen. The rating of a time budget is a complex assessment which is made user-friendly through a set of flexible rulers which can be manipulated to generate immediate feedback on a computer screen (see Figure 6E.1). The first action is to rate the different time portions according to a rank-ordering of the amount of time spent in the different situations. This is possible by setting the rulers. The software will compute the pie-chart accordingly. The next action is the fine-tuning, which is done by the same rulers. The rater can operationalize expressions such as: I spent twice as much time with family-members, compared to friends. Using this procedure, both the case manager and the patient rate the patient's time budgets for persons, activities and places over the past week. Due to technical problems with the assessment device during the first weeks of the study five subjects were lost.

³⁷ The current analysis was made possible by a grant from the "Nationaal Fonds voor Geestelijke Volksgezondheid".

³⁸ "Time of the day" was added in the "experience in context" - assessments.

Figure 6E.1 Computer screen to rate time budgets: left, the controls; right, the time-budget feedback.

The data of 15 subjects are analyzed in relation to two different aspects:

1. Assessment of time-budgets rated prospectively using ESM and retrospectively using the computer aided assessment by the patients themselves and their case manager (validity for *within*-subject differences in frequencies). We have used three different analysis strategies but will only present data on the third one — the meta-analytic approach:
 - We compared rank-order correlations (ρ) over the coding categories of a time budget by subject for ESM and computer aided ratings. Average ρ 's and average ρ 's after Fisher-z transformation³⁹ — to correct for the minus 1.00/+100 restricted range in correlations — were computed and their significance tested;
 - We compared Pearson r correlations to include the absolute differences in the rated frequencies of behaviors and contexts. Significance is computed for the average correlation without and with Fisher-z transformation;
 - We finally tested the significance of the proportion of significant correlations against the error proportion of 5% (meta-analysis).
2. Assessment of time-use by category *between* subjects: Most aspects of time budgets are group norms and in this sample, systematically skewed time budget patterns were observed (high frequency of “home” and “alone” time). We therefore assessed the ratings for specific categories (e.g. “being alone”) between subjects. We used the different analysis procedures described above to assess the ratings.

Rating experience by context

We have collected retrospective assessments of 4 different experiential states (anxiety - cheerfulness - psychopathology and motivation) in 4 sets of typical contexts (“Who?” (alone, family, friends and strangers) - “What?” (nothing, work, maintenance, leisure, other, transport) - “Where?” (home, work, network, other) and “When?”

³⁹ $r_z = \frac{1}{2} [\ln(1 + r) - \ln(1 - r)]$. However, Monte Carlo simulations have shown that this transformation is not satisfactory for average correlations (Koele, 1984; Silver & Dunlap, 1987).

(morning, afternoon, evening)). Our original intention was to let both the patient and the case manager make the retrospective ratings at the end of the subject's ESM week. However, in contrast to the rating of time budgets and after repeated attempts we concluded that this task was too complex for patients. Therefore, no assessments by the subjects of their own mental state in context are available. The data rated retrospectively by the case manager are compared with the subject's ESM data, of course, only when observations in those contexts were available. ESM is a sampling technique that is conducted in the natural flow of daily life. Therefore, some subjects never engage in specific situations. The rating procedure (see Figure 6E.2) presents the case manager with a set of situations and we ask him to rank-order those situations according to the expected emotional intensity (e.g. "anxiety") for the patient during the previous (ESM) week. Those assessments were made within 2 days after the ESM period.

Figure 6E.2 "Questionnaire" used to rate experience in context.

The data of 20 subjects are analyzed using two strategies.

- The first strategy computes the level of association (Spearman-rho) for each subject (retrospective rating compared to ESM). The mean of those rho's are computed and tested using a t-statistic with n minus 2 degrees of freedom⁴⁰. Because correlation coefficients are restricted to a -1.00/+1.00 range, the computation of means are not best estimators. As a partial solution for this, we computed a Fisher z-transformation on the rho's and tested the resulting mean against the null-hypothesis that there was no relation between rater-assessments and ESM data.

⁴⁰ $t = \rho \sqrt{\frac{n-2}{1-\rho^2}}$ ($z = \rho \sqrt{n-1}$ for $n > 30$) with $\rho = 1 - \frac{6 \sum D^2}{n(n^2-1)}$ and D = the difference between the ranks (Bruning & Kintz, 1977; Gibbons, 1993).

- The second strategy compares the proportion of perfect ordinal ranking (retrospective rating compared to ESM rating), with the random match ⁴¹. Two alternative computations are used. The first computation includes the cases for which the ESM data was only available for one contextual category (which gives an automatic perfect match which is corrected by the random control parameter). We also restricted the computation to the pairs for which at least two categories were rated in both situations. There was no difference in the results. A difference z-score is computed ⁴² and significance levels are determined accordingly.

Results

Within-subject analyses

Table 6E.1a contains the assessment of time use by the patient and the case manager. These assessments are compared with the ESM time budgets. The most important result from this analysis is that, irrespective of the analysis used, only the “who?” and “where?” assessments were reliable. In contrast, the assessment of activities (what?) was unreliable — a trend in the opposite direction was found. Both assessors — patient and case manager — were able to rate the frequencies of “persons” and “places” reliably, but not the “activities”. Interestingly, the retrospective ratings from patients and case managers were more similar to each other than to the ESM assessments. The mean rho’s for “who?”, “what?” and “where?” were 0.66, 0.45 and 0.87 respectively (all significant at the p<0.01 level). The figure for activities is the lowest from the group, but still significant. This means that the retrospective assessments of activities target an equal “knowledge set” shared by patient and case manager and that this set is different from the actual daily life activities.

			Who?	What?	Where?
Patient / ESM (n=15)	<i>rho</i>	range	-0.40 / 0.94	-0.50 / 0.52	0.31 / 1.00
		$\bar{X}$ (s.d.)	0.63 (0.31)	-0.07 (0.39)	0.75 (0.18)
		H ₀ : rho=0.00	z=7.51 ***	z=-0.68 n.s.	z=15.36 ***
		p 0.05 / p 0.01	0.20 (0.20)	-0.07 (0.00)	0.20 (0.40)
		H ₀ : p _o = p _e	z=2.67 **	z=0.29 n.s.	z=4.20 ***
Case manager / ESM (n=14)	<i>rho</i>	range	-0.22 / 0.89	-0.76 / 0.80	0.34 / 1.00
		$\bar{X}$ (s.d.)	0.61 (0.30)	-0.23 (0.41)	0.69 (0.16)
		H ₀ : rho=0.00	z=7.39 ***	z=-1.99 *	z=15.30 ***
		p 0.05 / p 0.01	0.21 (0.14)	-0.07 (0.00)	0.21 (0.43)
		H ₀ : p _o = p _e	z=2.35 **	z=0.29 n.s.	z=4.60 ***

• p<0.10; * p<0.05; ** p<0.01; *** p<0.001.

Table 6E.1a ESM compared to retrospective assessments: Time budgets.

⁴¹ The random match is computed using the probability of perfect matches in random situations

which can be computed as $\sum_{i=1}^n \frac{1}{c_i!}$ with $\begin{cases} n = \text{number of subjects} \\ c_i = \text{number of categories available for subject } i \\ c_i! = 1 \times 2 \times \dots \times (c_i-1) \times c_i. \end{cases}$

⁴² $z = \frac{P_1 - P_2}{\sqrt{\frac{p(1-p)}{n_1} + \frac{p(1-p)}{n_2}}}$ with $p = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}$ (Bruning & Kintz, 1977)

Table 6E.1b contains a between-subject analysis of the retrospective assessments by coded category. 6 of the 37 tests were significant ($z=1.45$; $p<0.06$). Interpretations are not straightforward and maybe even a little speculative because the meta-analysis was not significant. The case managers — not the patients — were able to assess the amount of time spent working (what? and where? category), which is not surprising because “having work” is a subject characteristic in this population. They were able to rate the amount of time the patient spent with friends. Patients were able to assess the amount of time spent alone and with strangers. Considering that this analysis is between-subjects, this is impressive. Subjects knew that they were more often alone than their fellow patients. Negative non-significant high correlations were found for “doing nothing” and “social contact” (case managers only). This means that patients who describe themselves as often “doing nothing” are, in fact, more active than those who think they are active. In those situations, the retrospective assessments were poor.

ESM versus	Who?		What?		Where?	
Patient (n=15)	Alone	0.56 *	Nothing	-0.38 n.s.	Home	0.46 n.s.
	MH professionals	-0.13 n.s.	Work	0.16 n.s.	Work	0.45 n.s.
	Inmates	0.32 n.s.	Maintenance	0.00 n.s.	Family	0.31 n.s.
	Family	0.08 n.s.	Leisure	-0.08 n.s.	Friends	.
	Friends	0.31 n.s.	Social contact	0.09 n.s.	Public in	0.59 *
	Strangers	0.63 *	Thinking	-0.22 n.s.	Public out	0.16 n.s.
			Meals	-0.10 n.s.		
Meta-analysis	(2/6)	$z=$ 1.25 *	(0/7)	$z=$ -0.60 n.s.	(1/6) (3/19)	$z=$ 0.65 n.s. $z=$ 1.09 n.s.
Case manager (n=14)	Alone	0.07 n.s.	Nothing	-0.51 n.s.	Home	-0.05 n.s.
	MH professionals	-0.01 n.s.	Work	0.56 *	Work	0.55 *
	Inmates	-0.33 n.s.	Maintenance	-0.07 n.s.	Family	0.26 n.s.
	Family	-0.06 n.s.	Leisure	-0.01 n.s.	Friends	0.19 n.s.
	Friends	0.56 *	Social contact	-0.34 n.s.	Public in	0.26 n.s.
	Strangers	0.25 n.s.	Thinking	-0.07 n.s.	Public out	-0.02 n.s.
			Meals	-0.06 n.s.		
Meta-analysis	(1/6)	$z=$ 0.65 n.s.	(1/7)	$z=$ 0.59 n.s.	(1/6) (3/19)	$z=$ 0.65 n.s. $z=$ 1.09 n.s.

* $p<0.10$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$.

Table 6E.1b ESM compared to retrospective assessments: Between-subject differences.

Between-subject analyses

Table 6E.2 contains the assessment of experience in context. Only the ratings of the case managers were available. Starting with the assessment of perfect matches, we see that the case managers were not able to rank the effect of context on depression (cheerfulness) and motivation. Only two ratings were significantly better than random, i.e. the intensity of psychopathology during different activities ($z=1.92$; $p<0.03$) and anxiety in places ($z=1.72$; $p=0.042$). Considering the fact that we computed 16 tests we should not attribute too much importance to those significant relations (testing the significance of significant tests: $z=0.75$; n.s.).

The first analysis, using observed and expected proportions, considers only perfect matches. The next analysis also takes partial matches into account. In this situation, we see that the case managers did a fair job in assessing activity and place-effects on mood and mental state. The assessment of the effect on persons (significant only for cheerfulness) and time of the day (significant for psychopathology) was less clear. Overall, the mean correlations are only satisfactory for the place relation of mood

(range: 0.57 - 0.82). Although significances were found for time and persons, the correlation coefficients ranged from -0.31 to 0.14 and -0.04 to 0.38 respectively.

		Who? (# cat =4)	What? (# cat =6)	Where? (# cat =4)	When? (# cat =3)
Anxiety	• $p_o (p_e)$	0.22 (0.20) $z=0.12$	0.22 (0.13) $z=0.64$	0.87 (0.57) $z=1.72 *$	0.27 (0.22) $z=0.38$
	• mean rho (n)	0.10 (17) $t_{(15)}=0.38$	0.52 (17) $t_{(15)}=2.35 *$	0.82 (11) $t_{(9)}=4.29 **$	0.14 (18) $t_{(15)}=0.56$
Cheerfulness	• $p_o (p_e)$	0.40 (0.20) $z=1.36$	0.21 (0.11) $z=0.76$	0.63 (0.53) $z=0.91$	0.10 (0.21) $z=-0.95$
	• mean rho (n)	0.38 (19) $t_{(17)}=1.69 *$	0.56 (18) $t_{(16)}=2.70 **$	0.57 (12) $t_{(10)}=2.19 *$	0.04 (19) $t_{(17)}=0.16$
Psychopathology	• $p_o (p_e)$	0.25 (0.20) $z=0.36$	0.35 (0.09) $z=1.92 *$	0.75 (0.53) $z=1.47$	0.30 (0.21) $z=0.60$
	• mean rho (n)	0.11 (19) $t_{(17)}=0.45$	0.44 (19) $t_{(17)}=2.02 *$	0.73 (13) $t_{(11)}=3.54 **$	0.12 (20) $t_{(17)}=0.51$
Motivation	• $p_o (p_e)$	0.11 (0.18) $z=-0.57$	0.17 (0.10) $z=0.64$	0.72 (0.55) $z=1.07$	0.11 (0.20) $z=-0.69$
	• mean rho (n)	-0.01 (16) $t_{(14)}=-0.03$	0.39 (16) $t_{(14)}=1.58$	0.75 (11) $t_{(9)}=3.40 **$	-0.31 (17) $t_{(17)}=-1.26$

• $p<0.10$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$.

Table 6E.2 ESM compared to retrospective rated assessments by the case manager: $p_o (p_e)$: observed proportion (p_o) of perfect match compared to expected proportion (p_e); and measure of association (rho) compared to no relation at all.

Discussion

In this comparison between ESM and retrospective assessments by the case manager and the patient self, we repeatedly switch from a position where the external criterion assesses ESM validity, to a position where ESM is the criterion to determine the validity of the other assessment strategies. In fact, ESM data can be considered the best criterion for evaluating both the retrospective assessment of time use, as well as the experience in context. Except for the situation of fraud there is no reason to believe that when someone says in ESM that he is alone he is, in fact, not alone. What can we say about the assessments of the patients and the case managers, compared to the ESM observations? The distribution of the ratings over the different subjects for a specific coding category were more similar for both retrospective ratings, compared to the ESM observations. The ESM ratings are a little more skewed (positive side) but especially flatter, giving more differentiation in the ratings.

	Rating		
	Patient	Case manager	ESM
Mean skewness	1.0	1.0	1.6
Mean kurtosis	1.5	1.1	4.0

Table 6E.3 ESM assessment compared to ratings: characteristics of the response distribution.

This similarity/dissimilarity pattern between ratings — both from the patient and the case manager — and the ESM observations was replicated in different places. The adequacy of the assessments by the patients, compared to those by the case managers, is striking for this highly invalidated group. However, the living environment for the patients was a hospital ward and this explains why the ratings of “places” and “persons” have a good quality. Those ratings are the reflection of the hospital living situation. In contrast, the assessment of activities was less adequate because chronic

mental patients on hospital grounds have a lot of free time between 7:30 and 22:30. Scheduled therapy takes a maximum of 25% of that time. However, the quality of the assessment of places and persons found in the within-subject analysis vanishes when we look at between-subject differences. It is easier to assess subjects being on the ward most of the time, than to rate who is at home most. In the assessment of time use the general trend is not interesting. We know that a person who lives alone will spend a lot of time alone and changes in that pattern as a consequence of an intervention, for instance, will only be marginal. Therefore, the assessment of small differences is more interesting. We conclude that the retrospective assessment is inadequate.

Baars (1994) assessed the social networks of chronic mental patients and compared the ratings of the patients themselves and their mental health professionals using the "Maastricht Social Network Analysis" (MSNA), a standardized network interview. Because he used a structural approach to social networks — emphasizing the number of persons — comparisons are difficult to make. He found differences in the sector of friends and acquaintances with lower estimates by mental health professionals compared to patients. Our retrospective assessments by patients and mental health professionals were more similar than prospective ESM assessments. While patients were good assessors of alone time, mental health professionals made better assessments of friends. Once again we have to acknowledge that we compared the frequency of interactions while the MSNA compared number of persons by category in the network. Even more interesting from a clinical point of view are the contextual differences in experiences. We want to know how a patient feels when he is alone at home. This information makes the difference between advising changing the living environment or accepting that environment as a fact and working, for instance, towards a better differentiation of the subject's activity pattern based on the assessment of positively experienced activities. The assessments for "time of day" and "persons" were poor. But using rank-order correlations the assessment of experience and mental state could be related to activities and places. We presented no between-subject assessments because these are even less adequate than the within-subject comparisons.

In comparison with the study by Baars (1994), overall the mental health professionals rated social interactions with lower affectivity than did the patients themselves. We, unfortunately, had no retrospective patient ratings of affect by interactions. None of our retrospective ratings by the mental health professionals were satisfactory.

In summary, while demonstrating the validity of ESM is a difficult job, we have illustrated that ESM data matches subject specific knowledge of living environments and, more specifically, experience in context. For the non-matching situations, we have argued that the ESM data are the most valid one. A new study using yoked sampling — self-observations and observations by others — may solve the uncertainties for the patterns for places and persons, but never for the experience in context question. Here, we can only rely on the subjects' verbatim expressions, for which no other external reference is available. Fortunately, the case manager's knowledge of experience in context is fair, but only for global categories. This has consequences for the assessment of "quality of life". QoL not is reflected by the living situation of our patients, but by the activities performed in this situation and the emotions that go with them. Retrospective assessments by case managers will not yield sensitive QoL measurements.

F: DISCUSSION: SAMPLED EXPERIENCE IN OUTCOMES OF CLINICAL TRIALS

Double blind randomized clinical trials

Modern psychiatry increasingly needs information that allows us to optimize treatment selection in clinical practice. This is important to offer patients the best possible service and alleviate suffering as soon as possible. The availability of differential outcome information is also essential to be accountable for the insurance companies. Double blind randomized clinical trials are the methodology of choice to assess scientifically if a specific intervention is effective. The clinical trial is designed to ascertain that the intervention (the independent variable) is the only influence that causes the differences in the results between the experimental and the control conditions (the dependent variables) (Pockock, 1983).

The reason why we randomize in clinical trials is to maximize the likelihood that the subjects that are referred to the different experimental and control conditions have the same characteristics before the intervention. As a consequence, not the preexisting subject attributes but only the intervention can be held responsible for the observed difference in outcome. While matching subjects can be seen as a good alternative, randomization also allows control over subject characteristics of which the effect on the outcome of the intervention is unknown.

The aim of the double blinding is to limit the knowledge of the actual clinical intervention and conditions to the subjects as well as to the evaluating observers involved in the trials. Since expectations about mechanisms of action and expected effects of the experimental conditions could skew the responses that are given on the assessment instruments, *a mechanism of built-in naïvety is essential. When the anticipation of the treatment effects are randomly divided over the different experimental conditions — the raters guess the membership of the subjects randomly — this represents no real problem.* However, they are important confounding influences when such expectations follow a systematic trend or bias. To avoid such bias the double blind study seems ideal.

Two fields are important for the improvement of our therapeutic potential in the clinical practice with schizophrenia. In both fields, the successful implementation of double blind clinical trials is not self-evident. First, we depend on the advances of biologic psychiatry such as the new neuroleptics for depression and schizophrenia (Kane, 1991; Addington & Addington, 1993). The central place of pharmacologic agents requires renewed attention to testing, assessment and the quality of information clinicians receive to implement treatment. The challenge for research is particularly true for new classes of drugs that have markedly different clinical and side effect profiles (Awad, 1993). Secondly, biologic psychiatric treatments should be supplemented with psychosocial interventions (Lieberman et al., 1994a,b; Falloon & Fadden, 1993). The complexity of these interventions and the problems related to the standardized implementation of those protocols make them very difficult to research.

In this section we will discuss some problems related to the implementation of the randomized clinical trial methodology for testing new neuroleptics and psychosocial interventions. We will argue that blind experimental designs have many drawbacks and that we need a huge number of subjects to assess the experimental effects with enough power in the current research context. This has resulted in expensive multicenter strategies in psychopharmacologic research. Multicenter studies should also be used in socio-psychiatric evaluation research in order to generate the information for clinical decision-making⁴³. We will plead that the patient's subjective experience is an ally, not a confounder, of research and can yield outcome measures that are more sensitive to change and can result in a reduction of research costs.

Problems of keeping subject and rater blind to conditions

Individuals, both subjects and raters, who contribute to the generation of the outcome measures should be naive or blind concerning the experimental condition. Double blind strategies assume that neither the patient nor his treating psychiatrist knows which intervention is applied. This is necessary, so that only the effect of the experimental condition, and not prior knowledge or expectations, will influence the reports. A blind assessment strategy aims at avoiding unintentional systematic influences on the results, jeopardizing the validity of the research. But, no matter how well planned, in current clinical practice a double blind clinical trial is problematic since the prerequisites of the double blind assumptions are often difficult to meet.

Today, in evaluations of neuroleptic drugs for psychosis and depression the double blind condition is extremely difficult to maintain. Since the pharmacologic and clinical profiles of the new antipsychotic and antidepressive drugs are radically different from their classical alternatives that are used as control drugs, a significant error component may be introduced (Kane, 1991; Awad, 1993). As a consequence, the blind condition cannot be reliably maintained for the treating physician, who cannot fail to recognize the substance. Moreover, the patients who have often had other treatment experiences will also easily recognize the new drug, if he or she has previously been treated by a traditional compound and knows what effects to expect. Paradoxically, the gains of the new medicines make them difficult to study and challenge the validity of the current clinical trial paradigm. As a result, the responses to clinical trial assessment instruments are likely skewed by the respondent's experience and related anticipation of the working mechanism of the drugs.

Because the therapist is actively involved in the application of psychosocial interventions it is impossible to maintain blinding for him. Equally so, the subjects themselves will be informed of the different therapeutic options under study and will easily recognize the experimental condition that applies to them. The working mechanisms of psychosocial interventions cannot be dissimulated that easily as the nature of the pharmacologic compound. Keeping subjects and researchers blind to the experimental conditions, therefore, is an impossible endeavor in psychosocial research.

⁴³ Of course, due to the difficulties in standardizing the interventions in socio-psychiatric research such studies would be difficult to conduct.

However, controlling measurements by blinding is important since most assessment instruments for clinical outcome, side effects and quality of life rely on the subjects' responses rated by the clinicians. One must then entertain the possible influence of unintentional bias as a consequence of the fact that the blind condition was difficult to maintain. These threats to validity may have a major impact on assessing the effect of the interventions in core psychiatric conditions. Moreover, since the relevant difference in clinical effectiveness in psychiatric outcome research is small — when comparing between active interventions as well as against placebo conditions — one can assume that a small bias introduced by deficiency in the blindness could seriously affect the results. While bias potentially affects the patient and the physician alike, the problem is compounded for the treating psychiatrist who has full knowledge of the assumed differential working mechanisms of the experimental conditions.

Solutions for ineffective masking of the experimental conditions

There are different strategies available to circumnavigate the problems that occur if blinding is broken. The first one is to use objective — rater independent — assessment instruments. This strategy is useful for some target measurements, such as counts of sick days, discharge and readmission rates or measuring neurotransmitter binding using MRI assessments. But, as developers of new neuroleptics testify (Meltzer, 1991), these objective measures are poor predictors of clinical effects, since subtle side effects and changes in experienced quality of life are often shown. Such assessments rely on the subject's own introspection and as a consequence cannot be assessed without some bias.

Another strategy is to introduce a blind rater who has no other knowledge of the clinical course of the patient, than the observations he makes during the assessment interviews. Contrary to the patient and treating psychiatrist (who must know the “ins and outs” of a study from a practical and ethical point of view), the persons hired only for assessments can be kept in the dark or may even be deceived by false information about the kind of intervention. Raters may be chosen with no prior knowledge of the assumed working mechanisms of the experimental conditions. The use of naive assessors, of course, has drawbacks. Non-professional raters could still generate bias during the interview and especially under repeated assessment conditions. Furthermore, the subjects themselves can influence the raters with biased self-assessments in response to the interviewer questions. Experimenting with different rater strategies could in any event improve the validity of the study but new assessment techniques are required. Those techniques should be less influenced by subjective bias due to prior knowledge and more sensitive to assess change.

Innovative assessment: the Experience Sampling Method (ESM)

We believe that new assessment instruments should rely more on the subject's self-assessments and less on the assessments by the psychiatrist (see Davidhizar, 1985). Since normally raters seem to be subject to bias, the best choice is to base assessment strategies on the patients themselves who have the least vested interest as to which treatment is being used — as long as they are adequately helped. However, most often

schizophrenics are considered poor assessors of their mental state (Frith, 1992). Reasons for the lack of reliability of subjective reports have been discussed by Reis & Wheeler (1991). They argue that the generation of retrospective assessments is a complex cognitive process. The raters have to select the appropriate events that may not be well-defined for the subject, remember them and aggregate the materials into a summarizing rating. For instance, it is possible that a schizophrenic subject may not be aware that specific events were hallucinations, he may not remember the 2 days of florid hallucinations of the previous week, or hallucinations due to their length and intensity may be weighted differently.

Given these problems, the prospective monitoring of mental state is a solution for the reliability problem in the assessment of such subjective experiences. For this, we have developed the Experience Sampling Method (ESM) (deVries, 1992) a random time sampling technique in which subjects are randomly and repeatedly signaled to register these answers on a questionnaire over a period of 1 week, 10 times each day to self-monitor their current mental state and context. This method avoids problems of memory or reconstruction. Only one moment is described at a time and no selection or aggregation of experience is involved.

Today, the validity of ESM for the assessment of clinical effectiveness has been demonstrated in a wide range of psychopathologic illnesses such as schizophrenia, depression, anxiety, somatization disorders, pain, stress and addiction (deVries, 1992). Using the ESM assessment strategy I have demonstrated that the lack of reliability holds primarily for retrospective assessments ("How did you feel over the last week?") but less for prospective assessments ("How do you feel now?") (Chapter 6, section B). In this study I observed a trend towards a response drift and significantly reduced variability, suspicion of answers that are influenced by social desirability, even after very short delays of 15 minutes ("How did you feel 15 minutes ago?"). I also analyzed ESM data in comparison with retrospective assessments collected daily using a traditional diary technique (Chapter 6, section C) and a sophisticated assessment technique to assess the past mental state and activity patterns at the end of the week (Chapter 6, section E). Compared to ESM, both assessments are less intensive strategies which make them inexpensive assessments that should be considered more often in evaluation research. However, ESM assessments were often substantially different from other instruments. Because the moment-to-moment assessment of ESM is embedded within the flow of experience and minimizes the cognitive aggregation, the validity of the retrospective instruments should be doubted. A fourth study compares ESM with the RIR, an intensive event sampling strategy (Chapter 6, section D). In that study I was able to demonstrate that ESM was able to assess social interactions reliably and I showed that different mental state items of the ESF's might have a differential sensitivity to change over time (characteristic of the subject, the event or the moment within the event).

When we aggregate the moment-to-moment mental state reports over a 1 week period ESM yields stable assessments of subject characteristics as I demonstrated in a comparison of 15 subjects over a 6-month period in which no intervention was made (Chapter 6, section A). The combination of the sensitivity of moment-to-moment individual reports and the stability over time of aggregated summaries makes ESM a

good candidate to generate outcome variables that can be used in evaluation research in psychiatry.

ESM in clinical trials

In clinical trials, we are interested in comparing the mental state of the post-intervention period with the baseline mental state before treatment. The one-time prospective assessments currently used — clinical interviews and standardized questionnaires — provide skewed reports that have great error variability and threaten the stable and representative assessments of a subject's mental state, making clinical trials cumbersome.

Tests of the usefulness of ESM for the measurement of side effect profiles and quality of life are currently under development. The optimal research design for applying this method in clinical trials is currently being tested. The ESM random time sampling method is designed to de-emphasize or obscure the goal of the assessment and instead emphasize the current emotions and behaviors of the subjects. This enables the evaluation of moment-to-moment changes. Rather than focusing on the assessment of clinical change, the frame of reference in ESM assessment is the description of their immediate personal experience at that moment. Patients rate the intensity of their emotions, cognitions, and activities on 7-point Likert scales. At the same time they describe what they are currently doing, where they are, and with whom. Patient profiles of behavior, emotional and cognitive reactivity, as well as symptoms in natural settings may be manufactured. Moreover, the moment assessment de-emphasizes the more general time frame of the pre- or post-intervention period as well as the specific research targets (therapeutic effect), thus avoiding untoward influence of response style. Assessing only the current moment and the research instruction to immediately respond to signals minimizes the influence of response bias and defense mechanisms rendering a response similar to free association in psychoanalysis. Although ESM is not a panacea for obtaining error-free assessments, we are convinced that the prospective monitoring of experience under different experimental conditions substantially improves current assessments.

The repeated assessments of current mental state that are sampled over a week are probably the best possible assessment of the subject's actual mental states. Equally so, the summary scores that are mathematically computed using those sampled reports are an optimized reflection of the actual mental state of the subjects for the assessment week than are the traditional one-time retrospective measures obtained from interviews or questionnaires (deVries & Delespaul, 1989, Discussion of Chapter 3).

Moreover, ESM responds to the need that the time course of outcome factors requires special consideration in assessing change in clinical trials (Gaebel, 1993). The possibility of predicting clinical change in earlier stages of intervention, would be a very important finding for clinical practice. Predictions could be made and tested, about the expected time frames of changes for the different types of outcome markers currently in use. Restricted to the assessment of neuroleptics, biologic markers such as plasma concentration levels are immediate and highly sensitive to medication intake, but predict clinical change poorly (Van Putten et al., 1991). We think that the data

generated by ESM approximate the almost immediate response of biologic markers to medication, but uniquely at the symptom and behavioral experiential level (Delespaul & deVries, 1987) and can be used in evaluation of neuroleptic as well as psychosocial interventions. If this can be substantiated in further studies, ESM will be promising as a sensitive clinical change predictor at the early stage of interventions, because it provides quantitative information on behavioral reactivity, emotional experience and time use early in treatment. Those parameters represent a more detailed picture of the dependent variable, the outcome condition. Such data result in production of behavioral and emotional profiles of experience and activity, that more fully illustrate therapeutic effects.

Conclusions

Prospective repeated measures are not only possible but also provide the means with which to measure the subjective experiences and behaviors of individuals with minimized reporting bias. ESM provides this data in a quantitative format that can be used in clinical evaluation research. Subjective experiences including assessments of quality of life are vital to understanding the effects of treatment interventions. Measured using ESM they are sensitive to differences between diagnostic groups and allow the assessment of within-subject differences over time — possibly also the assessment of treatment effects.

We argue that it is important for psychiatry to reevaluate the current practice of double blind clinical trials and argue that in most current applications the prerequisite of the blindness is no longer feasible while the number of subjects needed in each study makes it virtually impossible to plan enough studies to allow the differential weighting of a large number of intervention strategies. New assessment techniques such as ESM may, however, provide the most significant advance. ESM serves as a reliable self-assessment adjunct that alleviates key problems related to defining dependent variables and outcomes in clinical trials. The sensitivity of ESM to detect changes its application in psychiatric evaluation research is promising.

Figure 1. The figure shows a horizontal line with a textured, noisy appearance, likely representing a data series or a plot element.

Chapter 7: Analyzing ESM data ⁴⁴

“Danger, lives have been lost!” reads a sign on the rocks by a raging cascade near one of the authors’ place of birth. A similar warning should accompany Experience Sampling data, for a novice could easily squander many years on the slippery terrain of ESM data analysis. Such a warning should be in bright neon for students trained in the positivist framework of experimental science; like the daily experiences they attempt to mirror, ESM data have a complexity which defies traditional textbook analysis. Rarely do they fully oblige the stringent assumptions of inferential statistical tests. The many factors impinging on daily life, and thus upon these data, demand a constant vigilance to statistical artifact and possible confounds. Therefore, even the seasoned ESM researcher takes many trips back to the computer before finalizing an analysis in a way that most clearly expresses the relationship he or she wishes to capture.

Having issued this Surgeon General’s warning, we can now go on to say that, given the right frame of mind, ESM data analyses can be rich, rewarding, and fun. This article is intended as a kind of *Joy of Cooking* for the uninitiated. It lays out basic approaches to statistical analysis of ESM data, identifying the potentials and shortcomings of each. Our underlying philosophy is that the study of daily lives is an interpretive science — there are many valid and interesting ways to address the same question and no final or “correct” answers. The goal is not prediction and control, but description and deeper levels of understanding. The process of analyzing ESM data, therefore, is a process of continual rethinking what findings mean, what factors might be influencing an apparent relationship, and how an analysis might be redone to better represent the phenomenon of interest. It involves a kind of “creative worry,” which, given a sufficient computer budget and some capacity for humor, can be highly engaging and enjoyable.

At heart most statistical issues are conceptual issues, thus in this chapter we first talk about the clarification of research questions. One of the most frequent problems of ESM data analysis is confusion about what specifically is being asked. We differentiate between two basic types of questions that can be addressed: a) *questions about situations*; and b) *questions about persons*.

In the second section, which forms the body of the paper, we discuss how these questions about situations and questions about persons can be analyzed using different

⁴⁴ This chapter provides a guidebook for doing simple analysis with the complex data obtained by experience sampling. We describe the basic types of questions investigators may ask, and discuss the strengths and limitations associated with different approaches to testing these questions. Because ESM data reflect the complexities of everyday life, no analysis can fully capture the underlying trends, and few analysis will fully meet the assumptions of statistical tests. Nonetheless, we argue that careful and creative analysis of ESM data can provide provocative information about the daily lives of a research population. It was originally published as: Larson, R. & Delespaul, Ph. A. E. G. (1992). Analyzing Experience Sampling Data. A guidebook for the perplexed. In M. W. deVries (Eds.), *The Experience of Psychopathology*. Cambridge: Cambridge University Press.

subsets of the complex ecological dataset. The discussion will be illustrated with examples and is written to provide rules of thumb as to when different types of approaches might be appropriate.

We should note that the focus of guidebook is the design of an analysis that addresses a single question and results in one finding, table or graph. For a paper or article one will want to combine several such analyses, probably synthesizing questions about situations and persons. Indeed the richness of an article often comes from the juxtaposition of findings and the tensions or congruence found between the situational level and the person level. Here, however, we are able to comment little on the aesthetic of bringing together a set of findings, and speak more narrowly to the statistical technique of finalizing one analysis.

FORMULATION OF RESEARCH QUESTIONS

Psychologic states can be properties of situations and of persons. There are some situations, for example, that create anxiety in most people (e.g. driving in fog) while other situations are specific to groups of subjects (e.g. elevators); and there are some people who experience anxiety in most situations (e.g. anxiety patients). The excitement of ESM is the potential convergence it offers between these questions about situations and questions about persons: the same data set allows one to consider both. This combined data set contains, however, a major pitfall: we can easily forget the distinction between both sets of questions and carry out analyses that blur person and situation together. Thus, do not mistake a delirious moment with the trait of delirium. Prior to conducting an analysis, it is essential to determine whether it will address a question about one or the other. There are some complex circumstances in which one might want to consider the two within the same analysis, but these are fewer than one would think and the beginner is strongly encouraged to frame the analysis around either the situation or person before going near the computer.

We will present some examples of questions about situations and questions about persons to clarify what each is. Afterwards we will discuss the possible confusion that arises between the two types of questions when the aim of the analysis is unclear.

Questions about situations

For our purposes here we define “situation” very broadly. Questions about situations include comparisons between contexts of daily life, as defined by people’s activities (e.g. work vs. leisure), companionship (being with family vs. friends), location, or time of day. “Situations” might also be differentiated by subjective states (e.g. the experience of craving or panic), or by any other criterion one can think of that differentiates moments in time (e.g. the smoking of a cigarette, a stressful event). The independent variable for these questions will most often be a datum recorded on the ESM self-report form, such as a contextual variable or a mood score, but might also be derived from a prior ESM sheet, a daily diary report, or some other criterion that differentiates categories of experience in daily life.

Here are some examples of questions about situations:

- Are people more motivated in work or in leisure?
- Do drug users experience a greater craving for the drug in the morning or evening?
- What psychologic states precede an epileptic episode?
- How long does a panic attack typically last?
- How are patterns of daily moods and time use different before and after a divorce?

Observe that the concern in these questions is not with comparisons between groups of people. Rather they entail comparisons between moments in time, but within individuals. They are concerned with describing or contrasting some pattern of experiences that differ systematically between situations for a target population.

Questions about persons

Now we switch to looking at the other type of question. Questions about persons include comparisons between groups of people, for example comparisons between people having different traits. One's independent variable may be a demographic variable (age, sex, SES), a diagnostic category (schizophrenic vs. normal control), a score on a test, or even a score derived from the ESM data (subjects who report frequent panic attacks vs. those who do not). Here are some examples:

- Are women more emotionally variable than men?
- How is the daily experience of handicapped children different from non-handicapped children?
- Do patients treated in a hospital experience more social interaction than those treated in the community?
- Do the patients' reports of agoraphobia in a clinical interview correspond to greater agitation when they are with people?

Questions apt to create confusion

In some cases researchers may be unclear whether they are interested in situations or persons. Consider the next two questions:

- Is TV viewing correlated with depression?
- Is frequent hallucination associated with anxiety?

In our experience it is not uncommon for students to get excited about questions phrased like these, without being clear about what they are really asking. The wordings used here leave it unspecified whether the issue being addressed is depressed and hallucinating people or depressed and hallucinatory situations. Thus it is essential to decide whether your questions are about persons:

- Is the amount of TV a person watches associated with the severity of the person's depression?
- Do psychiatric patients who describe themselves as frequent hallucinators also experience more anxiety [than those hallucinating less]?

Or about situations:

- Are [depressive] persons more depressed when watching TV?
- Do psychotic patients who describe themselves as frequent hallucinators also experience more anxiety when hallucinating?

Failure to make this differentiation can lead to erroneous conclusions. In the ESM literature researchers sometimes have, for different purposes, published results on both person- and situation formulations of the same general question. Two examples illustrate how an ill-defined question can lead to findings which obscure the underlying relationships.

EXAMPLE 1: A novice investigator might naively ask, "Is the experience of subjective control correlated with happiness?" This question fails to distinguish between two separate issues: (1) Are happy experiences associated with greater control? (a question about situations) and (2) Are happy people also people who experience more overall control? (a question about persons). Larson (1989) demonstrates that for normal samples the answers to these two questions are quite different. Happy experiences are not strongly related with subjective control, but happy people *do* experience substantially greater average control. Failing to distinguish the two questions would have resulted in these two separate relations being blurred and probably in the findings being misinterpreted.

EXAMPLE 2: Larson (1979, see also Csikszentmihalyi & Larson, 1984) provides another example of the difference between these two levels of analysis in his investigation of how time alone is related to affect among adolescents. Across all categories of teenagers in his sample the immediate experience of being alone was one of lower affect: they felt worse alone than at other times. From this finding one might expect that individuals who spent more time alone would report lower overall affect, but this was not the case. Among older adolescents, those who spent an intermediate amount of time alone (25% to 35% of waking hours) reported the highest average affect. Paradoxically, spending time in this *relatively* negative context was related to higher, not lower overall average mood states. Differentiation between the situation of being alone and the personal "trait" of spending time alone was critical to understanding the underlying relationships.

From these examples it becomes clear that ill defined questions may become a threat to the internal validity of an analysis. The conclusion that those adolescents who spent much time alone must feel worst because the overall sample shows lower affect when alone can not be derived from the data and is therefore an error. We will use the term "Perspective Fallacy" for the error of mixing up person- and situation-level questions in the wording of the research question, in the setting up of the analysis, or in the formulation of concluding comments. In the next section we will discuss how to analyze both questions about persons and questions about situations and show how the "Perspective Fallacy" may easily be introduced in the actual setting up of the analysis.

ANALYZING RESEARCH QUESTIONS

The research questions in ESM studies, both those about situations and persons, can be analyzed using two different configurations of data. One type of analysis uses the individual self-report, the "beep", as the fundamental unit or case. The other uses the subject (the person) as the unit of analysis. The two types of analyses correspond to two types of datafiles, a beeper file and a subject file, which are described in the appendix. We will adopt the term "beep-level" to refer to analyses using the beep as the unit of analysis and "subject-level" to statistical approaches using scores aggregated for each person or one time subject measurements (e.g. age, diagnosis, sex) as the reference case in the analysis. Readers are warned not to confuse questions about

situations and persons with analyses that use the beep or the subject as the unit of analysis. All four combinations of questions and analyses are considered below.

In general, beep-level analyses are simpler, more flexible and easier to understand and communicate than most of the subject level analysis. Subject level analyses, however, are more complex considering both the (unexpected) effects of aggregating variables and the sophistication of the analysis, but typically provide a statistically more conservative test and reduce the probability of both type I and II errors. In the following section we will discuss and illustrate how we can approach the analysis of questions about situations and questions about persons respectively and show the possible strategies that are available for the most appropriate tests.

Analyzing questions about situations

Are people happier with their friends or families? How long does an anxiety episode last? In the following sections we will discuss how these kinds of questions about situations can be analyzed with a beep and a subject level approach.

Beep-level analysis

EXAMPLE 3: Csikszentmihalyi & Larson (1984) used beep-level analysis to describe first, how normal adolescents divide their day between different activities and second, how they feel in these activities. In these analyses the activities define the "situation". For the first one the authors computed a frequency distribution on their Beeper-File for the variable "Activity", which contained codes for what the teens were doing at the time of each signal. The frequency of these codes over all beeps ($n=2734$) forms a "time budget" for adolescents as a group and suggests that socializing is the most frequent activity for them. The authors did not compute significance figures for these percentages, but might have done so by computing confidence intervals based on the standard error of a percentage.⁴⁵

Second, the adolescents' feelings in each of these activities were computed using analysis of variance on the z-scored version of their beeper file, with mood as the dependent variable and activity as the independent variable⁴⁶. Positive scores point to feelings above the subjective average, while negative scores are below average moods. Because the n for these analyses was large, these ANOVA's were wildly significant. In order to provide a more specific test for each activity they computed whether the average state for each situation was significantly different from 0. These analyses revealed, among other things, that sport was one of the most positive experiences in the adolescents' lives (Table 7.1).

Other similar analyses on the daily experience of enjoyment are provided by Massimini et al. (1987) and on the experience of drug and alcohol use by Larson et al. (1984).

⁴⁵ $S\% = 100 \times \sqrt{\frac{p(1-p)}{n}}$, with p the proportion and n the sample size. The Normal distribution fits this statistic. The confidence interval is $p \pm 1.96 \times S\%$ (at 5%) or $p \pm 2.58 \times S\%$ (1%). The statistic is inadequate for extreme values, especially when $n \times p$ or $n \times q$ is less than 5 (Ferguson, 1959).

⁴⁶ The normalization process which is referred by the z-scoring in this chapter is computed by the formula $z_{ij} = \frac{(x_{ij} - x_{i.})}{sd_i}$ (with i the subject index, j the beep index within subject, x_{ij} the assessment of subject i at beep j , $x_{i.}$ the mean of subject i over all the beeps j and sd_i the standard deviation of subject i). The process is repeated for each subject and each item independently.

Activity	Frequency	Affect	Activation	Concentration	Motivation
Studying	346	-0.27 ***	-0.13 **	0.49 ***	-0.44 ***
Eating	153	0.30 ***	0.06	-0.29 ***	0.45 ***
Rest & Napping	87	-0.59 ***	-1.37 ***	-0.94 ***	0.57 ***
Sports & Games	93	0.22 **	0.80 ***	0.36 ***	0.48 ***
Art & Hobbies	41	-0.01	0.77 ***	0.61 ***	0.65 ***

* p<0.05; ** p<0.01; *** p<0.001

Table 7.1 Mean z-score data by situation (Csikszentmihalyi & Larson, 1984).

EXAMPLE 4: deVries et al. (1987) used beep-level analyses to examine the recovery of psychiatric patients from anxiety episodes. In order to do this they identified a subset of occasions in which anxiety was above +1.00 s.d. compared to the subject's own mean. They then draw curves of the mean levels of anxiety on subsequent self-reports. These analyses demonstrate that these patients had a very slow rate of recovery from anxiety episodes, much slower than that of normal controls and, surprisingly, slower than their recovery from depressive episodes (Figure 7.1). Procedures for evaluating the significance of these analyses are suggested by similar analyses carried out by Larson et al. (1980).

Figure 7.1 Recovery of anxious & depressive episodes of highly anxious patients (deVries et al., 1987).

Discussion — The most important criticism of beep-level analysis, particularly as employed in the second part of Example 3, is that it has an inflated n. The significance tests used assumed that all data points were independent, but in this case they were not, since the 2734 observations came from only 75 individuals. If signals occur close together in time, there may also be some interrelationship between adjacent reports (see: Larson et al., 1980; Larson, 1987). Nonetheless, in certain circumstances violating this assumption is almost unavoidable and may represent the best possible presentation of the data. In such circumstances it is understood that the significance figures do not represent a strict Bayesian test of probability, but rather provide a loose and probably inflated index of confidence. In some instances investigators in this situation will raise their p level to 0.01 to reduce the chance of Type I errors.

A related problem with beep-level analyses is that they give equal weight to all instances of a category of experience, thus those individuals who report that category more often will have more influence on the outcome. People who fill out more ESM sheets (who are in some cases the psychologically healthier subjects; Larson, 1989b) will also carry more weight. It is for this reason that z-scores are preferred for these analysis, because one does not want some small group of people with high or low

means on a variable to exert an overwhelming influence (resulting in “Perspective Fallacy”). Using z-scores, however, can create additional problems. For example, if you are examining anxiety episodes, as in Example 4, those people who are anxious most often will contribute the most instances to the pool and thus carry more weight. In some instances this may be conceptually acceptable— we are interested in high anxiety people after all — in other instances it may be unavoidable. It would surely not be the first time that a significant group result is caused by a small proportion of subjects. At the least, one should check how many people have provided instances of the focal category, how many each has provided, and whether the number of instances is associated with the dependent variable of interest. Whenever possible, however, we recommend subject level analysis.

Subject-level analyses

In many cases the problems of inflated n, unequal weighting, and the danger of a Perspective Fallacy can be diminished by employing a subject-level analysis. This involves computing appropriate aggregate scores for each individual and analyzing these scores using the person as the unit of analysis⁴⁷.

EXAMPLE 5: Larson (1983) used this approach in evaluating whether adolescents’ psychologic states with friends are different than their states when they are with family members. Nearly all of the 75 subjects in his study reported at least one occasion of being with friends and being with family members. Thus, for each subject he computed a mean for each variable of interest for all times the subject was with each type of companion using the raw data⁴⁸. Matched pair t-tests between these means revealed that when adolescents were with their friends they reported higher average ratings of affect, freedom, and openness than with their families (Table 7.2). Larson & Johnson (1985) used a similar approach in testing the moods of bulimics at home vs. at work vs. in public and deVries et al. (1987) did the same (using z-scores) for schizophrenics and anxiety subjects.

Items	$\bar{X}$	$\bar{X}$	With family	With friends	t (n = 73)
Self-Rating (1 to 7)	Clear (vs. confused)		4.65	4.88	2.06 *
	Open (vs. closed)		4.58	5.13	4.65 ***
	Free (vs. constrained)		4.70	5.14	3.87 ***
Ratings of the situation (1 to 6)	Goals same (vs. different)		3.89	4.50	3.25 **
	Feedback positive (vs. negative)		4.42	4.94	3.57 ***
	Talk jokingly (vs. seriously)		3.34	4.04	5.35 ***
Mood (-12 to +12)	Total affect score		7.75	9.64	5.57 ***

* p<0.05; ** p<0.01; *** p<0.001 (two tailed test of significance)

Table 7.2 Perceived interaction when with family or friends (Larson, 1983).

EXAMPLE 6: In some cases an investigator develops a single coefficient that tests a specific situational hypothesis within each subject. A significance test is then performed to evaluate whether a significant majority of subjects confirm the hypothesis. Portegijs (1988) used this approach to evaluate whether the intensity of pain in 24 lower back pain subjects was related to time of day. He computed correlations between time of day and ESM reports of pain for each person. Using a

⁴⁷ However, see warning on aggregating data in Appendix 6.
⁴⁸ Larson’s approach in this analysis controls for individual differences in the mean (because it used a matched pair t-test), but fails to control for individual differences in variance - which would be controlled had z-scores been used. The latter approach (norming by variance using z-scores), however, is also criticizable because we can prove that: $x_a - x_b \neq z_a - z_b$. As a consequence the choice option is not arbitrary. It is up to the researcher to make this choice consciously.

t-test he then found that the mean of these 24 correlations was significantly below zero, indicating contrary to Folkard (1976), that pain diminishes across the day. In another use of this approach Larson et al. (1982) simply counted the number of people showing one trend vs. another (whether a given activity was more frequent when alone or with people) and tested these counts against chance using a sign test. We think these simple descriptors are underused.

Discussion — Subject-level significance tests are conservative tests and tread on few statistical assumptions. Because the individual is used as the unit of analysis, the assumption of independence is not violated, as it is in the beep-level analysis. Therefore, we should try whenever possible to aggregate the situational information to a one score by subject datum. This requires creativity and can cause some consternation before the database is fitted for analysis, but it is worth the trouble because person-level tests that can be used then meet basic statistical assumptions. They will be more persuasive when significance is found.

From this it must be clear that whenever possible, the subject-level analysis is preferred above the beep-level approach⁴⁹. Unfortunately, subject-level significance tests are not foolproof and different possible pitfalls can be discussed for them.

It can be argued that aggregating data to the subject level squanders the repeated measurements, increasing the probability of Type II errors in which a true relationship is not detected. Folkard (1976) showed that the increment of pain by time of day is linked to the diurnal variation in body temperature. This, however, is a correlational observation. Environmental changes with a day pattern such as “being alone”-frequencies, which are lower in the evening, may also cause these differences. The adequate analysis of environmental effects on experience must therefore necessarily be conducted on the residual data after fitting a diurnal model upon the data. The techniques for these analysis are described by Gottman (1981).

The next possible problem is related to the actual computing of the aggregated scores at the subject level and here each author takes a different stand. The second author and others in the Maastricht group commonly set a minimum limit on the number of data points required to compute a summary score for an individual (e.g. a person must be alone at least 5 times in order to compute a mean for alone time). The argument is that a mean based on 1 or 2 self-reports is not a very reliable estimate of that person's true value. The first author, however, argues that this is not important since the issue of confidence in a relationship is addressed by the significance test. He further points out that setting such a minimum may exclude many individuals and lead to a the test being conducted on a biased sub-sample (e.g. people who are alone more often). In part this difference in opinion reflects our respective research backgrounds. The first author, primarily identified with developmental and social-psychological research, is concerned with revealing trends that characterize broad populations, while the second author, primarily a clinical researcher, must be much more attentive to individual variations that may be of clinical relevance.

This bone of contention identifies the major restriction on use of subject-level analyses — they require that a substantial number of participants have one or preferably more moments where they report the target categories of experience. In an ana-

⁴⁹ cfr. comment on MANOVA's.

lysis of relatively rare behaviors (e.g. using drugs, riding in a car), this may not be the case and beep-level analysis may be the best you can do. This restriction becomes particularly severe when you are interested in making comparisons across three or more situations using a MANOVA model. These statistical techniques do not tolerate missing data, yet it is likely that few individuals will have enough cases in each situation to compute a mean⁵⁰. Again, in such circumstances, unless one can invent a plausible algorithm for substituting missing data, one may have to settle for a beep-level analysis, with accompanying exploratory analyses to be certain that differences found are not an artifact of a person-level variable.

It might be noted that subject-level analyses dealing with percentages also run into severe problems if the category of interest is relatively rare (for example if you are interested in the amount of time spent talking to a parent vs. a spouse). Because the great majority of people will have no instances of these categories, the researcher will be left with skewed distributions in which nearly everyone has observed frequencies in those situations of 0%. Making things even worse, demographic correlating variables may cause these differences.

This brings us to our final problem, one that applies to all analysis of situation-level questions (and to all ecological research!): the possibility that an apparent relationship is actually the result of some third variable. In particular, it would be embarrassing and unconscionable if that third variable were something that was measured in the study. Hence, in Example 5, one might wonder whether the higher moods reported with friends might really be due to the fact that the subjects play sports more often when with them, which in Example 3 was reported to be a high affect activity. Or perhaps it is due to the time of day or the environments in which interactions with friends take place. Unfortunately, it is generally not feasible to create an analysis of variance that evaluates all of these independent variables simultaneously (because the time samples are not randomly distributed across all combinations of situational variables and the ANOVA is likely to be defeated by many poorly filled and empty cells).

The best one can usually do when confronted with such a complex database is to design alternative analyses and examine competing explanations for the finding on an ad hoc basis. For example, one could compare states with family members on an activity by activity basis, at least for those activities that are done and reported both with family and friends. Csikszentmihalyi & Larson (1984) found that such analyses supported the integrity of the original finding. This is the point at which creative worry and discussion with colleagues is useful and may lead you to a deeper and more complex understanding of the true relationships. It was only after much thought that the second author and colleagues realized that their counterintuitive finding — that agoraphobics feel best outside their homes— may be due to the fact that these patients only leave home when they feel good. To decide how, finally, to present your data, you may end up trying several different approaches, beep-level and person-level analysis, using z-scores and raw scores, splitting up the database to conduct independent

⁵⁰ The authors have typically worked with data sets with 30 to 50 time samples per person. Many of the difficulties identified here with subject level analyses of situations would be much less acute were it possible to obtain more data points per person.

analysis and attempting to control for various variables. Losing significance in seemingly parallel set ups should always be considered an important signal. The decision rests on which approach most satisfactorily presents the underlying pattern in the data. In some cases one may find it necessary to present the results of two or more different approaches to express this pattern.

Analyzing questions about persons

Questions about persons can also be analyzed using beep- or person-level data. Again we will show that the person level approach is preferable and provides a more persuasive presentation of the data, but there may be occasions when beep-level analysis is the method of choice, especially whenever low frequency events or situations are the focus.

Beep-level analysis

EXAMPLE 7: deVries et al. (1986) compared the psychologic states of chronic psychiatric patients and normal controls in relation to the number of persons in the environment at the moment of the beep. They used the z-score Beeper File in order to control for between subject differences in overall means and standard deviations. Analyses showed that the schizophrenics reported the most positive states when in small company (two persons present), while for normals there was a progressive improvement in state with additional companions up to a maximum of four, with stabilization above this.

deVries & Delespaul (1988) replicated the study at the subject level but encountered difficulties computing mean scores for larger social situations because of the restricted environment ambulatory chronic mental patients live in. The differentiation of companionship at the moment of the beep was almost impossible at a level higher than three people, because of the small number of observations in each cell.

EXAMPLE 8: In a special journal issue Larson & Richards (1989) were interested in describing the changing time use patterns associated with the transition years from childhood into adolescence. Their primary interest was in comparing the percentage of time pre-adolescents and adolescents spend in a variety of contexts. A large sample size ($n=483$ people, 18,022 time samples) allowed them to look at a wide range of contexts, including contexts that occurred quite rarely, for example, watching different categories of TV shows. The rarity of these contexts, however, made subject-level analyses difficult for reasons discussed above. While some of the analyses might have been performed at a person level, the investigators decided for consistency to do all analyses at the beeper level. Figure 7.2 shows an analysis from an article by Raffaelli & Duckett (1989) that shows a dramatic increase in the amount of time girls spend talking over this age period.

Figure 7.2 Age differences in "whom young adolescents talk with" (Raffaelli & Duckett, 1989).

Discussion — All of the limitations of beep level analyses discussed in section 1.i. apply here as well, so this discussion is primarily a recapitulation of the previous one. N's are inflated, thus statistical tests do not provide a precise representation of probabilities. There is a danger that certain subgroups of individuals may exert unequal weight in determining specific values. Nonetheless, particularly in instances when one is concerned with relatively rarely occurring categories of experience, this approach may be the only possible. In this situations the data must be analyzed with appropriate extensive checks for possible confounds.

Subject-level analysis

EXAMPLE 9. The most common type of analysis at this level is the comparison of groups in terms of average subjective states, variance in states, or the percentage of time spent in a specific context. Delepaul & deVries (1987) computed the mean mood states for each subject in a sample of chronic psychiatric patients and a sample of normal controls. They then computed a simple t-test between the sets of means and found that the psychiatric subjects differed significantly from the normal controls in their average subjective state (Table 7.3).

Item group	Schizophrenia (n=9)	Normals (n=11)	t-statistic
Thoughts	4.33 (0.85)	5.14 (0.70)	$t_{(18)} = -3.10$ **
Mood	4.47 (0.95)	6.27 (0.37)	$t_{(18)} = -5.76$ ***

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 7.3 Between group statistics (Delepaul & deVries, 1987).

This procedure was then repeated for the difference of mean z-scores across a set of dichotomized situations (alone/not alone and home/not home). Difference scores were not significantly higher in schizophrenics than in normals but the absolute value of this difference was. The authors concluded that patients are more reactive than normals to different aspects of the environment (Table 7.4).

Johnson & Larson (1982) provide a similar comparison of the average and the standard deviations of psychological states experienced by bulimics versus normals.

Item group	Schizophrenia (n=9)	Normals (n=11)	t-statistic
Thoughts	0.46 (0.13)	0.30 (0.24)	$t_{(18)} = 1.74$ *
Mood	0.53 (0.44)	0.08 (0.07)	$t_{(18)} = 3.31$ **

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 7.4 Difference in alone/not alone mental state (Delepaul & deVries, 1987).

One criticism sometimes expressed of these types of comparisons for subjective ratings is that they merely measure differences in response sets — in how different groups use the items. In this case observed differences are considered habitual narratives of individuals. Larson & Lampman-Petratis (1989) circumvented this criticism by having participants use the ESM scales to rate emotions in a standard set of faces and were able to demonstrate that the group differences they found in the ESM reports were not replicated in the ratings of faces, indicating that response set could not have caused the differences. The response set issue can also be avoided when one is comparing people in terms of their states in a specific context. By evaluating people's mean z-scores in a context (instead of their mean raw scores), one can determine whether states in that context differ between two groups — relative to the subjects respective normative states. In some cases, however, use of mean raw scores is preferable (for example Larson & Asmussen, 1992).

EXAMPLE 10: In a very creative use of ESM data, Graef, Gianinno & Csikszentmihalyi (1981) used each person's reports on their daily activities to estimate how much energy each person consumed in their daily lives. Activities such as daydreaming and socializing use little gas and electricity, according to data from other investigators; media use and shopping consume more. Their curious finding — one provocative from a philosophic standpoint — is that there was an inverse correlation between this estimate of energy use by women and their average happiness on the ESM sheets. Women who used more energy were less happy.

Discussion — Clearly these subject level analysis are powerful and convincing analyses in answering questions about persons. Nonetheless, there are potential problems. As we have said before, care should be taken to avoid confounds, and we propose never to carry it out without concurrent testing of competing hypotheses.

Confounding variables are a constant threat to all kinds of studies, even those using control or contrast groups. In studies of the influence of context on psychopathology we are faced with the fact that mental illness is always related with a set of co-occurring life style differences. For example, a higher proportion of sick people are unemployed and unmarried; unemployment has direct repercussion on the time spent at home, and being married influences the alone time significantly; both in turn are likely to be related with average moods. While one can not separate these variables using the classical experimental approach of random assignment, attempts to decompose these relations are essential to understanding differences in how groups spend their time and how they feel in different contexts. As was illustrated in Example 2, and in the paper by Larson & Asmussen (1992), an overall pattern of time use or mean moods may mask important situational differences or similarities. The only way to deal with these possible threats is to challenge one's conclusions constantly — think about them in the shower, present them to colleagues, discuss them with your research subjects.

The second limitation is related to the issue of response set discussed under Example 9 above. Data derived from the anxiety research (Dijkman & deVries, 1987) suggests the way the subjects perceive themselves and the surrounding world may influence their responses in clinical interviews and on one-time questionnaires. In these data-collecting situations subjects seem to present their individual "stories". We often get other responses when the same questions are asked repeatedly and linked to a real natural situation as in ESM. Another process seems to regulate the formulation by the subjects of ESM answers although it is difficult to understand how this should work, considering the fact that ESM assessment form often mirrors traditional questionnaires. No control situation exists to test the validity of experiential assessments in natural environments. Response sets can partly be ruled out by methods as described in example 9. Other, indirect approaches have to be used but most of them are handicapped by the fundamental difference between moment-to-moment experience and retrospection. In this situation the creative design of alternative evaluating strategies and the confluence of results is the method of choice to circumvent possible pitfalls.

INTERPRETING RESULTS

Good questions and optimal analysis are easily destroyed by poor discussions. A straightforward explanation for the finding in Example 10 is that less happy people use more energy, perhaps as an attempt to compensate for a lack of personal energy. It is also possible, however, that people who use a lot of energy may have more free time than others, perhaps as a result of lower employment and this may influence the findings. Hence jumping to the conclusion that energy use does not serve people's happiness may blur the real relationship, an error which these authors carefully avoid. The discussion part of papers presenting results derived from ecological datasets is usually the moment for free speculation. That is what discussions are for, you may say, but this does not mean that the above mentioned possible pitfalls should not be considered carefully. As we have stated repeatedly, ecologic datasets can be described, sometimes interpreted but almost never fully understood. We advise some humbleness in summarizing the data and caution in the projection of the utility. We advise authors to carry out an open discussion of rivaling interpretations, the description of discussions with colleagues, as well as the projection of future research lines with better controls of possible intervening variables.

FINAL COMMENTS

After hearing about all the possible confounding factors in the ESM analysis, you may easily be seduced into thinking that more advanced statistics might be better for these data — that some complex multivariate design will allow you to control for the whole world of confounding variables at once and test person and situation effects simultaneously. We would like to warn against this fantasy, because even skilled statisticians will probably fall back in lethargy if they have to conclude that most of the premises of the tests they want to use are not met. We advise investigators to start with basic descriptive statistical techniques, because their robustness is well known and the findings are closer to the original data and more easily understood. Most ESM questions can be translated into a set of alternative operationalizations, sometimes leading to divergent results. Our advise is to select multiple approaches when possible and look for results that converge upon the same conclusion.

Nothing can replace the basic process, supported with elementary statistics, of looking at the data and getting a feel for it. Writing out and submitting case studies should also not be neglected. After doing these simpler steps it may be possible to explore the new horizons of the sophisticated statistical techniques. Remember, though, that statistical procedures, after all, are a means not an end.

Part 3: Studying Schizophrenic Symptoms and Daily Life Adaptation

Chapter 8: Context and subjective experiences in schizophrenia ⁵¹

What clinician has not experienced with perplexity the chance meeting of a patient in a social or non-clinical setting when he or she appears remarkably different, more distressed or capable than the therapist expected from the clinical encounter? This experience may relieve or shock the therapists. We are likely to be philosophic about the encounter or muse over the problems of diagnoses and the perplexing nature of man. We may also mention it to the patient, depending on our therapeutic proclivity, but more systematic inquiries into the questions raised by the discontinuity of clinical perception are not considered. This situation is an understandable outcome of the assessment methods that we generally employ in psychiatry. Most often we rely on the clinical interview or at best a few observations to provide us with data about patients' lives. These standard procedures, however, simply do not provide sufficient access to the context of patients' lives.

Assessment strategies that adequately describe the person in context have proved difficult to realize scientifically and, therefore, research is scarce despite the fact that the interaction between the human organism and the environment is a commonplace idea in the behavioral sciences. Some fields in biomedicine and the social sciences — for example, neuroendocrinology and sleep research (Minors & Waterhouse, 1981; Kripke, 1983) — have advanced significantly since the advent of assessment methods which can more appropriately measure the actual time course and changes in their subject matter. Psychiatry, however, has not kept pace in developing parallel methods. By simply adhering to traditional psychiatric assessment methods, researchers may overlook important clues to the very phenomena that characterize certain diagnostic groups and psychological phenomena.

As Carpenter (1986) argued:

“It is only in the context of time, interpersonal sensitivity, and persistent inquiry that a comprehensive view of the nature of psychopathology and its impact upon the person and his surroundings can be appreciated.” [p. 23]

The most relevant clinical data provide an understanding of what it is like for the patient to be in “different” mental states, and make it possible to place this information within the context of the patient's normal life. Clinicians should be interested in knowing when a particular mental state was experienced: where and with whom the patient was at that moment, what he or she was doing and whether it was difficult or easy, what happened just before and what was expected to happen next, whether the patient was taking medication or under the influence of alcohol, and so on. All these data are necessary to sketch the mental state of a patient in context. A desirable strategy for psychiatric research would then be to provide a more precise description

⁵¹ This article was originally published as: deVries, M. W. & Delespaul, Ph. A. E. G. (1989). Time, Context and Subjective Experience in Schizophrenia. *Schizophrenia Bulletin*, 15(2), 233-244.

of subjectively experienced symptomatology, health, and variability in natural settings as has been urged by pioneering researchers in the field (Bannister, 1977; McGuire & Fairbanks, 1977; Wynne et al., 1978; Bellak, 1979).

Until recently, this kind of research was difficult to implement in psychiatry. With the advent of computer technology and advanced statistical methods that allow the handling of large sets of related data, time-allocation and time-sampling methods are moving to the forefront of behavioral science methodology (Gross, 1984; Pervin, 1985). This article presents data collected with such a method. The Experience Sampling Method (ESM) (deVries, 1987; Csikszentmihalyi & Larson, 1987) allows us to inquire into the dynamic fluctuations of pathology and experience in relation to the social and temporal changes in daily life. ESM is a data-collecting approach that assesses a representative sample of mental state and context gathered in the daily, natural environment of the subject. We hope it will, in the long run, lead to a contextualized "systematic phenomenology" of mental disorder.

VARIABILITY IN SCHIZOPHRENIC SYMPTOMATOLOGY

The issue of mental state variability is central to the understanding of psychiatric pathology in general and schizophrenia in particular. Historically, psychiatry has been interested in these phenomena. Kahlbaum (1863), Kraepelin (1896) and Bleuler (1911), almost a century ago, richly described the symptoms and course of schizophrenia. Since then, theorizing and studies of schizophrenia have been rich and many, but currently researchers tend to avoid phenomenologic descriptions. Only a small group has investigated the actual course, changes, stages and fluctuations in schizophrenia (Strauss et al. 1985). Even less is known about fluctuations of symptoms, mood, thought and behavior in real-life situations, although these issues are the bread and butter concerns of clinical psychiatry.

Currently biologic formulations are receiving much attention in schizophrenia research. Most studies emerging from prominent institutes use new diagnostic techniques such as MRI, PET scans, and CBF that promise to bring to fruition the search for functional and anatomical abnormalities in schizophrenia. Epidemiologic researchers ponder issues such as cross-national prevalence differences and unequally distributed birth rate frequencies over the months of the year, of persons developing schizophrenia later in life. Psychologic scholars focus on intellectual and cognitive factors such as loss of abstracting ability, perception and attention difficulties, and problems with faulty associations and thinking, all of which are tied to slow learning and adaptive liability (Rabin et al., 1979). These studies have enhanced our knowledge of schizophrenic symptomatology. Unfortunately, the interactive, social and behavioral aspects of the disorder that have been illuminated over the last 30 years have not been researched as intensively. Therefore, new studies in all fields of schizophrenia research need to be augmented as far as possible by phenomenologic and behavioral descriptions of schizophrenic symptoms over time.

The schizophrenia syndrome is today generally considered a disorder marked by a vulnerability to time-limited, acute illness episodes that occur throughout the life-

span. Studies have shown that specific factors leading to exacerbations of the illness do not differ from precipitating factors in other illnesses, but that persons with schizophrenia seem to react to these influences more severely and maladaptively. Although the meaning of the concept is not well worked out, a general agreement has evolved that patients with schizophrenia are environmentally vulnerable. Patients with schizophrenia are postulated to be less able than normal individuals to maintain homeostatic control in the face of environmental disruptions. This has led to successful treatment strategies that have stressed the management of environmental changes, particularly in the family (Leff & Vaughn, 1985).

Most of the studies on changes in mental state of subjects with schizophrenia have focused on the issue of expected outcome over time. In particular, the long term follow-up studies of Bleuler (1978) and Ciompi & Müller (1976) are instructive. Their work can be summarized as follows: 15-25% of subjects developing schizophrenia at some point in their life will be "cured" some years later and remain symptom-free for the rest of their lives; 30-40% will follow an episodic course leading to repeated rehospitalization but also potentially long symptom-free periods during which they have no residual handicaps; 10% will stay chronically incapacitated for the rest of their lives.

Unfortunately, we don't know how we can change these patterns of outcome. Most factors associated with poorer outcomes are demographic, such as never having been married, month of birth and place of birth, and it is clear they are not easily changed. Pathology aspects such as acute versus chronic onset and subtyping are also related to outcome. Finally, Cloninger et al. (1985) have suggested that maximum predictability of poor outcome can be reached using an index composed of symptoms such as delusions of persecution and control, specific mood-incongruent delusions and auditory hallucinations. These outcome related characteristics all deserve much attention but fall short in their usefulness for designing therapeutic interventions.

Until now, studies of the course of schizophrenia have been less common. Except for the staging of the onset of psychosis (Chapman & Chapman, 1980; Docherty, 1978) we have not learned much. Bleuler (1978) differentiated a linear course from an episodic one. Strauss et al. (1985) described 8 specific principles of longitudinal course during the first year following a psychotic episode. Strauss stresses key components such as the nonlinearity of course, identifiable phases such as "moratoria", "changing points", "ceiling effects" and "getting a foothold", the time-decay of vulnerability and phases of environmental responsiveness to the patient. This description of course over an entire year is a step forward in clinical understanding but more is needed, for it is still not enough to describe the individual's long-term relationship to contextual and temporal factors while not understanding their subjective illness experience.

Can we indeed find an increased vulnerability, evidenced by cognitive and mood disturbance in certain settings, that differentiate patients with schizophrenia from non-psychiatric populations? Setting effects have not been well-studied, despite the strong therapeutic claims made by the family and expressed emotion researchers (Leff & Vaughn, 1985). Supportive evidence for setting effects have come from behavioral observation. Reynolds (1965), while seeking periodicity in schizophrenic symptoms with behavioral observation techniques, noted instead considerable contextual influences on maladaptive behavior. Setting-dependent symptom reactivity has also been

reported in observational studies of schizophrenia (McGuire & Fairbanks, 1977; McGuire & Polsky, 1979). These authors suggest that persons with schizophrenia are less able to stabilize their behavior after a stimulus change and therefore demonstrate greater thought disorder in the presence of increased affect and interpersonal stimulation. This often results in a rigid and withdrawn defense strategy that is the hallmark of the illness. In all these studies, variations between patients in reaction patterns and sensitivity were also noted. Similarly, deVries (1983) observed in a pilot study of paranoid and undifferentiated subjects with schizophrenia that the undifferentiated group was more sensitive to social stimuli.

Subjective self-reports of symptom fluctuations as they occur in relation to environmental influences can be illuminated by the repeated non-retrospective self-reports of experience that we employ. Since settings have been shown in retrospective, cross-sectional studies to differ in their pathogenicity, we would expect to find this, too, in prospective studies. So we expect, as Leff & Vaughn (1985) did, that some settings increase pathologic responsiveness, while other settings do not.

STUDY

Purpose of the study

In this paper we will present a study demonstrating that it is possible and interesting to set up research on repeated non-retrospective self-reports of experience, even with severely disordered mental patient populations. The core issue is the description of variability and context relatedness of symptoms over time. The different subtopics studied were: 1) the mental state description of the subjects; 2) the use of time and context; and 3) the effects of context on mental state.

Subjects

The study was conducted on a group of 9 severely disordered persons. All psychiatric patients were in active treatment in health care facilities during participation in the study. All were diagnosed as having schizophrenia by DSM III criteria and had a 7 "plus" score on the Strauss & Carpenter (1981) 12 symptom list, a cross-cultural research criterion used in the International Pilot Study of Schizophrenia. The patients were drawn from extremely ill ambulatory and hospital populations who had illness careers exceeding 10 years; all were under 40 years old and on medication. In addition, 7 non-psychiatric subjects (academic-level professionals) were sampled. Cross-sectional measures were gathered before the repeated measurement period.

The Experience Sampling Method

The Experience Sampling Method (ESM) is a data collecting method using random time sampling procedures to gather repeatedly and non-retrospectively self reports of experience (deVries, 1987; Csikszentmihalyi & Larson, 1987). It borrows from a va-

riety of time sampling approaches that date at least from the ethnographies of Malinowski (1935) to the large time budget studies of today (Robinson, 1977; Szalai et al., 1972). Further contributions have been made by the more systematic application of time observation techniques in naturalistic field studies (Barker, 1968, 1978; Chapple, 1970; Munroe & Munroe, 1971a, b), by ethologic research strategies applied in psychiatry (McGuire & Polsky, 1979; Reynolds, 1965), by "spot observations" in ethnographic fieldwork (Whiting et al., 1977) and by the behavioral monitoring techniques (Nelson, 1977). These approaches have described the behaviors of subjects and patients in great detail, but not the experience of individuals. ESM, however, is intended to catch the experiential dimension of mental state variability.

The ESM is an assessment technique designed to obtain repeated self-reports, using an electronic beeper that signals subjects to fill in small questionnaires at preselected but randomized time points each day over a number of days. The subjects carry a beeper and a booklet containing the report forms into the places and during the activities of his normal day and makes the required assessments within this context. For method, application, validity and reliability, see the appropriate chapters in this volume.

ESM-procedure — In the ESM application used in the study reported here, subjects were signaled 10 times a day (once in each 90-minute timeblock between 7:30 and 22:30) for 6 consecutive days. Randomizing the occurrence of beeps throughout the day minimizes subject reactivity, assures contextual and concurrent validity and presents a clear picture of situational and diurnal variations in experience.

Subjects were seen by the researchers for a briefing session in which the diagnostic assessments were done and the sampling procedure as well as technical details concerning the beeper, were explained. This face-to-face contact is crucial in order to develop a good research alliance. After the sixth sampling day we met again with the subject for a debriefing session. Then we checked the missing data with him as well as the answers that are expected to lead to coding difficulties.

ESF: questions, variables and coding — The repeated self-report questionnaire is based on the mental state exam in psychiatry and called an Experience Sampling form (ESF). The ESF is constructed using closed Likert-type scales and open questions.

- *CLOSED LIKERT-TYPE SCALES* — The Likert-type scales are set up using a modular format, with subsets requesting responses on scales referring to "thoughts", "mood" and "physical concerns" (Table 8.1). Data of individual scales belonging to a "module" are summarized to a "module score" by averaging Likert scales in such way that a high score is always a positive emotion or affect. Factor analytical studies confirm the internal consistence of these different modules. The thought assessments follow dimensions described by Hurlburt (1980), while mood can be traced back to the classic studies by Izard (1972). A specific deviance and psychopathology-module was developed for schizophrenic symptomatology.

Standard cueing questions	Pathology related cueing questions
Thoughts: pleasant; clear; excited; normal	I hear voices
Moods: cheerful; secure; social; relaxed; calm; friendly	I'm suspicious
Physical concerns: hungry; tired; not feeling well	I cannot express my thoughts
	My thoughts are influenced (by others)
	I can't get rid of my thoughts
	I feel unreal

Table 8.1 ESF 7-point Likert scales (1= not at all; 7= totally true) for thought, mood and physical concern modules (left) and schizophrenic symptomatology (right).

- OPEN QUESTIONS* — The open ended questions are coded by the research team and assess the social physical context in which the above mentioned aspects of mental state take place. The coding categories are briefly described in Table 8.2. The assessment of “thoughts” (“What are you thinking?”-question) is difficult to operationalize. Two additional variables are derived. The *first* one assesses the congruency between the thought and the activity. For example, when a subject is washing dishes and thinking about a quarrel on the previous day it is assessed as “off-line”. Conversely, a subject washing dishes and asking herself where she left the towel is clearly “on-line”. The *second* variable assesses the pathologic aspects in the thought process. It forms an ordinal scale from no pathology to severe illness. This variable, as well as the previous one, refers primarily to the thought content. Some examples can elucidate this “thought-pathology”-variable. The woman washing dishes can think of the towel and will be assessed as “focused”; she can think of her children at school and then will be assessed as “daydreaming”; when expressing some disturbance referring to the children who are late from school she is “worrying”. When she thinks without cause of what could happen to the children at school she is preoccupied and when this is described as happening all the time — a rumination that can’t be gotten rid off — it is assessed as “circular thoughts”. Considering a situation where she describes her hands dissolving in the water she is probably hallucinating and will be scored “derealized”. Coding this variable usually requires additional information on the habits of the subjects and often further questioning in a debriefing interview (Delespaul & deVries, 1987).

Cueing question	Coding information
Where are you?	• home; relatives; network; health care facilities; public places
What are you doing?	• nothing; self-care; household; work; leisure; travel
Who are you with?	• alone; with family; friends; colleagues; strangers
What are you thinking?	• congruency of thought and activity (“on-line”) • pathology (“goofs”): focused thinking; daydreaming; worrying; preoccupation; circular thoughts; psychotic/derealized • content

Table 8.2 Open-ended questions with coding information.

Results

Mental state — The problem operationalized in this part is: do severely disturbed psychiatric patients describe themselves differently, in comparison to the normal controls, on selected experiential dimensions?

Based on the Likert scale self-assessments, only the module assessing physically related experiential aspects separated both groups significantly. The other ones were in

the expected direction but failed to reach significance, probably because of the small sample size. The interpretation of the only significant module is difficult because of the broad range of aspects covered by the scales (hunger - tiredness - not feeling well). After the removal of "tiredness" the difference was no longer significant. A separate treatment of individual Likert scales would be more appropriate in the future. As we expected, the schizophrenic subjects were rated less "on-line", indicating less congruency between their thoughts and activities. Normals showed less psychopathologic thought content on an assessment of thought quality, the "goofs"-variable. Only the "on-line" variable significantly differentiated the two groups (Table 8.3).

Item		Schizophrenia (n=9)	Normal (n=7)	t-statistic
<i>Thoughts</i>	7-point Likert Module	4.73 (0.65)	5.35 (0.68)	$t_{(12)}=-1.73$ n.s.
	"on-line"	0.52 (0.18)	0.74 (0.15)	$t_{(14)}=-2.62$ **
	"goofs"	1.75 (0.83)	1.13 (0.15)	$t_{(12)}=-1.82$ n.s.
<i>Mood</i>	7-point Likert Module	4.66 (0.37)	5.12 (0.67)	$t_{(11)}=-1.57$ n.s.
<i>Somatic Concerns</i>	7-point Likert Module	3.43 (1.72)	5.96 (0.91)	$t_{(10)}=-3.18$ *
	Likert excluding "tiredness"	5.53 (0.75)	5.96 (0.91)	$t_{(10)}=-0.90$ n.s.

* $p<0.05$; ** $p<0.01$; *** $p<0.001$

Table 8.3 Between-group statistics.

Use of time and context — The information gathered concerning the social environment at the moment of the beep was coded and collapsed into the categories "alone"/"not alone". As expected, patients with schizophrenia spent more time alone (41 %), as compared with 31 % for normals ($t_{(14)}=0.65$, n.s.). This last figure matches the "alone" time in other studies of normal populations (Csikszentmihalyi & Larson, 1984). Differences between in- and outpatients should be noted. The clinical population, although small, is almost never alone, while the ambulatory-treated patients with schizophrenia were even more often alone than their normal counterparts ($t_{(7)}=1.60$, n.s.). This low "alone"-frequency for clinically treated schizophrenics is a consequence of the living situation on the crowded hospital wards where being alone and consequently to enjoy some privacy is almost impossible. Contrary to their clinical colleagues, ambulatory treated schizophrenics are alone very often, even more than normals. The meaning of this difference in frequency of alone time should be checked further in relation to the mental states occurring in this situation to determine its clinical significance.

Setting effects and experience — The influence of context on mental state was assessed using the "number of persons present at the beep"-aspect of context and a global mental state assessment composed of the combined thought and mood scales. We hereby hoped to determine "the context of disability" (Figure 8.1). Both schizophrenics and normals had relatively "depressed" mental states when alone, reaching an optimal level in a 1 to 3 persons company range. The groups differ when a larger number of persons is present. Then we see a clear drop in mental state self-evaluation for subjects with schizophrenia but not for normals.

Many interpretations are possible, but the data are primarily of interest because it seems that a crowded setting is extremely troublesome for these schizophrenic patients. Crowdedness is an aspect of the setting where we most often treat them, the acute hospital unit. Ambulatory treatment without alternatives for social contacts may

also lead to troublesome situations because the frequency of “being alone” is extremely high in this group and this also gives relatively depressed mental states. This information, if replicable in a larger standardized sample, could provide us with important data for optimizing living-environments for subjects with schizophrenia. We plan to study specific social activities and contexts to further clarify this relationship.

Figure 8.1 Thought/Mood combined score related to the number of persons present.

CASE PRESENTATIONS

Other contextual aspects were explored on a casuistic basis to demonstrate more complex situational relationships with mood and thoughts. They warn us of the pitfalls of interpretations based on group-level aggregated data.

Case 1

A woman in her mid-forties, formally educated as bacteriologist, who had a ten year history of paranoid schizophrenia, was sampled over three consecutive days with a mean beep frequency of 6 per day. She showed a repetitive late afternoon increase in problematic thoughts. The morning periods, during which she felt better, were spent carrying out domestic tasks. In the late afternoon, she was generally alone without specific tasks and in this situation she became preoccupied with herself physically. Her delusions had a professional component such as “bacteria crawling on her skin”. Revealing are the reports of the third day where beep signals capture a delusional period followed by the patient subsequently studying an anatomy text. On debriefing it was confirmed that her studying, functioned as an intellectual coping strategy, which served as an attempt to organize disturbing delusions.

To substantiate and confirm adaptive patterns of this sort, a longer sampling period with a higher beep frequency is required (10 times for 6 days). In this way we could collect more observations in the situation of interest and enhance the power of the test of the “environment/mental state”-covariation hypothesis.

Case 2

The complexity of the relationship between environmental context and fluctuations in experienced mental state is also illustrated in the following case:

A 53-year-old married woman was seen in ambulatory care for an endogenous depression with psychotic features. The graph (Figure 8.2) presents her thought scores (mood data were similar: $r=0.56$) for the "alone"/"not alone" conditions in a within day format computed by collapsing the six sampling days to a "mean" day. The within day data can be divided into three time-blocks. During the first phase, until 17:00 the most positive thoughts and feelings occur when alone. In the second phase (17:00 to 20:00 hours) social contact clearly enhances feelings, although the alone condition also shows a slight positive trend. After 20:00 a drop in experienced mental state in the "not alone" situation can be observed.

Figure 8.2 Within day pattern of thought self-assessment in "alone" versus "not alone" situations (case 2).

Clearly the simple contextual association with the number of people, as found in the first study, can not account for the complex interactions we observe in this idiographic analysis when "time of the day" is introduced as an additional variable.

Case 3

ESM data may simply be presented by graphing a person's coded activities and plotting the frequency of severity of symptoms over the time period assessed. These plots of troublesome thoughts and mood in relation to social activities and context in turn become a powerful clinical tool for the therapist and the patient alike. For the therapist, it actualizes the patient's life and generally produces empathy; for the patient, it enhances positive transference, the feeling of being cared for and often provides useful information. An example suggests how these data may function clinically.

Alfred is a 28-year-old man, first diagnosed at the age of 22 with undifferentiated schizophrenia. His other DSM-III diagnoses were passive dependent personality disorder and hypertension. He had been employed as a dishwasher and a store clerk, and had lived alone or with his parents in a relatively unstressful environment over the last 8 years. Since he was diagnosed the patient was treated steadily on 20-30 mgr. of Stelazine, which allowed him to function socially at a marginal level. The patient was referred for two reasons: the exacerbation of his schizophrenia as well as a hypertensive crisis which had required intensive medical intervention. On ESM assessment it was found that, rather than being present throughout the day, the patient experienced anxiety and increased blood pressure primarily at 11:00 each day while at work. At this time, as a dishwasher, he had to sort silverware - spoons from forks - on a automated conveyor belt system during the pre-lunch rush. In this anxious state his blood pressure increased markedly, often to 180/120. In the clinical setting, the agitated patient with rampant high blood pressure appeared puzzling. The time assessment, however, linked the problem to a specific environment that the patient experienced as socially and cognitively distressful and pointed to a clinical intervention. The circumstances were explained to his employer with the suggestion to adapt the patient's work

tasks. The employer cooperated and over the next 3 weeks the patient's blood pressure returned to near normal, a quite remarkable change since his extremely high blood pressure over the previous 2 months had led to intensive treatment with Aldomet, Propranolol and Valium in addition to the Stelazine. After 3 weeks, the patient was again assessed with the ESM method and an increase in the patient's agitation between 10:00 to 11:00 was again noted. At this time, the lunch-work shift, made up primarily of female workers, came to prepare the noon meal. This brought up a host of fearful, sexual and aggressive feelings in the patient which he experienced as psychologically disorganizing. This situation uncovered new psychotherapeutically useful information about the nature of the patient's relationship with women and particularly his mother, relationships which he had previously been unable to discuss. A focused psychotherapy approach was thus possible and resulted in a relatively symptom-free year.

While this information could have been derived from careful clinical interviewing, other medical workers had missed the importance of the environmental influence. The ESM assessment method brought it to the fore. In this case, it proved doubly powerful since it provided a chance for both a medical and a social intervention as well as guiding a psychotherapeutic approach. It is important to remember that the patient during interview had been unable to communicate the linkage of these events. We think, the case of Alfred is a striking example of the clinical usefulness of ESM.

DISCUSSION

In this study we have demonstrated the promising usefulness of repeated self-report methodology in natural environments, for the elucidation of questions related to the experience or mental state of chronic patients with schizophrenia in the context of changing social aspects in their daily life.

As a group, our patients felt best in a small social environment, characterized by the presence of 1 to 3 persons. Being in the company of more people, as well as "being alone", was more troublesome. These results, if replicated in larger studies, have direct implications for the organization of care taking facilities for the chronically mentally ill. As far as these data are considered, the crowded hospital ward is not the optimal choice for a treatment environment, while ambulatory treatment can lead to a dramatic rise of "alone" time, which is equally disturbing. Optimized care-taking facilities should be organized to avoid these "contexts of disability" and create opportunities to let the patients live and be treated in small scale environments.

These results are not unexpected. Nevertheless, the cases presented show us that optimal environments are more differentiated and complex at the level of individual patients. Furthermore, even the global conclusions at the group level may be inaccurate due to the nature itself of the ESM sampling technique. Events such as being "alone" or "not alone" do not occur in a randomized way in the daily life of patients, but may be an active choice, even a coping strategy, reacting upon occurring mental state changes. This way the depressed mood in the alone situation can be a result of the active withdrawal of the patients when they are bothered more by delusions and hallucinations. The same hypothesis can be made concerning the more depressed mental state in the large social environment, where withdrawal can occur in the anonymity of the crowd.

Chapter 9: Daily Life in Chronic Mental Illness⁵²

Current psychiatric opinion holds that mental illnesses is best understood as caused by a genetically driven vulnerability. Growing evidence exists to support this. The actual emergence or onset of full-blown symptoms is, although dependent on this genetic vulnerability, largely caused by situational cues and stressors. The study of the influence of the environment on vulnerable individuals suffering schizophrenia and their emotional and symptomatologic adaptation to immediate contextual influences are the topic of this study. We are convinced that the in-depth study of the dynamic adaptational process of the individual in interaction with his environment — in illness and well-being — provides an invaluable source for understanding the fundamental processes of mental illness in general and schizophrenia in particular. These help us to elucidate the heterogeneity obscured in the current diagnostic systems. This adaptation to daily life is also the frame of reference for therapeutic successes or failures.

More than 5 years ago we published a preliminary study on the effects of time and context on the subjective experiences of chronic mental patients (deVries & Delespaul, 1989; Chapter 8). Since 1989 we have invested further in the comparative study of daily life experiences and activities of subjects representing different mental health populations: the “Basic Properties of Illness and Well-Being” research program. The present chapter updates the discussion presented in Chapter 8, by submitting data collected in a new and larger sample of chronic mental patients that includes schizophrenic as well as non-schizophrenic chronic mentally ill patients. We will assess the patients’ moment-to-moment experience and time use patterns and relate these to subject characteristics. We will also assess how sampled experience can be understood in relation to concurrent contextual cues and changes in these contextual cues over time. Our research is driven by the large number of contextual theories currently available in psychiatric literature. Whenever possible we will compare our current findings with data collected previously from a spectrum of less chronically ill psychiatric or somatic patients. The data in this study enables us to understand the reality of daily life experiential adaptation and assess the impact of subject characteristics, illness process (chronicity) and illness type (schizophrenia and depression).

METHOD

Subjects

Two hundred ESM periods of chronic mental patients were sampled in 5 years using the procedures described in Appendix 1. Thirty-three of these ESM episodes were fol-

⁵² Thanks to Don Hedeker for the stimulating discussion leading to this analysis and to Bart Leunissen for the data-crunching.

low-up samples, resulting in a set of 167 different cases. The patients are grouped in three subcategories. The first group is labeled “schizophrenic spectrum disorder” ($n=81$) and contains in addition to schizophrenic subjects (295.xx) the following DSM-IV diagnoses: 298.90 (psychosis — NOS / $n=7$), 297.10 (delusional disorder / $n=3$) and 298.80 (brief reactive psychosis / $n=4$). As we argued in Chapter 1, the DSM-IV diagnosis does not give the most optimal delineation of the illness. Thus we study symptom variability, environmental reactivity and coping of the schizophrenic spectrum disorder. In the second group ($n=29$) — labeled (major) and chronic depression (296.xx) — we included patients with dysthymia (300.40 / $n=4$) and adjustment disorders with depressed mood (309.00 / $n=1$). The third group is called “Mental Disorder Controls” ($n=23$) and is a heterogeneous group containing among other diagnoses, addiction ($n=4$), anxiety ($n=5$), conduct disorders ($n=2$) and somatization ($n=4$). All these patients have long term histories of mental institutionalization. A fourth group are the normal controls ($n=34$). Those subjects were primarily the nursing staff of the wards who participated in the study. Their age is younger than the patients.

The groups had differential drop-out rates due to the strict criteria used in ESM (see below). From the original 81 schizophrenic patients, 57 could be used in the analyses (70%). In the chronic depressive group the number of patients dropped from 29 to 23 (79%). In the remaining group of chronic mental controls, data from only 12 out of 23 subjects (52%) were valid. Finally, 4 normal control subjects did not pass the validity criteria, leaving 30 valid subjects (88%). The resulting sample has 122 cases.

Demographics — Fifty-one percent of subjects were male and their mean age was 35.7 years. There were no significant differences between the pathologic groups but the normal control subjects — primarily nursing staff — were younger ($F_{(3,116)}=8.20$; $p<0.001$). Living situation and work was a reflection of the global adaptation by the subjects and were a consequence of illness. Therefore, significant differences for those variables between the psychopathologic groups and the normal control subjects are unavoidable. One third of the subjects in all three of the psychopathologic groups are living in health care facilities. Chronic schizophrenic patients who were not in a mental health care facility live primarily with their families (44%) while the largest group of the chronic depressive and the other chronic mental health control subjects live alone (40%). No healthy control subject lives in a health care facility. The majority (66%) live with their family ($\chi^2_{(9)}=22.01$ $p<0.05$). Work is a complex variable (see Appendix 5). As could be expected, the schizophrenic subjects had the worst adaptation, while most of the normal control subjects had a paying job. Fifty-five percent of the schizophrenic patients had no day-structuring activities while this was only 4% for normal controls ($\chi^2_{(3)}=20.24$; $p<0.001$). Sixty-six percent of the schizophrenics did not leave home on a structured daily basis ($\chi^2_{(3)}=28.25$; $p<0.001$) and 72% had no income based on their activities (compared to 27% in the normal control group — $\chi^2_{(3)}=15.23$; $p<0.01$). In each comparison the other chronic mental samples took a position in the middle.

Cross-sectional psychopathology — The psychiatric diagnoses were determined by the research staff during the briefing interview, usually 1 day before the ESM sampling

week, and confirmed independently by the treating physician. Although not significantly different, the chronic schizophrenic subjects are the most severely ill — as measured by the BPRS ($F_{(2,82)}=2.68$; $p<0.08$). We computed a positive and negative symptom subscore based on the BPRS item scores (Appendix 5). As a group, the schizophrenics scored higher on both subscales (positive: $F_{(2,81)}=4.91$, $p<0.01$; negative: $F_{(2,81)}=14.95$, $p<0.0001$). Interestingly compared to the positive scale, the scores on the negative subscale were higher in our sample. This might be the consequence of chronicity. Overall, schizophrenia was associated with greater severity. Schizophrenics also had the lowest level of social adaptation — measured by the SFCL ($F_{(2,79)}=3.17$; $p<0.05$). At different places throughout this study we will describe the schizophrenic subjects and relate the data to subgroups (see the DSM-IV definitions in Appendix 5). Relevant differences based on the literature are the paranoid/non-paranoid distinction and the course. In our sample we had 67% non-paranoid and 33% paranoid schizophrenics. Considering the phase-related criteria we found: unspecified 29%; subchronic 12%; chronic 52%; subchronic and acute 5%; and chronic and acute 2%.

Experience Sampling Method

The subjects were sampled using four different, but related, ESM booklets (see Appendix 4). Only the subjects who answered more than 1/3 of the beeps (this is 20 beeps or more) within the ESM reliability window — 5 minutes before or 15 minutes after the beep time — were accepted as valid subjects. Overall, data from 73% of the subjects were valid. There was a differential subject drop-out rate based on this criterion between the selected groups ($F_{(3,163)}=3.44$, $p<0.02$). The difference was caused by a high drop-out rate in the heterogeneous residual “Mental Health Control” group (48%) and does not affect our most important samples: schizophrenia, chronic depression and normal controls ($F_{(2,141)}=2.23$, n.s.).

For the subjects included in the analysis the number of valid beeps did not differ between the groups (43.61 beeps from a total of 60: $F_{(3,121)}=2.17$, n.s.). We assessed which beeps were missed within the day and over the course of the ESM week (Appendix 3). Within each day, the highest levels of drop-out rates were found for the first and last beeps of the day, primarily because the subjects were asleep ($F_{(9,1713)}=10.76$, $p<0.0001$). There was no higher amount of missed beeps in the weekend nor for the first compared to the last days of the sampling week. Also, we did not find a difference in average mood level over the course of the sampling week. However, we did find a significant drop in variability after the first 2 days of sampling for a number of items. For the current analyses these differences have been ignored because we expect that most of the contextual conditions assessed in the following analyses are independent of the sequence number of the day in the ESM sampling week.

Analysis

The purpose of this study is to assess the activities and experiences in the daily lives of chronic mental patients. We want to understand how the subjects feel from moment to moment, which symptoms they experience during the day and how these relate to what they do, who they meet and where they go. Most importantly we want to under-

stand the influence of the subject's illness characteristics on those variables and differentiate between the effects of chronicity and specific illness type (depression and schizophrenia). Those questions are "questions about subjects" (Larson & Delespaul, 1992). For the analyses we used a conservative approach. Parameters such as mean experience and time use codes were computed by subject and evaluated as dependent variables with subject characteristics as independent variables in an analysis of variance. When our purpose was to evaluate specific influences and interactions of different subject characteristics on daily life experience and activities we used regression techniques with dummy variables (Hand & Taylor, 1987; Hardy, 1993).

For the assessment of experience in context — questions about moments (Larson & Delespaul, 1992) — the traditional analysis of variance approach (with repeated measures) is unsatisfactory because too many subjects are thrown out of the analysis due to missing data (Searle, 1987). As in the situation of "questions about subjects" regression techniques are more suited here but for such an approach the data have to be independent. Because we have more but also a different number of beeps for each subject this is hardly the case in most situations (Hedeker et al., 1994). Normal regression does not adequately accounts for the hierarchy that exists in the data: "experimental" units such that multiple observations (beeps) are grouped within upper-level units (subjects or samples) (Gibbons et al., 1988, 1993; Kenny et al., 1994). A solution to this problem is to include the upper-level unit explicitly as a variable in the analysis. Even then interactions between lower- and upper-level variables are not adequately assessed. The solution we propose here is a random regression model in which outcomes at the beep level are modeled in terms of both the beep and the subject level variables. Theoretically, clustering is possible at multiple levels and authors have generalized random regression models for this situation (Goldstein, 1987). In practice, studies with more than two levels are rare and those with more than three levels almost impossible to conceptualize. In this study the analyses were conducted with MIXREG (Hedeker, 1993), a software program that estimates random-effects regression models for clustered data. Applied to ESM data the basic matrix model is:

$$y_i = X_i \beta + 1_i \alpha_i + \varepsilon_i$$

- y_i : an $[n_i \times 1]$ vector of responses for subject i ($i: 1 \dots N$) and n_i the by subject varying amount of (repeated) responses for subject i ;
- X_i : a known $[n_i \times (p+1)]$ covariate matrix with p covariates + 1 for the intercept see β ;
- β : a $[(p+1) \times 1]$ vector of unknown population parameters (to be estimated);
- 1_i : a $[(n_i + 1) \times 1]$ vector of dummies (ones);
- α_i : an unknown subject effect normally distributed, mean 0 and variability σ^2 : $\mathcal{N}(0, \sigma^2)$;
- ε_i : a $[(n_i + 1) \times 1]$ vector of independent residuals distributed as $\mathcal{N}(0, \sigma^2 I_{n_i})$.

In this model both the y_i and X_i vectors are subscripted, allowing a different number of beeps for each subject. The covariate matrix X_i can include covariates measured either at the subject level (demographic information or cross-sectional assessments of illness and social functioning) or at the beep level (for instance, the time of the day or the concurrent context and activities). X_i can also contain, for instance, logistic or quadratic transformations of variables or dummy codes to define appropriate contrasts between the independent variables. Finally, interactions can be included that allow the

assessment of a differential impact on the dependent variable for specific combinations of the independent variables. Parameter estimations for α_i (subject level effects) are obtained using empirical Bayes methods (or “expected a posteriori”), while the intercept (β) and estimations of beep level variances (σ^2 and σ_α^2) are obtained using Maximum Marginal Likelihood techniques (Hedeker et al., 1984). The standard error of the estimation allows the computation of a z-score. This allows us to assess the significance of the estimate.

The selection of the best fitting model — maximum fit of the observed data with the lowest number of explaining variables — is an iterative process driven both by theory and statistical evidence. Variables with significant estimators that result in a (statistically) significant improvement of the overall fit of the model are retained in the step-wise process to build the final equation. We start with the inclusion of main effects (for instance, men higher than women) and include in a next phase interactions in the equations (for instance, men higher than women in schizophrenia but inversely in depression). In a sequence of nested models the improvement in fit of the more saturated model (h_2) over the less saturated model (h_1) is approximated by a χ^2 -distribution. The number of additional parameters in Model 2 compared to Model 1 is the “degrees of freedom”. The formula to test the additional significance of a model is:

$$\chi^2 = -2 \sum_{i=1}^N \log \left[\frac{h_1(y_i)}{h_2(y_i)} \right]$$

with $h_1(y_i)$ and $h_2(y_i)$ as the marginal likelihoods of Model 1 and 2. The variance explained by the model can be computed by the fraction of the variance of the model — or elements of it — and the total variance including the residual variance. The fit of the final model will be described by comparing this model with the first model computed — the intercept alone considered as a fixed effect.

$$R^2 = 1 - \frac{S_{\text{res}} [\text{Model}]}{S_{\text{res}} [\text{Intercept alone as fixed effect}]}$$

RESULTS

Schizophrenia in daily life: subject-related characteristics

Schizophrenic subjects are assumed to be poor raters of their own mental states. Moreover, their emotional life can best be characterized as blunted and invariable in affect and withdrawn and inactive behaviorally.

In the first section of this chapter we investigate if the patterns in their reports or are random — as might be expected from theoretic formulations that argue that they are poor raters of their own mental state — or in contrast related to what we would expect from the cross-sectional assessments of demographic characteristics and clinical

symptomatology. We assume from clinical experience that schizophrenics have, in fact, a more diversified emotional life than is often thought. Furthermore we will check the assumption of behavioral withdrawal by assessing the activity patterns, the diversity of persons they meet, and the places they frequent.

Our first analysis assesses the relation between sampled experiences as they occur within daily life and subject characteristics. We will discuss between-subject differences on demographic variables and on psychopathologic subject characteristics.

Moment-to-moment experience

Moment-to-moment mood states — We first assess the relation of sampled moods and demographic information for mental health patients and controls subjects. None of the moods — feeling “cheerful”, “satisfied”, “anxious”, “lonely” and “demotivated” — assessed in daily life — were related to the subject’s sex or age. We found differences related to the demographic characteristics that are a result of the illness process for two of the five items. “Loneliness” was related to the subject’s living situation. Living with family protected patients from moment-to-moment “loneliness” ($F_{(3,111)}=3.37$; $p<0.02$). Also, subjects with a structured work situation ($F_{(1,110)}=4.94$; $p<0.03$) or — a trend — who worked out of home ($F_{(1,110)}=2.99$; $p<0.09$) were less lonely. Those significant differences were caused by the normal control subjects and vanished in the pathologic subsample. Higher moment-to-moment sampled anxiety were found in the subjects who were at home all day ($F_{(1,110)}=4.63$, $p<0.034$), earned no salary ($F_{(1,110)}=4.02$, $p<0.047$) and had no work-related day structure ($F_{(1,110)}=5.78$, $p<0.018$). These, of course, are also related to illness severity. When we restricted the analysis to mental health patients, all these relations disappeared. In contrast, a trend emerged for sex ($F_{(1,88)}=3.55$, $p<0.063$). Mentally ill men (1.91) were less anxious than women (2.38).

In the next section we will assess the moment-to-moment experiences of mental state in relation to subject characteristics of psychopathology (between subjects). Mood, sampled in daily life, differed in relation to cross-sectional psychopathology (Table 9.1). As could be expected, the depressive chronic mental patients were less “cheerful” compared to all the other groups ($F_{(3,105)}=2.78$, $p<0.05$); however, only the comparison with normals was significant. Interestingly, schizophrenics almost matched normals for that item. For anxiety we see high levels for schizophrenic patients and the residual category of chronic patients (in which a number of anxiety subjects were included). Depressive chronic mental patients took a middle position and normal controls had almost no anxiety at all ($F_{(3,118)}=6.91$, $p<0.001$). For loneliness we see a significant difference between pathologic and normal subjects ($F_{(3,118)}=4.72$, $p<0.005$). Mentally ill patients feel more lonely. Pairwise comparisons were only significant between normal subjects and schizophrenics.

The differences found cannot be explained by differences in severity alone but the specificity of the different illness conditions are also important. We assessed the relation with illness severity (as measured by the mean item score of the BPRS) for the three subsamples of chronic mental patients. Significant differences were found for “satisfaction” ($F_{(2,43)}=4.83$, $p<0.013$) and “anxiety” ($F_{(2,82)}=4.68$, $p<0.012$). The relation was curvilinear. In both cases the worst figures were for the medium intensity

group. This illness severity range contained the largest proportion of depressive and anxiety subjects. In contrast, assessments of global adaptation (measured by the mean item score of the SFCL) was linearly related to sampled levels of mood — a better adaptation is also a better overall mood. We found social impairment related differences for “anxiety” ($F_{(2,79)}=4.03$, $p<0.022$) and “loneliness” ($F_{(2,79)}=4.25$, $p<0.018$).

$\bar{x}$ (s.d.)	Alone			Family-In			Family-Out			Friends			Strangers		
• Schizophrenia	4.12	(1.49)					2.22	(1.28)		2.48	(1.53)				
• Depression	3.88	(1.60)					1.65	(0.79)		2.22	(1.62)				
• MH controls	4.36	(1.53)					2.35	(1.26)		2.19	(1.13)				
• Normal controls	4.30	(1.41)					1.22	(0.59)		1.35	(0.47)				
						4.44	(1.32)								2.69 (1.05)
	$F_{(3,105)}=2.78 *$			$F_{(3,55)}=1.00$ n.s.			$F_{(3,118)}=6.91 ***$			$F_{(3,118)}=4.72 **$			$F_{(3,118)}=0.29$ n.s.		
Scheffé's test	S	D	M	S	D	M	S	D	M	S	D	M	S	D	M
• Depression															
• MH controls															
• Normal controls		*					**		*	**					

• $p<0.10$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$; **** $p<0.0001$; gray areas are not important according to tests.

Table 9.1 Sampled experience by diagnostic group.

Moment-to-moment symptoms — In the next analyses we assess sampled symptomatology and — of course — restrict the study to the three subsamples of chronic mental patients: schizophrenia, depression and the residual mental health control group. The aggregated intensities of hallucinations were related to the cross-sectional assessment of illness severity — as measured by the mean item score of the BPRS (visual (VH): $F_{(1,41)}=19.17$, $p<0.0001$; auditory (AH): $F_{(2,74)}=5.01$, $p<0.01$). Interestingly, in our sample of chronic mental patients, reports of hallucinations in daily life were not restricted to the schizophrenic subjects, but “suspicion” was ($F_{(2,88)}=3.25$, $p<0.05$). We also compared cross-sectional ratings of positive and negative symptomatology (using BPRS item scores) with sampled psychiatric symptomatology. Both hallucination modalities were higher in the positive symptom and mixed group (high on both positive and negative symptom sets) (VH: $F_{(3,38)}=8.85$, $p<0.0001$; AH: $F_{(3,72)}=7.93$, $p<0.0001$). This type-related trend was replicated in the analysis of the schizophrenic subsample alone. AHs ($F_{(3,45)}=3.58$, $p<0.03$) and VHs ($F_{(3,19)}=3.29$, $p<0.05$) had higher intensities in the subjects rated with high on positive symptoms. For “suspicion” we found a trend in the whole sample (elevated in both high positive as well as high negative symptom groups: $F_{(3,80)}=2.27$, $p<0.10$) but this trend was not confirmed in the schizophrenic subgroup ($F_{(3,48)}=0.95$, n.s.). Because we have repeated measures of sampled experience for each subject, we can compute the variability over time. We operationalized variability using the standard deviation of the ratings over the beeps of one subject. These variabilities were related to the cross-sectional assessments of symptom type. Subjects characterized by positive symptoms were more variable: “suspicion” ($F_{(3,80)}=4.08$, $p<0.01$), hallucinations (AHs: $F_{(3,72)}=10.63$, $p<0.001$; VHs: $F_{(3,38)}=12.05$, $p<0.001$) and “fear of losing control” ($F_{(3,46)}=4.35$, $p<0.01$). The same trend but less pronounced was found in the schizophrenia patients.

Sampled symptomatology has a relation with global social adaptation, as measured by the SFCL. Both modalities of hallucinations — AH ($F_{(2,71)}=5.84$, $p<0.01$) and VH ($F_{(2,37)}=7.66$, $p<0.002$) — have a relation with levels of social functioning. A trend in the same direction is found in the schizophrenic subsample taken above: the sampled

intensity of AHs ($F_{(2,44)}=2.85$, $p<0.07$) and VHs ($F_{(2,18)}=2.65$, $p<0.10$) are related to social adaptation. This relation is less clear for “fear of losing control” which is significant for the whole mental health sample ($F_{(2,45)}=3.50$, $p<0.04$). Both “medium” and “low” social functioning individuals were different from the high performers, a significance that is lost in the schizophrenic subsample ($F_{(2,21)}=1.46$, n.s.).

Discussion — Most of the differences we have found for our sampled moment-to-moment mood states are self-evident at first glance and in line with the expectations from cross-sectional retrospective assessments using questionnaire and clinical interview data. However, the results are impressive because they are based on samples of moment-to-moment experiences collected by the subjects themselves. In our hypotheses we have formulated that the self-assessments of chronic schizophrenic subjects might be unreliable because they lack the insight that is necessary for self-monitoring. Responses should therefore be random. Furthermore and in contrast to this, self monitoring could be undifferentiated due to the blunted affect that characterizes chronic schizophrenic and depressive subjects. In contrast, the data show clearly that self-monitoring responses of chronic schizophrenic subjects are not undifferentiated and of comparable variability with other mental health patients (Table 9.1). As a consequence the hypothesis of blunting and undifferentiation can not be sustained. Compared to normal control subjects, the between subject variability in the sampled experiential aggregated characteristics was up to threefold greater in the mentally ill group. Some might argue that this high between subject variability is a reflection of a random answer pattern. If this would be true, it is not specific for schizophrenia because the other mental health patients also share this characteristic. Wildly random answer patterns, however, should primarily be found in the within subject moment-to-moment ratings and as a consequence lead to clusters of subject averages in the middle of the 7-point Likert scales. The result of this would be a small between subject variability which was not found.

Furthermore, when we compared the individual aggregated moment-to-moment subject ratings with cross-sectional ratings made by a skilled clinician — the traditional strategy in psychiatry to circumvent the unreliability of patient’s self-ratings — we saw that both matched to a large extent. In the study of sampled experiences and living conditions we demonstrated differences for “anxiety” and “loneliness”, but not for “cheerfulness”, “satisfaction” and “motivation”. Due to the nature of the ESM data collection process, no causal attributions were possible. The subject’s living condition is a consequence of the illness process but the living environment also shapes the moment-to-moment experiences of the individuals. We can assess the likelihood of these causal relationships but should be aware that changes in living condition, for instance quitting a job, have usually arisen elsewhere in the subject’s illness history and will, therefore, not always be related to the subject’s current mental state. Some interpretations, therefore, are more likely than others. We can imagine anxious people quitting their jobs but it is difficult to imagine lonely people choosing to live alone. A differential pattern of aggregated sampled emotions was also found for illness severity and global levels of (social) adaptation. Those differences were not mere reflections of severity, but in contrast mirrored the cross-sectional typology of the illness.

In conclusion, sampled experience and mental state — even in this very ill but homogeneous group of chronic mental patients — were not “blunted” nor random. The sensitivity demonstrated between-group, differed psychopathology. The observed differences were not mere reflections of illness severity.

Time use patterns

Time budgets (TBs) — best visualized with pie charts — are descriptions of the proportions of time that subjects spend in different situations. Previous studies have shown that time budgets are powerful indicators of the subjects’ daily life adaptation and differ between normal subjects in different cultures (Szalai et al., 1972) and ages (Larson, 1990), but also between subjects with different kinds of psychiatric diagnoses (deVries et al., 1988). There are different theories about time use in psychiatry. The principal one related to schizophrenia states that these patients are withdrawn. This should result in quantitative differences in time use patterns in the schizophrenic patients: less differentiation in activity patterns (unproductiveness, negative symptomatology); contact with less people (social isolation); and frequenting a less differentiated set of places (restricted mobility). Other theories postulate specific situations in which chronic mental patients spent their time. Mental illness results in a shrinking of social role performance which results in different activity, social contact and mobility patterns. Finally, specific theories combine time use and experiential appreciation. This topic, however, will be assessed in the next section. We will first discuss the TBs for persons, activities and places.

i. Persons

Figure 9.1 Time budgets of persons by subgroup.

Comparing time budgets for persons — We first analyzed who the subjects spent their time with. Figure 9.1 shows large differences between groups. When we compared the situation for the different categories of persons that were present at the moment of the

beep, between-group differences were striking for all codes except for “family out”⁵³. The time spent “alone” was larger for depressive subjects and significantly so in comparison with normal controls and schizophrenic patients. The time spent with family members who were part of the patient’s own household is greater for people with schizophrenia compared to chronic depressive subjects, who were alone more often, and normal controls, who spent more time with friends, colleagues and acquaintances. Finally, the normal controls were in the presence of strangers more often compared to all the pathologic groups.

$\bar{X}$ (s.d.)	Alone			Family (in)			Family (out)			Friends			Strangers		
• Schizophrenia	41.04 (24.80)			24.24 (25.72)						17.02 (20.17)			3.44 (5.92)		
• Depression	59.91 (25.82)			6.97 (13.94)						17.19 (17.97)			1.85 (3.29)		
• MH controls	45.49 (29.46)			11.20 (20.24)						24.21 (22.27)			1.16 (2.36)		
• Normal controls	34.84 (17.82)			6.15 (8.86)						31.74 (20.14)			8.35 (9.87)		
							6.68 (11.44)								
	$F_{(3,96)}=3.84 *$			$F_{(3,96)}=5.16 **$			$F_{(3,96)}=0.77$ n.s.			$F_{(3,96)}=2.54 *$			$F_{(3,96)}=4.22 **$		
Scheffé's test	S	D	M	S	D	M	S	D	M	S	D	M	S	D	M
• Depression	*			*											
• MH controls															
• Normal controls		*		*									*	*	*

• p<0.10; * p<0.05; ** p<0.01; *** p<0.001; gray areas are not important according to tests.

Table 9.2a Analysis of time budgets: persons.

• *COMPARISON WITH THE “BASIC PROPERTIES” STUDY* — In a previous study — the “Basic Properties” (BP) study — we compared the time budgets in different samples (Table 9.2b). In that study all the estimations were different for the different pathologic groups. Anxiety, pain and control subjects were similar on “being alone”-time, but not on “family in” where control subjects had the same frequencies as depressives and schizophrenics. When interpreting these figures we had to realize that other aspects of the samples can be responsible for the observed trends. Significant sex differences were found in “family out” (W>M: $F_{(1,201)}=6.69$; $p<0.01$) and “friends” (M>W: $F_{(1,201)}=3.94$; $p<0.05$). Age differences were found for “alone”, “family in” and “family out”. Having work led to different times “alone” and with “friends” (which included colleagues). Interestingly, “living alone” did not necessarily lead to more “alone” time. Those subjects matched time budgets of subjects living with their family, while both differed from the family-related time budgets of subjects living in groups ($F_{(2,172)}=6.40$; $p<0.002$). It was interesting to assess the influence of different sampling characteristics between the BP study and our current sample on TB data. The BP schizophrenic sample overlapped to some extent (45%) with the sample presented in this chapter. The depressive subjects, however, were not chronic patients but subjects who presented themselves recently for ambulatory care. In the BP study, alone time was not illness related. The time budget patterns of anxiety and pain subjects were similar to the control sample, while depressive subjects matched the schizophrenics. The high rates for social withdrawal (being alone) in our chronic depressive sample was an aspect of long-term illness course as was the reduction of time spent with the family. In contrast, “anxiety” and “pain” subjects spent a lot

⁵³ Family members that do not live in the subjects own household (see Appendix 5).

of time with family. Finally, the control sample of the BP study was selected to match some characteristics of the pathologic subjects. In comparison with our control sample, fewer subjects of the BP study worked and most of them spent large parts of the day at home.

$\bar{X}$ (s.d.)	n	Alone				Family (in)				Family (out)				Friends				Strangers				
• Schizophrenia	33	49.63 (28.33)				29.90 (28.47)				3.10 (6.12)				13.27 (17.02)				4.10 (7.17)				
• Depression	40	46.95 (21.36)				30.93 (23.45)				10.93 (14.72)				9.30 (9.54)				1.89 (2.85)				
• Anxiety	50	34.40 (18.96)				44.41 (25.91)				6.95 (9.74)				12.69 (13.54)				1.55 (2.93)				
• Pain	35	35.25 (15.39)				50.87 (20.98)				0.10 (0.42)				10.18 (13.45)				3.59 (7.37)				
• Norm. control	44	37.77 (15.82)				26.78 (21.29)				12.88 (19.92)				20.25 (12.50)				2.14 (3.13)				
		$F_{(4,197)}=4.65$ ***				$F_{(4,197)}=7.38$ ***				$F_{(4,197)}=6.75$ ***				$F_{(4,197)}=4.44$ **				$F_{(4,197)}=1.98$ *				
Bonferroni comp.		S	D	A	P	S	D	A	P	S	D	A	P	S	D	A	P	S	D	A	P	
• Depression										*												
• Anxiety		**	*			•	•															
• Pain		*	•			**	**				**											
• Norm. control		•				•		**	***	**			***	***	*	**						

• $p<0.10$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$; gray areas are not important according to tests.

Table 9.2b “Basic properties” of time budgets: Persons (Delespaul et al., 1992; deVries et al., 1993).

One might wonder if the demographic characteristics of the subjects such as living and work situation explained the differences between the groups. However, in this assessment we also have to account for the fact that the demographic situation of the subjects is often a consequence of the illness process (e.g. having work, living alone or with the family). We analyzed this for “alone” time in the next section.

Assessment of “time alone” — We analyzed the time that the subjects spent “alone” using a regression analysis with dummy variables for diagnostic groups, work (at home or out of home) and living situation, while trying to control for subject characteristics that are not illness-related (age and sex) and for illness severity (mean BPRS item score). We first discuss the *main* effects and later the *interaction* effects:

- **Main effects** — Significant main effects are found when — irrespective of other influences — average levels of all subjects with one characteristic is different compared to the other subjects, e.g. irrespective of diagnosis, men are different from women.

SEX AND AGE — We first analyzed the influence of non-illness related subject characteristics. The sex of the subjects in these samples was not related to “alone” time ($F_{(1,104)}=0.01$, n.s.) but there is a trend for age ($F_{(1,104)}=3.42$, $p<0.067$; a difference of 0.40% per year).

SOCIAL ADAPTATION — Subject characteristics of social adaptation such as work and living situation are influenced by the patient’s illness history. Global levels of adaptation (using the SFCL) do not explain the amount of time subjects spent “alone”. The fact that someone lived alone resulted in higher levels of “alone” time ($F_{(1,103)}= 17.73$, $p<0.0001$; a difference of 20%) as did the fact that the subject worked outside the home ($F_{(1,99)}=3.67$, $p<0.058$; a difference of 9%). Both effects had an independent influence on “alone” time.

ILLNESS — Illness severity (based on the BPRS) did not explain the amount of time subjects spent “alone”. Being (chronically) ill, however, resulted in a higher amount of alone time compared to the normal control subjects ($F_{(1,103)}=$

5.68, $p < 0.019$; approximately 14% more). No differences were found between the control chronic mental patients but there was a difference between chronic schizophrenia and refractory depression ($F_{(1,97)} = 4.71$, $p < 0.032$; $\pm 13\%$). However, for recent onset depression (see later in Table 9.2b) “alone” time was more comparable to schizophrenia. Although this was a cross-sectional comparison — we did not follow the time budget development of a cohort over time — our data indicated that refractory depression may lead to more social isolation than, for instance, schizophrenia.

- *Interaction effects* — Significant interactions are found when the pattern in one group of subjects is different compared to the patterns of another group, e.g. schizophrenic men are more, but depressive men less alone than women.

INTERACTIONS WITH ILLNESS TYPE — We first assessed whether social isolation (alone time) evolved differently by age between chronic mental patients and normal controls. We found no differences. When we restricted the analysis to the set of mental patients we saw that the amount of “alone” time was different by age and between chronic schizophrenics and refractory depressive subjects. Moreover, the best model included an interaction term showing that depressives and schizophrenics had a different evolution of social withdrawal over time ($F_{(3,93)} = 5.49$, $p < 0.0016$; $R^2 = 0.388$). Chronic depressive patients became more socially isolated by age — as normal subjects would be (see Larson, 1990) — but schizophrenics are not.

Figure 9.2a Evolution of “time alone” by age. Regression analysis results.

INTERACTION WITH SOCIAL ADAPTATION — We further analyzed how the indicators of social adaptation had differential influences on “alone” time. As demonstrated in the main effect analyses we found no differences with general social adaptation (SFCL-items). When taking living situation and work into account we found no significant interaction effects between normal and mental health subjects, but did within the group of chronic mental patients between depressive and schizophrenic subjects ($F_{(5,87)} = 5.86$, $p < 0.0001$; $R^2 = 0.502$). Figure 9.2b shows that the alone time of schizophrenics is affected by their living situation and working out-of-house. In contrast, those conditions do not matter very much for depressive subjects. We found no significant interaction between work and living alone in that group.

- *Summary model* — When we included higher order interaction terms and combined variables from different origins the interpretations became more complex,

especially since most of the demographic variables were illness-related. Our final regression ($F_{(7,85)}=6.557$, $p<0.0001$; $R^2=0.592$) included main effects for work, living situation and diagnosis (chronic depression or schizophrenia) and interaction terms, for the living situation with having work, age and diagnosis, and for work with diagnosis (Figure 9.2b).

Figure 9.2b Evolution of "time alone" by work, living situation and age. Regression analysis results for depression and schizophrenia.

The finding that the amount of time spent alone is not related to age in schizophrenics (Figure 9.2a), is not true irrespective of the subject's demographic situation. In contrast, it is also related to their living and work situation. When controlled for these demographic characteristics the age related patterns for depressive and schizophrenic subjects were the same. Having work raises "alone" time over age. This is a marginal effect when the subjects do not live alone, but extreme for subjects who live alone. We also found more alone time by age for subjects living alone who had no work but an inverse relation for the subjects who lived with other persons. While having work protected young subjects from being alone too often, it was not so for old people. By age the living environment — families or group homes — became more important to reduce social isolation. While the age related patterns in time alone were the same in chronic schizophrenic and depressive subjects, large differences were observed for the level. The schizophrenics had more extreme patterns. When living alone, they were more alone compared to depressive subjects, but when they did live with others they were less alone. For depressive subjects something paradoxically came from the analyses. The trend found in the schizophrenics was only replicated for elderly. In the young depressive group, the subjects who lived with other persons were more alone.

Discussion — Of course, the nuances in the relations can be caused by characteristics of our sample. We, therefore, restricted ourselves to the most robust trends. This means that non-significant findings such as the patients living environment — hospitalized or not, for instance — was not a significant predictor of alone time. Our final estimations show that the clear pattern of Figure 9.2a only holds

when we do not take the demographic differences between subjects into account. However, the differences might also be indicative of a different social adaptational pattern mediated by age for depressive and schizophrenic patients, because both having work and living conditions are related to the subject's illness process or history. Our original hypothesis concerning the TBs of schizophrenic patients stated that they are socially withdrawn while for depressive subjects we might expect a longing for social support. In both situations the trend should become more and more pronounced when the illness becomes chronic over time. Our data did not support this. In our sample the chronic schizophrenic patients spent a moderate amount of time alone and — except for the possibility of a non-interacting satellite position in a social group — should be considered as not withdrawn from this point of view. Moreover, this pattern does not seem to change over time. The effectiveness of depressive patients to organize social support around them is also marginal and even worsens over time. These patterns might be indicative of pathology-related indicators of (failing) coping which are of relevance to enhance our fundamental knowledge of the illness process and to guide the selection of treatment options.

Assessment of "time with family (household members)" — The family of schizophrenic subjects has been central in the formulation of theories on environmental etiologic theories of schizophrenia. Most of these theories have been abandoned over the last decennium because prospective and comparative studies could demonstrate that the assumed pathogenic factors in the family environment of schizophrenia patients were not specific to their families. The characteristics were also observed in all environments where people lived together in intensive situations — such as a severe illness. As a consequence, the causality attribution formulated as a consequence of the concurrent observation of specific family communication characteristics and schizophrenia — the illness being a consequence of the environment — could no longer be hold. Currently, one theory still exists. The "expressed emotion" (EE) theory is not a theory about illness onset, but about relapse. Specific communication characteristics exist in families where patients rapidly relapse after being discharged. This theory reflects an empirical finding and causality attributions are unimportant, especially since consequent studies have shown that the family's communication style can be altered by training, resulting in lower relapse rates. Also, reducing the face to face contact to less than 35 hours per week results in a comparative protection. An extensive overview of EE literature can be found in Leff & Vaughn (1985). Our own time budget based estimates of the presence of family members at the moment of the beep (Table 9.2a) were 25 hours and 30 minutes for the schizophrenics and only 7 hours and 20 minutes for the depressive patients per week. At the group level both estimates are within the protective range for relapse. In the next analysis we will assess which were the best predicting cross-sectional characteristics for the amount of time the patients spent with their family.

Frequent contact with family members that share the same household was, of course, related to the living situation of the subjects. The general level of social adaptation (SFCL) played a crucial role in the equation. Unexpectedly, the

frequency of contact with these family members was not related to illness severity (BPRS) but was related to illness quality (comparing chronic depression to chronic schizophrenia). The resulting equation was highly significant ($F_{(4,71)}=14.162$, $p<0.0001$; $R^2=0.666$). We saw that overall the schizophrenic subjects spent more time with family members. This frequency was related to their living situation and the level of social adaptation for both pathologic groups. Contact with family was lower for the subjects with the highest level of social adaptation who lived in a group (more independent subjects). The lower than zero predictions for the high social adaptation subjects were caused by the fact that in that range no observations by the mental patients were available to allow adequate estimations of “time with family”.

Figure 9.2c Evolution of “time with family” by living situation and general social adaptation. Regression analysis results for depression and schizophrenia.

Discussion — Our data indicate that the amount of time the patient spent with his family is related to his social skills. Patients with a poor social adaptation were more likely to spend more time with family. Frequently, this was so for schizophrenic subjects. This seems to confirm that the amount of time the patient and his family members have contact with each other on a daily base should not be understood as an indication of a overprotective mother but is more likely a consequence of the poor self-reliability skills — not illness severity — of the patients. Face-to-face contact for this subpopulation can raise up to 80 hours per week — well over the amount of time that was considered protective for relapse in previous research. Thus, social invalidation “causes” relapses. Treatment should therefore also focus on the training of daily living skills in those subjects who most need them. This might enhance their community survival — a finding clinicians are aware of for years, of course.

Summary — Our most interesting finding in the TBs of persons is the difference in alone time between schizophrenic and depressive patients by age. Young depressive subjects were less alone than schizophrenics but this changed dramatically by age. Also, a chronic course maximized the amount of alone time in depressive subjects. The high amount of alone time was even more intriguing for the depressive subjects who lived in a group. In contrast to schizophrenics in the same situation, they were alone more often. We consider this difference to be the result of an adaptational process both for the patients and for the persons who lived with them. Being alone could be a free choice or an unintentional consequence of environmental — not patient — withdrawal, caused by the illness characteristics. An influence of illness type (depression

and not schizophrenia) was warranted because illness severity was not significant in the equations. We found that having work reduced alone time of young subjects but not of older people. For those subjects their living situation was more important. The schizophrenics who lived alone were more alone irrespective of age. For young depressive subjects, however, not living alone led to more social isolation, which was unexpected. Here, too, the only possible explanation was the dynamic interactions between the patients and their environment which could be problematic in depression.

ii. Activities

Comparing time budgets for activities — The TBs of activities were largely similar between our groups (Table 9.3a). As we expected, differences were found for “doing nothing” and for “work”. In all those situations, the pattern differentiated between control and psychopathologic subjects, the latter group doing nothing more often — which was only significantly different for the depressive subjects — and less often at work. We also found a significant difference for the amount of time the subjects spent eating. None of the pairwise comparisons between samples for eating, however, were significant.

Figure 9.3 Time budgets of activities by subgroup.

$\bar{x}$ (s.d.)	Nothing			Work			Maintenance	Leisure	Social	Meals		
• Schizophrenia	25.25	(23.87)		9.28	(14.19)						12.20	(11.07)
• Depression	34.49	(40.14)		14.09	(14.59)						5.68	(8.77)
• MH controls	34.65	(37.79)		9.05	(11.08)						9.93	(12.68)
• Normal controls	5.78	(7.41)		31.97	(17.85)						5.49	(4.29)
							15.67	(12.81)	20.89	(15.69)	9.01	(8.49)
							$F_{(3,96)}=3.71 *$ $F_{(3,96)}=10.23***$ $F_{(3,96)}=0.60$ n.s. $F_{(3,96)}=1.15$ n.s. $F_{(3,96)}=0.76$ n.s. $F_{(3,96)}=3.15 *$					
Scheffé's test	S	D	M	S	D	M	S	D	M	S	D	M
• Depression												
• MH controls												
• Normal controls		*		***	**	**						

• p<0.10; * p<0.05; ** p<0.01; *** p<0.001; gray areas are not important according to tests.

Table 9.3a Analysis of time budgets: Activities.

- We might argue that chronicity partially compensated for the low levels of work and maintenance time in the acute population. In contrast, social time but also meal duration was drastically reduced by the evolving illness process. Most of the time was compensated by “doing nothing” but “work” and “maintenance” time were also higher in the chronic population. In contrast with the more acute patients we studied before, we could argue that role functioning is marginally restored in the chronic evolution of the illness, but important impairments remain.

* $p<0.10$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$; gray areas are not important according to tests; - no data published

Assessing the frequency of “doing nothing” — When we evaluated the time that subjects do nothing, we saw that there was a large amount of within-group variability. We therefore engaged in an assessment of other subject factors, within the population subset of a diagnostic group, that might have influenced the likelihood of doing nothing. We explored the time subjects spent “doing nothing” by assessing subject characteristics (sex and age), illness related assessments of social adaptation (work and living situation) as well as illness itself.

- **SUBJECT CHARACTERISTICS** — Irrespective of the group the subjects belonged to the frequency of “doing nothing” was related to the subject’s sex and age ($F_{(2,103)}=4.22$, $p<0.02$; $R^2=0.275$). Women reported doing nothing more often (an overall difference of 9%). Furthermore, independent of the sex of subjects the reports of doing nothing became more frequent with age (5% in 10 years).
- **SOCIAL ADAPTATION** — The subjects who had a structured daily activity were doing “nothing” less often ($F_{(1,99)}=17.05$, $p<0.001$; $R^2=0.383$). This difference was not related to the subject’s living environment nor to their social adaptation (SFCL).

- **ILLNESS** — We found, however, that the amount of time that the subjects spent doing nothing was related to illness severity (BPRS) and illness type (depression or schizophrenia) ($F_{(3,75)}=13.50, p<0.001; R^2=0.592$). Illness severity influenced the amount of time “doing nothing” more for depression compared to schizophrenia.

FINAL MODEL — The addition of having “work” to the equation did not improve the fit of the model for the prediction of doing noting, but the sex of the subject did. Our final model included sex, illness severity and illness type ($F_{(2,76)}=21.988, p<0.0001; R^2=0.605$). The amount of time patients spent “doing nothing” increased with illness severity and faster for depressives than schizophrenics and for women than for men. Estimated frequencies above 100% are once again due to the fact that not enough data are available from severely ill women.

Figure 9.4a Evolution of “time doing nothing” by illness, illness severity and sex.

Assessing the frequency of “working” or “maintenance” activities — The within-group variability for the time spent “working” had less variability than time “doing nothing”. Therefore, for “work” the influence of pathology was more striking. The factors that influenced the amount of time a subject spent working were largely defined by the fact that he or she was less ill and had work ($F_{(2,98)}=38.06, p<0.001; R^2=0.661$). No additional information enhanced the fit of the model significantly. In contrast to our expectations, maintenance was not related to the subject’s living situation but to illness severity (BPRS) and global levels of social adaptation (SFCL) ($\Delta=3.5\%$ by severity point; $F_{(2,73)}=9.46, p<0.001; R^2=0.454$). Only at low illness levels does a higher social adaptation lead to more maintenance activities. At higher illness levels the system collapsed. Because social adaptation is correlated with illness severity ($r=-0.478, p<0.01$) the likelihood of high illness and high adaptation was low and not enough data were available for an adequate estimation.

Figure 9.4b Evolution of “time doing maintenance” by illness severity and level of social adaptation.

Social activities were related to social adaptation (2% by scale point; $F_{(1,74)}=6.81$, $p<0.011$; $R^2=0.290$) but not to living with others, having work, illness severity, or illness type. Time spent eating, in contrast, was related to illness type (compared to depressive subjects the schizophrenics spent 7% more time eating: $F_{(1,106)}=8.72$, $p<0.004$; $R^2=0.276$). Here, too, there was no effect of illness severity or living situation and only a small portion of the variance was explained by the model.

iii. Places

The last time use pattern we will discuss here is the assessment of places. Data are presented in Figure 9.5 and Table 9.4a. The most striking observation was the lack of mobility in the patient samples. All the patients were home most of the time. Compared to the normal population the chronic mental patients were home most often ($F_{(1,97)}= 25.03$, $p<0.001$, $R^2=0.45$ a difference of almost 30%).

Figure 9.5 Time budgets of places by subgroup.

$\bar{X}$ (s.d.)	Home			Network			Workplace			Public Place			Transport		
• Schizophrenia	69.32	(22.57)					9.08	(13.40)		12.26	(11.16)				
• Depression	63.56	(22.79)					13.63	(12.23)		10.43	(7.58)				
• MH controls	57.59	(30.88)					14.15	(13.69)		4.87	(7.15)				
• Normal controls	36.55	(19.81)					34.47	(18.77)		10.19	(8.28)				
				8.05 (11.33)									7.90 (22.02)		
	$F_{(3,118)}=9.00$ ***			$F_{(3,118)}=1.16$ n.s.			$F_{(3,118)}=8.59$ ***			$F_{(3,118)}=4.16$ **			$F_{(3,103)}=1.97$ n.s.		
Scheffé's test	S	D	M	S	D	M	S	D	M	S	D	M	S	D	M
• Depression															
• MH controls										*					
• Normal controls	***	*					***	**	*						

• $p<0.10$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$; gray areas are not important according to tests.

Table 9.4a Analysis of time budgets: Subselection of places.

- **COMPARISON WITH THE "BASIC PROPERTIES" STUDY** — The lack of mobility in the chronic psychopathologic group was replicated in all our previous pathologic samples (Table 9.4b). Here, the differences between each category was signifi-

cant. We would like to have a better matched control population to adequately assess how the type of illness or illness-related secondary consequences such as the loss of a job affected those figures. Most of our theories of space and mental illness were specific for anxiety (agoraphobia) or schizophrenia (“social withdrawal”). We could not replicate the specificity of these relations. All the subjects were at home most of the time. Between groups, only the control subjects differed in amount of time spent at home, compensating the gain primarily by visiting their network and, of course, going to work. Mobility is low in schizophrenia and, interestingly, they spent a lot of the remaining time in “anonymous” public places, this contrary to the other groups. Here also, demographic factor-related differences were found.

Group	n	Home				Network				Workplace				Public Place				Transport			
• Schizophrenia	33	76.76 (21.78)				4.56 (8.69)				3.81 (8.51)				12.75 (13.77)				1.73 (3.42)			
• Depression	40	63.56 (24.35)				14.79 (20.89)				7.29 (14.11)				10.47 (9.81)				3.90 (4.23)			
• Anxiety	50	68.31 (21.21)				11.70 (12.03)				6.83 (13.53)				8.02 (7.98)				4.33 (5.32)			
• Pain	35	72.59 (19.51)				7.11 (10.14)				8.62 (15.73)				7.13 (8.25)				4.22 (5.09)			
• Normal control	44	46.80 (21.47)				22.65 (18.46)				13.12 (13.44)				9.69 (6.96)				7.45 (7.21)			
		$F_{(4,197)}=11.52 ***$				$F_{(4,197)}=8.66 ***$				$F_{(4,197)}=2.56 *$				$F_{(4,197)}=1.97 *$				$F_{(4,197)}=5.97 ***$			
Bonferroni compar.		S	D	A	P	S	D	A	P	S	D	A	P	S	D	A	P	S	D	A	P
• Depression		*				*				-	-			-				-	-		
• Anxiety										-	-			-	-			-	-		
• Pain										-	-	-		-	-	-		-	-	-	
• Normal control		***	**	***	***	***		**	***	-	-	-	-	-	-	-	-	-	-	-	-

• p<0.10; * p<0.05; ** p<0.01; *** p<0.001; gray areas are not important according to tests; - no data published.

Table 9.4b “Basic properties” of time budgets: Places (Delespaul et al., 1992; deVries et al., 1993).

Assessing the frequency of being home — What are the determinants for being at home in our current study? Generally and restricted to an analysis within the subgroup of mental health patients being at home was related to age (not to the subject’s sex) and to their living situation (being institutionalized or not — not to the fact that they worked). Finally, being at home was related to illness severity but not illness type ($F_{(3,48)}=7.31, p<0.001, R^2=0.56$).

Figure 9.6a Evolution of “being at home” by age, living situation and illness severity (Results of a regression analysis).

Assessing the frequency of being at work, friends or in public places — There was no difference in visits to friends and relatives, but there was for work. Normal subjects were at their workplace more often compared to the patients. Of course, this difference was largely explained by the fact that some subjects worked ($F_{(1,110)}=52.77$, $p<0.001$, $R^2=0.57$, a difference of almost 20%). The time observed in public places, finally, showed a significant difference between schizophrenic and mental health control subjects. This difference might be caused by a number of anxiety subjects in the latter group. Being in “public places” was more generally related to sex and age and, when this was controlled for, we found differences between chronic depression and schizophrenia in men ($F_{(5,111)}=5.38$, $p<0.001$, $R^2=0.44$).

Figure 9.6b Evolution of “being in public places” by age, sex and illness (results of a regression analysis).

Experience in context: persons present at the beep

In the next section we will assess how a specific aspect of the social environment — the number of persons present at the moment of the beep — affects the subject’s concurrent mental state. Moreover we will demonstrate how this process is different for normal and chronic depressive or schizophrenic subjects. Data will be analyzed using random regression models (MIXREG, Hedeker, 1993). The number of persons present at the moment of the beep was counted and recoded to 8 when more persons were present in the subject’s environment. Because alone/not alone differentiations are assumed to be important, we included a log-transformation of the number of persons into the regression equation⁵⁴. Logarithmic transformations are adequate when we assume a process which rises (or drops) quickly at start but diminishing differences in larger values. We will compare logarithmic estimation solutions with models that include squares of the number of persons to assess change when more people were present. The combination of linear, logarithmic and quadratic terms in the equation allows an optimal modeling at all levels of social environment density.

Mood and number of persons present

- “*Cheerfulness*” — When we modeled the effects of social environment on “cheerfulness” ($\chi^2_{(17)} = 808.52$, $p<0.001$), we saw a main effect for the normal/pathology differentiation — normal subjects feeling more cheerful overall. Moreover, the mood of normal subjects was not affected by the social environment. In

⁵⁴ The log transformation for the “Alone” condition (code 0) would result in minus ∞ and is therefore recoded to minus 6 the mirror value of the 1-7-point Likert scale maximum.

contrast, both groups of chronic mental patients were differently affected by the number of persons in the proximity of the subject at the moment of the beep. The alone/not alone differentiation was dramatic for depressive subjects. Mood dropped, however, as soon as more than one person was present, reaching a dip that almost equaled the alone situation when five persons were present. A marginal improvement was observed for six or more persons. Depressive subjects felt best when only one person was present. This was in contrast with schizophrenics who felt increasingly better when more persons were present, with a slight drop after six persons. For patients suffering from schizophrenia, living in large social environments is less critical than for depressive subjects (Figure 9.6a).

Figure 9.6a "Cheerfulness" related to number of persons present: differences between groups using logarithm of the number of persons.

- *"Satisfaction"* — We found a similar pattern for "satisfaction" ($\chi^2_{(16)} = 686.97$, $p < 0.0001$). The only difference was that normal subjects felt more satisfied when one person was present and this diminished when more persons were present. A mirror of this was found for schizophrenics. They felt less satisfied when one person was present while feelings of satisfaction were higher both in the alone and larger social environments.
- *"Anxiety"* in daily life could be modeled with an adequate fit using a small set of prediction variables ($\chi^2_{(6)} = 923.21$, $p < 0.0001$). Women were more anxious than men (a difference of 0.365 scale points). The highest levels of anxiety were found for the schizophrenic patients, the lowest for the normals, with depressive subjects taking a middle position. With no significant interaction terms, the anxiety levels were equally distributed over the groups for the number of persons present. All were less anxious when one to three persons were present. Explorations with non-significant higher order models including interactions between the number of persons and the pathology groups showed much higher context-related variability for the depressive subjects. For them, anxiety was higher in larger social environments than when alone.
- Feelings of *"loneliness"* were high for mentally ill subjects compared to normal controls. The normal subjects felt more lonely when alone but were not affected

by the number of persons when more people were present. For schizophrenic subjects the optimal range was between three and five persons. Interestingly, they felt lonelier in a larger party compared to when alone. In contrast, the depressive subjects felt less lonely when one person was present or more than seven. The resulting model was highly significant ($\chi^2_{(16)} = 1064.41, p < 0.0001$).

- The last item we analyzed in this assessment of moods in context was the “*lack of motivation*” (“I’d rather do something else”). The model we assessed for this item — we used the square of the number of persons — showed high environmental reactivity for all three groups ($\chi^2_{(17)} = 1367.28, p < 0.0001$). All modeled groups were less motivated when alone or when more people were present. The most motivating company was two to three persons for normals and schizophrenic subjects, but between four and six for depressives. Once again, the departure from the normal picture was more pronounced for the depressive subjects than for the schizophrenics. Activity motivation, of course, was related to the kind of activities. We modeled an equation for the mental patients that included different activities — doing nothing, work and maintenance, and leisure (Figure 9.6b). Our best fitting model included a log transformation on the number of persons. As a consequence, the complex reactivity to the social environment, created by the quadratic modeling, disappeared. Irrespective of the activity and the diagnostic group, demotivation became worse with the number of persons present. For depressive subjects, being alone was the best situation from a motivational point of view. For schizophrenics, “alone” was worse than a small group of persons, but equal from five or more subjects on. Finally, the schizophrenics were less motivated when doing nothing, a little bit more for work and most for leisure. This is also what we would expect. In contrast, chronic depressive subjects were extremely demotivated for work, but almost equally motivated when doing nothing or leisure activities (a difference of 0.20 scale points). For the depressive subjects doing nothing or leisure activities when alone seemed the best state from a motivational point of view. Schizophrenics were most motivated in small groups during leisure activities.

Figure 9.6b “Not motivated” related to activities and the number of persons present: differences between pathologic groups using the log of the number of persons.

Discussion: — The models presented in this chapter seem different from the model we published previously (deVries & Delespaul, 1989; Chapter 8). To understand this difference we should take three elements into account. Firstly, the method that we used in the previous study was a descriptive method without formal tests. Testing at that moment in time was impossible due to the small sample size and the heterogeneity of the sample. In this study we used the best possible statistical technique to assess the effects and confounds of between and within subject variability on the data — the random regression multilevel statistical techniques. Secondly, the data in our previous presentation were presented with a sketch of the trends we found. In this analysis we used real data in the graphs. Finally, in our previous analysis we collapsed the data over different items while in this analysis we preferred an item-by-item description to allow a discussion of between-item differences. When we collapse items we see that for alone/not-alone, a difference was also found in this replication. Generally speaking our patients felt worse in the alone situation. Moreover, for a number of items such as “anxiety”, “loneliness” and “demotivation” the comfort range⁵⁵ of the previous study was confirmed. For other items we did not find a difference for the number of persons present when the subject was not alone. Finally, only “satisfaction” did not match our original model.

The results that we have presented are modeled regression solutions and large between-subject differences exist. Having issued this warning, from those figures we can advise differentiating between the most optimal living environment for chronic depressive and chronic schizophrenic subjects — patients that are usually not differentiated in the residential wards of mental hospitals. Also, the optimal working environment from a motivation point probably will be different from the optimal living environment. The results, however, are less clear — leaving room for even more relevant between subject variation. In general, the optimal living environment for chronic schizophrenics is larger (3-5 persons) than for depressive subjects (1-2), while the work environment for schizophrenics should be small (1-2) but can be larger for depressive subjects (5+). As we have demonstrated ESM allows this differentiation at the group level showing the influence of the illness itself on daily life adaptation. In addition, the ESM assessment procedure can also be used for individualized assessments which would make it possible to customize the living environment of our patients to their personal preferred mental state. Of course, this is not always possible but we have used this strategy in personalizing rehabilitation programs. Chapters 12 and 13 discuss the possibilities but also the limitations of these strategies.

Pathology and number of persons present

Compared to moods in daily life the assessments of psychopathology could be fitted with a less extensive model. With the exceptions of auditory hallucinations, all items were unresponsive to the number of persons present. A significant slope was found but its estimate was marginal ($\Delta=0.007$ per additional person present). Small increases in symptoms were found when the number of persons was larger for “visual hallucinations” ($\chi^2_{(5)}=319.15, p<0.001$), “suspicion” ($\chi^2_{(6)}=951.84, p<0.001$) and “fear of lo-

⁵⁵ The environment in which the most positive mental state was observed.

sing control" ($\chi^2_{(5)} = 366.76, p < 0.001$), but decreases were found for "derealization" ($\chi^2_{(6)} = 684.78, p < 0.001$). Because the psychopathology assessments were targeted at schizophrenic symptomatology, understandingly the schizophrenic subjects rated highest on all five items. The depressive subjects reported almost no "visual" or "auditory hallucinations", while for "suspicion" and "fear of losing control" they assumed a medical position between normal and schizophrenic subjects. Finally, depressive subjects matched the ratings of the schizophrenic subjects for "derealization".

Figure 9.7 Psychotic symptomatology related to number of persons present.

In contrast to the other psychopathology items the model for auditory hallucinations was much more complex ($\chi^2_{(16)} = 491.35, p < 0.001$). Here, a log transformation of the number of persons was included in the equation. We found the lowest intensity of hallucinations when one or two persons were present, very similar to the picture we found for "demotivation". Interestingly, estimations of more complex models for the other symptomatology items showed a similar pattern although they yielded no significant improvements over the original model and were not graphed in the Figure 9.7. With the exception of "derealization", the best environment was when only one person was present. For "derealization", in contrast, the best environment was when four or five persons were present. This matches the process described for the moods in this schizophrenic population.

DISCUSSION

In this chapter we have assessed the influence of subject characteristics and properties of momentary contexts on illness experience. These daily life observational findings are important. We generally accept the fact that the daily life functioning reflects the specificity of the mental illness process. The inclusion of axis V in the DSM system, the assessments of disabilities and impairments in the WHO classification of diseases and the growing awareness that assessments of "quality of life" should be included in outcome research in somatic medicine, all are illustrations of the fact that daily functioning is accepted as an important determinant of illness. Moment-to-moment chan-

ges in cognitions, emotions and behaviors are indicative of the subject's adaptation to his or her illness. Some might argue that the relevance of information on daily life functioning is restricted and only useful as a characterization of the illness in epidemiologic research or as an indicator of the daily reality of our patients in therapy. We are convinced that the way subjects experience and cope with their environment is the most basic datum for a future in-depth study of the etiology of psychopathology.

What happens in daily life is dynamic. This adaptational process is the result of a continual reciprocal interaction between the subject's illness and the environmental responses to his or her cognitions, emotions and behaviors. An enhanced understanding of this daily life experience allows us to determine the different components that form the illness nosology in a renewed way. Hopefully we will enhance our knowledge of the relation between symptoms and etiology including brain dysfunction and determine better assessments of the nosologic heterogeneity. A possible spin-off of studying the relation between brain, environment and illness experience from a practical point of view is that it might help us to design more appropriate sheltered living environments and customize therapeutic interventions to maximize the subject's optimal experience.

In the present study we were not successful in replicating the standard characterization of the chronic mentally ill schizophrenic as a withdrawn and affectively blunted individual. Their daily life experience was unrelated to subject characteristics that had no relation to illness history such as age and sex. In contrast, moment-to-moment sampled experience was related to the subject's living situation and employment. However, these subject characteristics were considered consequences of the illness process — especially illness severity. In most situations no significant differences were found between the different pathologic groups, although the experience of the normal control subjects was different from the mental patients. Illness severity and the differentiation between positive and negative illness types were important variables in explaining illness experience in daily life.

Time use patterns, too, were significantly different between groups. Some of the variables showed dramatic effects which reflected the daily life adaptation. We demonstrated that the overall amount of time a subject spent alone changed by age for depressive subjects — an increase from 20% to 80% between age 20 to 60 — but not for schizophrenics. Although this data was not collected in a prospective cohort followed over time but was based on a cross-section of different subjects by age, we can still argue that illness differently affects the daily life adaptation. The use of multilevel random regression techniques guaranteed an optimal assessment of within-subject environmental differences, of course limited to the conditions in the sample.

We think that those figures are sensitive indicators of the illness process as it becomes chronic over the years. We assume a reciprocity between the subject and his environment: being alone can be a consequence of the subject's withdrawal or the retraction over time of persons that were originally in the subject's environment. Considering the fact that living with a depressive person is not rewarding, we can conclude that chronic depressive subjects are increasingly at risk of being more alone with age. Growing solitude with age — although in a less extreme form — is a normal development (Larson, 1990). Our data show, however, that a far-reaching generalized interpretation

is premature. The problem of solitude is greater for subjects who live alone — depressives as well as schizophrenics. No age-related differences were found for subjects who lived with their family and when they also did not work we even saw a reduction of alone time by age. The living and work situation of subjects are “protective factors” for growing solitude with age. Of course, we should take into account that a therapeutic use of this finding is limited by the fact that these protective living conditions are also a consequence of the illness process. Reversal of their effect on solitude by changing the living condition of our patients might, therefore, be of marginal utility.

In chronic psychiatry the prevalence of schizophrenia is high. As a consequence the illness process and quality of life of other diagnostic groups such as chronic depressive subjects is largely underestimated. Due to the effects of the evolving chronicity the realities of daily life of chronic depressive patients match the impairments found in chronic schizophrenia. It is, however, an inadequate strategy — as is often done in chronic psychiatry — to disregard the original or current psychiatric diagnosis. Our study has demonstrated that specific risk factors apply to the non-schizophrenic chronic mental patients and rehabilitative interventions should be adapted to the specificities of their illness. Using our observations we might even consider building customized living environments for the different subgroups.

We have also found that the amount of time subjects spent with family was unrelated to their general level of social adaptation when they live alone. In contrast, for subjects who have a high level of social functioning, the amount of time they spend with their family was low, even when they lived in the same house. An inverse relation was found for subjects with a low level of adaptation. Considering that the amount of face-to-face contact is problematic for subjects who live in high “Expressed Emotion” (EE) environments (Leff & Vaughn, 1985), the subjects — both schizophrenic and depressives — are at risk of relapse, especially when they have a low level of functioning. This seems to indicate that the critique of parent groups on the EE concept is true (Hatfield, 1987; Hatfield et al., 1987). High EE is also a consequence of a more severe illness development and criticizing parents accordingly is inappropriate. Nevertheless, ESM can be used to define risk groups and to account for the realities of daily family life. It allows us to customize interventions — not based on a undifferentiated cross-sectional diagnosis as high EE — but accounting for actual activities, emotions and cognitions of the individual family members and their mutual interactions.

Importantly, the risk assessments made by ESM can be improved above the differential effect of the physical environment by including the subject’s emotional reactions in the evaluation. Using ESM we were able to assess the effects of the number of persons in the environment on the patient’s mental state. We found differences between groups and for different mental states. Optimal experience in depressive subjects was best when a small number of subjects were present. In contrast, the moods of schizophrenic subjects were best when four to six persons were present. Interestingly, most of our therapeutic environments are designed using less than adequate contextual characteristics (see also Kooyman, 1993). Both chronic depressive and schizophrenic individuals often live in groups with more than 10 subjects and their work environments are usually even larger (up to 30 patients). Those are non-rewarding situations and should be avoided.

Although assessments at the group level are important to understand the specificities of the illness process, we want to warn against broad and generalized interpretations over a whole diagnostic group. ESM is a method of choice to assess the specific environmental responses of individual subjects. Because of the large within-group variabilities in the data we should select the most appropriate living and work environments using this individualized information. The final test of such a strategy would be a higher success rate of ESM-mediated rehabilitation (“personalized rehabilitation”) in a randomized clinical trial.

Chapter 10: Hallucinations in Daily Life ⁵⁶

Schizophrenia is a heterogeneous disease. Different authors have argued that the current diagnostic tradition might be a hindrance for further research in schizophrenia (van Praag, 1992; Strauss, 1994). Subtypes that are currently evaluated prove more useful in linking the clinical phenomenology of the disorder to underlying pathophysiologic processes (Fenton & McGlashan, 1991). Costello (1993) defends an even more open approach to diagnosis and argues that “researching symptoms of schizophrenia — rather than syndromes — addresses two issues: a) current psychiatric diagnostic systems produce data of questionable reliability and validity; and b) schizophrenia, as it is currently defined in various systems, is such an intricate and loose concept that the professional who finds significant correlates of the diagnosis has little idea what causal mechanisms might be involved or even where to look for them.”

In this study we focus on hallucinatory experiences. Hallucinations — when lasting more than a month if not effectively treated⁵⁷ — are important but not sufficient symptoms to determine if a subject is schizophrenic. According to DSM-IV (APA, 1994) only one additional symptom is necessary to meet the criterion A (“acute phase”) symptoms. Even verbal hallucinations alone are sufficient if they consist of “a voice keeping up a running commentary on the person’s behavior or thoughts, or two or more voices converse with each other”. Hallucinations are defined as:

“... a sensory perception that has the compelling sense of reality of a true perception but that occurs without external stimulation of the relevant sensory organ. Hallucinations should be distinguished from illusions, in which an actual external stimulus is misperceived or misinterpreted. The person may or may not have insight into the fact that he or she is having a hallucination. (...) The term hallucination is not ordinarily applied to the false perceptions that occur during dreaming, while falling asleep (hypnagogic), or when awakening (hypnopompic). Transient hallucinatory experiences may occur in people without a mental disorder” (APA, 1994, p. 767).

As recognized in the DSM-IV, non-psychiatric individuals are susceptible or vulnerable to hallucinations (Launay & Slade, 1981; Romme & Escher, 1989). The psychiatric literature is filled with long lists of names of famous and sometimes unexpected individuals who hallucinated (Johnson, 1978)⁵⁸. What differentiates the pathologic from the non-pathologic hallucinator is the amount of time hallucinating (Lindsley, 1963). In a large prevalence study of symptomatology in different psychopathologic

⁵⁶ An abbreviated report was published by Delespaul, P. A. E. G., & deVries, M. W. (1994) (The Experience and Social Context of Auditive Hallucinations. *Schizophrenia Research*, 11(2), 187) and preliminary results were presented in conferences in Maastricht (1991), Geneva (1991), Les Diablerettes (1994) and Lunteren (1994).

⁵⁷ Primarily treatment with a neuroleptic; withdrawal of a psychotropic substance can result in Criterion A symptoms but these last less than a month and are not indicative of schizophrenia.

⁵⁸ Socrates, Plato, Aristotle, Byron, Dumas, Mozart, Wagner, Hitler, Attila, Napoleon, Joan of Arc, Luther, Calvin, Descartes, Newton, Churchill, Stalin,... (Johnson, 1978).

diagnoses from times before DSM-III-R, Zigler & Phillips (1961) found hallucinations in 19% of the psychiatric patients over all categories (n=793): 35% in schizophrenics, 11% in manic depressives, 12% in character disordered and 4% in neurotic subjects. Astrup et al. (1962) reported a prevalence of 87% in schizophrenia. The divergence in both studies dating back from the same period is striking. This is not improved in the DSM-III-R: Andreasen (1990) reported a prevalence of hallucinations with severe intensity in 49% of the schizophrenic subjects while Fenton & McGlashan (1991) found only 28% meeting the same criterion. Theoretically, hallucinations are possible in each sensory organ. DSM-IV lists the following hallucinations:

- Auditory* — A hallucination involving the perception of sound, most commonly of voices. Some clinicians and investigators would not include those experiences perceived as coming from inside the head and would instead limit the concept of true auditory hallucinations to those sounds whose source is perceived as being external. However, as used in DSM-IV, no distinction is made as to whether the source of the voices is perceived as being inside or outside of the head.
- Gustatory* — A hallucination involving the perception of taste (usually unpleasant).
- Olfactory* — A hallucination involving the perception of odor, such as of burning rubber or decaying fish.
- Somatic* — A hallucination involving the perception of a physical experience localized within the body (such as a feeling of electricity). A somatic hallucination is to be distinguished from physical sensations arising from an as-yet undiagnosed general medical condition, from hypochondriacal preoccupation with normal physical sensations, and from a tactile hallucination.
- Tactile* — A hallucination involving the perception of being touched or of something being under one's skin. The most common tactile hallucinations are the sensation of electric shocks and formication (the sensation of something creeping or crawling on or under the skin).
- Visual* — A hallucination involving sight, which may consist of formed images, such as of people, or of unformed images, such as flashes of light. Visual hallucinations should be distinguished from illusions, which are misperceptions of real external stimuli (APA, 1994, p. 767).

Auditory (AHs) and visual hallucinations (VHs) are the most common types in schizophrenia. Slade & Bentall (1988) computed a meta-analysis of 16 different studies (also from pre-DSM-III-R times) and concluded that the prevalence of AHs is 63% and of VHs 29% in schizophrenia. Because of the high prevalence, specific interventions were developed to control and minimize auditory hallucinations (Haddock et al., 1993; Bentall et al., 1994; Chadwick & Birchwood, 1994; Kingdon & Turkington, 1994). They focus primarily on coping with these symptoms (see Liberman, 1988). How can we understand the pathogenesis of hallucinations? Different monographs on the topic have been published (Ey, 1973a, b; Siegel & West, 1975; Johnson, 1978; Slade & Bentall, 1988; Asaad, 1990; Romme & Escher, 1993) and it is far from my ambition to give an exhaustive overview of the etiologic theories or the empirical research data. Heilbrun (1993) argues that “a fully satisfactory theoretical explanation of psychotic hallucinations that would be symptomatic of schizophrenia still eludes us” (p. 61). He gives an overview of proposed explanatory models. Single state theories assume that one condition generates a hallucination. In a first subgroup the working mechanism is primarily psychological. Examples are sensory deprivation, suggestibility as a personality trait, a cultural belief system or induction by hypnosis or somnolence, misattributions due to cognitive deficits in the self-monitor process — even the mislabeling of inner speech — or decompensatory reactions to stress-induced

arousal. The second subgroup in the single state etiologic theories of hallucinations assumes a biologic condition such as brain dysfunctionality — a tumor or epilepsy — and the consumption of psychotropic substances such as LSD and mescaline that trigger the serotine system and amphetamines in the dopamine system. Multifactoral theories of hallucinations integrate different aspects into a comprehensive model. Examples are Sarbin's Ecological Theory (Sarbin, 1967), Heilbrun's Cognitive Vulnerability Theory (Heilbrun 1973, 1984) and the Slade-Bentall Five Factor Theory of Hallucinations (Slade, 1976; Slade & Bentall, 1988). This last model suggests that stress events (factor 1) lead to a generalized elevation and disturbance of mood (internal arousal) and interact with a hallucinatory predisposition (factor 2) and the expectancy (a perceptual set in reality discrimination — factor 3) to emerge as hallucinations. However, external stimulation (factor 4) can block this process due to an overload in the limited capacity of consciousness. Internally generated events such as hallucinations will only emerge when no external stimulation is provided (e.g. in sensory deprivation). Finally, hallucinatory experiences might be reinforced (factor 5) over time because often these experiences lead to the reduction of aversive mood states — subjects feel relieved afterwards. We will use the Slade & Bentall model as a frame of reference for the discussion of our data.

When studying schizophrenic hallucinatory symptomatology we are faced with serious obstacles, because the anomalous perceptions are part of the private experience of the patient. We generally ask patients to condense their experience into a description of the phenomena as they often or usually occur. This is particularly true when using questionnaires, interview or survey data. The research reported in this chapter suggests that we should examine not only the retrospective characterization of an individual's experience, but that phenomena such as visual or auditory hallucinations and delusions should be described as they occur under the circumstances of daily life. Such a method frees the data as much as possible from the person's own beliefs. In this chapter we will present data — collected by the Experience Sampling Method (ESM) — on schizophrenic experiences in real life situations. We will present data in an attempt to characterize the psychologic context of auditory and visual hallucinations: time of the day, places frequented, persons met and activities performed, as well as additional information on the concurrent experiences of those hallucination moments. In a concluding discussion we will dwell upon the influence of emotions on hallucination, presenting a theory that can be used to explain the emergence of hallucinatory experiences and the perseveration of the phenomenon over time. This will lead to treatment options for the psychologic management of this important and invalidating symptom for patients suffering from schizophrenia.

METHOD

Subjects

The present study on daily life hallucinatory experiences was conducted on the same sample of chronic mental patients described previously (Chapter 9). All patients were

in ambulatory or clinical treatment in Maastricht during the study assessment period. After the normal data quality control procedures for ESM we selected 57 chronic schizophrenic subjects and 35 chronic controls. In the latter group we merged the original chronic depressive and mental health control subsamples.

The Experience Sampling Method

Data was collected using the Experience Sampling Method (ESM), a random signal contingent sampling technique of cognitions, emotions and behaviors collected in the natural flow of daily life. All subjects participated in ESM for a period of 1 week, being sampled 10 times each day. Contingent on a signaling beep, the subjects fill in reports that include measures of ongoing thought, mood, motivation and current activity, as well as the social circumstances and places that a person frequents. Experiential assessments include ratings of anxiety, depression, tension or, in contrast, relaxation, motivation and activation. ESM studies of verbal hallucinations were previously presented (Delespaul & deVries 1991a, b, 1994a, b; Visser, 1993). No studies of visual hallucinations have been presented yet.

Self-reports of hallucinations

“Hallucinations” are percept-like experiences which occur in the absence of an appropriate stimulus. Those percept-like experiences have the full force and impact of an actual perception. They are not amenable to direct and voluntary control by the subject” (Slade & Bentall, 1988, p. 23). Self-assessments of hallucinations, as well as of all other positive schizophrenic symptoms, are problematic due to the assumed lack of insight of these patients (Kafka, 1993; Birchwood et al., 1994). Therefore, they are assumed unable to recognize that what they see or hear, is not based on an external perception and what they think has no root in facts. According to Frith (1992) the central deficit in schizophrenia is their inability to self-monitor. Schizophrenic subjects are not able to compare their experiences with the external reality, both perceptually and cognitively. This deficit is the root of the symptom-formation. It is also the reason why their assessment through self-monitoring — without the help of an evaluating clinician — is problematic. Consequently, crucial for the assessment of these symptoms is an evaluation by a clinician. He or she must ascertain that the phenomenon is not an “illusion” — an actual external stimulus misperceived or misinterpreted.

Does the nature of schizophrenic symptomatology make the assessment through self-monitoring impossible? According to the DSM-IV definition (APA, 1994) hallucinations can be recognized as such by the patients. Frith (1992) points out that in schizophrenia abnormal perceptions are assimilated through normal logic. Therefore, the emergence of cognitive dissonance is still possible and the schizophrenic subject can become aware that something is wrong with his perceptions or cognitions. This might not be so in acute phases of the illness, but it is definitely the case in chronic patients. The majority of chronic schizophrenic subjects have learned that certain perceptions do not match external reality and are not shared by others. Intriguingly, this does not correct the actual hallucinatory observations, nor the emotional reactions to those perceptions. Although (chronic) schizophrenic subjects usually know that what they

see or hear is not real, they will still be anxious and act accordingly, for instance, by withdrawal or responding to the verbal hallucinatory instructions. This ambiguity and double-level experience is central in chronic schizophrenia and can confuse a diagnosis. While the corrective emotional response to the visual and auditory phenomena is disturbed, self-monitoring is still possible because subjects have learned to recognize illness-related phenomena. Neuropsychologically, assessed deficits are fundamental in schizophrenia and the performance of different tasks often seems impossible. Unexpectedly, complex activities are still possible when assessed in real life (van de Gaag, 1992). This is also true for self-monitoring. Our previous ESM studies have proved that schizophrenic subjects can self-monitor and that the collected data are valid (Delespaul & deVries, 1987; deVries & Delespaul, 1989).

How can we measure visual and auditory hallucinations with ESM? Ideal ESM questions are short, non-technical cues. Because the subjects will have to fill out the ESM-forms repeatedly, long instructions will not be read and, therefore, will not enhance validity (see Chapter 3). Using clinicians and schizophrenic patients as informants⁵⁹ we selected the phrases that the patients use to describe their hallucinations. In this way we formulated the questions "I hear voices" and "I have apparitions" ("verschijning" in Dutch) to rate auditory and visual hallucinations repeatedly. The formulation for auditory hallucinations, while restricted to verbal hallucinations, is more effective than, for instance, "I hear noises". Subjects implicitly know that a normal conversation or a voice on the radio is not the "voices" that are intended in our question. The formulation "I hear voices" was successfully applied in different cultures. We assume this will be not true for "I see apparitions", which might only work in a Catholic culture such as the south of The Netherlands. There, the Dutch word "verschijningen" (apparition) is commonplace and refers to apparitions of the Holy Mother. It is effective by analogy as a cue for the repeated assessment of visual hallucinations.

Analysis

The analyses for the study of hallucinations in daily life are based on elementary statistical techniques as proposed by Larson & Delespaul (1992). Whenever possible we have computed summary scores by subject for the appropriate parameters. These are used in standard statistical group comparison techniques such as χ^2 -tests and analyses of variance. In a specific analysis we were interested in the kind of variables that best predicted hallucinations in daily life. For this we used discriminant analysis techniques. Symptom moments were defined using cut-off scores on the Likert type items that corresponded to 25% of the beeps for all subjects. This resulted — both for auditory and visual hallucinations — in "no hallucination" with a score of 1 and "hallucination" for scores of 2 and higher⁶⁰. Comparisons between concurrent mental states, including the use of psychotropic "drugs", in hallucinatory and non-hallucinatory

⁵⁹ An even better strategy is to use a "focus group" to define the items. In such a group different members of the target population meet and define together the best matching description for the concept under study. Often, a focus group gives good and unexpected descriptions. This strategy is especially useful when sampling in other cultures.

⁶⁰ Cut-off scores for the other pathology items were: "suspicion" (3+); "thoughts difficult to express" (4+); "obsession" (4+), "thought control" (4+); "derealization" (2+) and "fear losing control" (2+).

moments were assessed using traditional uni-level regression techniques on aggregated data at the subject level. In these analyses the hallucination condition was defined as a within-subject effect. Symptom episodes were defined using the cut-off procedure to define uninterrupted sequences of symptomatology. Codes were given to define the last report before the episode, the first report in the episode, reports in the middle, the last and the first after the episode (see Appendix 5). Questions about moments were assessed with multilevel random regression models (see Chapter 9).

RESULTS

Hallucinating subjects (hallucinators)

Sampled hallucinatory experiences: frequency and duration

Compared to the other non-schizophrenic chronic mental health patients — primarily depressives — more schizophrenic subjects reported hallucinations (visual (VH): 62.5% and 10.5%, $z=3.46$; $p<0.001$; auditory (AH): 49.1% and 17.1%, $z=3.09$; $p<0.001$). Unexpectedly but non-significantly, we found more subjects with VH than with AH in our schizophrenic sample (62.5% and 49.1%, $z=1.10$; n.s.). An equal number of subjects had auditory hallucinations but no visual ones, or visual but no auditory ones (18% each).

Figure 10.1 Prevalence of sampled experiences of hallucination: numbers of subjects who hallucinate and frequency of hallucinations.

As we could expect from the previous data, the amount of time each hallucinating subject is observed hallucinating is different between the groups. Overall, the frequency of hallucinatory moments happened around one third of the time. Subjects of both groups were hallucinating visually at equal frequency ($F_{(1,15)}=0.00$; n.s.), but schizophrenic subjects had more AHs than the non-schizophrenic chronic mental patients (Figure 10.1). The difference was not significant, however, due to the large within-group variability and the unequal n's ($F_{(1,33)}=0.96$ n.s.). Three patients with AHs (11%) hallucinated on each reported beep — one of them constantly at the maximum intensity level. Ten patients (36%) heard voices more than half of the time. One patient had VHs all the time and five (33%) more than 50% of the time. Both distributions were comparable. Although the frequency of hallucinating subjects differed between schizophrenic and non-schizophrenic patients, we found no difference for the

overall amount of time, the number and duration of episodes, and the intensity and moment-to-moment variability of either visual or auditory hallucinations.

Because of the small number of hallucinating non-schizophrenic subjects we restricted the further analyses to the schizophrenic patients who “hear voices”. Within this group, too, we found no difference for the overall amount of both visual and auditory hallucination time, the number and duration of episodes, and the intensity and moment-to-moment variability of the pathologic experiences. Nevertheless, hallucinating schizophrenics — excluding those who hallucinated all the time — were experiencing such episodes 26.32% of the time for visual and 33.20% for auditory phenomena. Assuming a sampling window between 7:30 to 22:30 — 900 minutes each day — we concluded that 3 hours and 57 minutes a day were spent in a state of VHs and one hour more — 4 hours and 59 minutes — during AHs. An average visual hallucination episode lasted 144 minutes and an auditory episode 190 minutes.

Discussion — Due to the ESM time sampling procedure the assessment of episode frequency and duration was approximated and was not always accurate (Appendix 3). The frequency could be overrated due to our operationalization of events (Appendix 5) that assumed that the first positive beep of a day or a positive report after a missed beep always started a new episode. On the other hand, episode frequency could also be underrated because we interpreted two consecutive positive beeps as a lasting hallucinatory experience. The ESM randomization schedule generated beeps at intervals of 15 minutes to 3 hours. So the likelihood of symptom-free moments between two positive beeps also had to be considered. Only a small proportion of hallucination episodes were sampled totally and did not include missing data. From the analysis of those non-missing episodes we concluded that we underrated the amount of episodes and that short hallucinating periods occurred more often than our own rating. Therefore, the dynamic moment-to-moment changes in hallucinatory experience are even more impressive. Most schizophrenic patients are part-time patients. Unfortunately, the repeatedly recurring symptoms are often so disturbing that the patients are unable to capitalize on the daily occurring symptom-free periods and develop appropriate coping strategies.

Why, considering the problems related to the adequate assessment of hallucination episodes with time sampling techniques, have we not used event sampling? In Chapter 6 we have argued that event sampling strategies are unreliable for assessing event duration. Moreover, hallucination experiences are not discrete events and when monitored continually the recognition of an event can be problematic.

Subject characteristics and sampled hallucinatory experiences

Who are the hallucinators? Can we predict sampled hallucination experiences from demographic subject characteristics, cross-sectional assessments of psychopathology, sampled experience and sampled psychic symptomatology? To answer this, the results of a stepwise discriminant analysis are presented.

Demographic characteristics such as sex, age and life circumstances — work and living environment — were unrelated to the subjects who reported hallucinatory experi-

ences in daily life. In contrast, significant predictions could be built using cross-sectional assessments of psychopathology. The most severely ill and positive type schizophrenic subjects were the subjects who hallucinated most often. Social functioning (SFCL-rating) was less important but still significantly related to sampled experiences of hallucinations (VHs: $\chi^2_{(3)}=17.98$; $p<0.001$; AHs: $\chi^2_{(3)}=19.04$; $p<0.001$). We also estimated stepwise discriminant functions to predict hallucinators who hallucinated more or less often using averaged ESM ratings of experience. A significant equation was computed for VHs ($\chi^2_{(5)}=14.57$; $p<0.015$) but not for AHs ($\chi^2_{(5)}=8.77$; n.s.). The most important emotion was experienced “anxiety”. Anxious schizophrenic patients had a high likelihood to be a hallucinator. At the next step, sampled experiences of “loneliness” and “demotivation” were entered in the equation. Finally, “cheerfulness” completed the model while “satisfaction” was skipped because it had a negligible effect. The prediction of sampled frequencies of auditory and visual hallucinators using averaged ESM ratings of other psychopathology items offered an intriguing difference between VHs ($\chi^2_{(3)}=8.37$; $p<0.04$) and AHs ($\chi^2_{(3)}=9.38$; $p<0.03$). “Suspicion” was a salient item for both kinds of hallucinations but “fear of losing control” was the most prominent symptom to predict the frequency of AHs but not of VHs. In contrast, “derealization” was related to VHs and only to a lesser extent to auditory ones.

Dependent variables	Visual hallucinations		Auditory hallucinations	
• Overall models for VHs and AHs	BPRS severity	0.321	BPRS severity	0.370
	ESM anxiety	0.304	Pos./neg.-group	0.351
	pos./neg.-group	0.300	ESM anxiety	0.284
	ESM suspicion	0.232	ESM fear losing control	0.278
	SFCL adequacy	-0.130	SFCL adequacy	-0.278
	$\chi^2_{(17)}=42.22$; $p<0.001$		$\chi^2_{(17)}=32.59$; $p<0.015$	

Table 10.1 Stepwise discriminant analysis: subject characteristics of hallucinators (max. 5 predictors).

Discussion — Overall, the predictors for subjects that hallucinate auditi-vely or visu-ally were similar. This is understandable considering the overlap between the subjects in both conditions. More than 80% of the subjects who hallucinated auditi-vely also had visual hallucinations. Table 10.1 illustrates that no demographic characteristics were related to sampled experiences of hallucinations, but illness severity (assessed using the BPRS) and illness type (positive symptom type) was. In addition, moment-to-moment anxiety predicted subjects with high frequency of both forms of hallucina-tions. Also, moment-to-moment reports of psychopathology were predictive but diffe-rent items were included for VHs — “suspicion” — and AHs — “fear losing control”. Finally, the general level of social adaptation, measured with the SFCL, was included as the fifth predictor for both hallucination alternatives.

Hallucinating moments (hallucinations)

Concurrent mental state of hallucinatory moments

How do subjects feel when they hallucinate? Is there a difference from moments which they do not hallucinate? We assessed the difference between both samples, schizophrenic and non-schizophrenic chronic mental patients, and evaluated the effect of concurrent emotions on hallucination and non-hallucination moments using analy-

sis of variance techniques. Due to the large amount of tests, we selected salient significances using meta-analytic techniques (Hunter, 1990; Petitti, 1994).

Visual hallucinations — The moments which the subjects reported having VHs were not related to concurrent assessments of cognitions, moods, activity appreciation or assessment of characteristics of the social interactions. In contrast, the subjects rated themselves more obsessive ($F_{(1,10)}=4.08$; $p<0.07$), more “derealized” ($F_{(1,10)}=10.58$; $p<0.09$), more “fear of losing control” ($F_{(1,14)}=7.48$; $p<0.02$) and more “hearing voices” ($F_{(1,14)}=5.42$; $p<0.04$). Intriguingly, the strongest averaged ESM items to predict hallucinating subjects — “anxiety” and “suspicion” — were not significant in predicting hallucinating moments.

Auditory hallucinations — In contrast to VH, AHs were related to self-ratings of thoughts. Concurrently with “hearing voices” thoughts were less “pleasant” for schizophrenic subjects but more for non-schizophrenic control subjects ($F_{(1,30)}=3.31$; $p<0.08$). Thoughts were rated as “racing” more often during hallucinations in the non-schizophrenic mental health control subjects but no differences were found in the schizophrenic patients ($F_{(1,25)}=4.97$; $p<0.04$). Both subject groups felt more lonely during AHs ($F_{(1,30)}=19.06$; $p<0.0005$), a difference that was more pronounced for the non-schizophrenic chronic mental patients ($F_{(1,30)}=9.66$; $p<0.004$). No significant other concurrent emotions were found.

We found significant differences for the appreciation of activities, social interaction and concurrent psychopathology, but none of these passed the meta-analytic criterion that the number of significant tests should be significantly larger than the significance level that was used. With the exception of concurrent psychopathology, the differences were more pronounced for the non-schizophrenic subsample. For concurrent psychopathology all subjects experienced more “thought control” ($F_{(1,26)}=7.71$; $p<0.01$), “derealization” ($F_{(1,26)}=3.16$; $p<0.09$) and of course “visual hallucinations” ($F_{(1,14)}=7.19$; $p<0.02$). Here, too, the two items that were most prominent in the prediction of hallucinators were not important for assessing hallucinating moments (“suspicion”: $F_{(1,30)}=0.00$; n.s.; “fear of losing control”: $F_{(1,14)}=0.40$; n.s.). “Suspicion”, however, was high concurrently for the schizophrenic subjects, but low for the non-schizophrenic chronic mental control patients.

Figure 10.2 Feelings of “loneliness” (left) and “hunger” (right) during auditory hallucinations.

An unexpected finding that is difficult to understand was that in both samples the subjects rated themselves as being more “hungry” during hallucinations ($F_{(1,30)}=5.99$;

$p<0.02$), a difference that was more pronounced in the non-schizophrenic control sample ($F_{(1,30)}=3.87$; $p<0.06$).

Concurrent psychotropic substances — We also assessed the concurrent or previous consumption of psychotropic substances and related this to reports of hallucinations. Overall, a significant meta-analysis effect was found (at $\alpha=0.10$: $z=2.02$, $p<0.02$). There was less alcohol consumption during hallucinations (VH: $F_{(1,14)}=3.07$; $p<0.10$; AH: $F_{(1,29)}=8.97$; $p<0.006$ — more pronounced in the non-schizophrenic group: $F_{(1,29)}=2.94$; $p<0.10$). Equally so, but only for the AHs, we found less consumption of medication ($F_{(1,27)}=12.88$; $p<0.001$) and coffee ($F_{(1,27)}=2.76$; $p<0.11$) during or just before a hallucination, differences that were once again more pronounced in the non-schizophrenic group (medication: $F_{(1,27)}=2.85$; $p<0.10$, coffee: $F_{(1,27)}=5.64$; $p<0.03$). No significant effects were found for the use of tobacco.

Figure 10.3 Proportions of concurrent use of psychotropic substances and hallucinations.

Discussion — The variables that were salient for selecting hallucinators were different from the items that were rated high for hallucinatory moments. In the previous section we have demonstrated — restricted to sampled ESM mood ratings — that anxious subjects were also the subjects who hallucinate more often. Equally, from the psychopathology related items, “suspicion” was the most salient subject characteristic. However, neither “anxiety” nor “suspicion” predicted hallucinatory moments.

As a general rule non-schizophrenic subjects were more reactive emotionally to hallucinatory moments, while schizophrenic subjects seemed less affected. We assumed that schizophrenics were more emotionally blunted, but this was not confirmed. The moment-to-moment variability in the items was the same in both groups. Also, we expected that the non-schizophrenic chronic mental patients would suffer more from the extreme experience created by the hallucinations. However, significant interaction terms were found for “pleasant thoughts” and “suspicion”, reflecting, unexpectedly, more positive states during hallucinations. Auditive — not visual — hallucinations were less threatening for the non-schizophrenics. Non-significant trends in the same direction were found for “anxiety”, “guilt” and “angry/irritation”. In each situation the emotion was stronger for schizophrenic subjects during hallucinations but less than baseline for non-schizophrenic subjects. No such differences were found for visual hallucinations. The small number of hallucinating subjects in the non-schizophrenic sample can be held responsible for these idiosyncratic findings.

When considering the assessments of concurrent psychotic symptomatology both subgroups were equally affected. Moment-to-moment variability in hallucinatory experiences was embedded in psychotic moments. However, the patterns for visual and auditory hallucinations were not the same. In the previous section we described the difference in prevalence and moment-to-moment occurrence of both positive psychotic symptoms. Here, moments with visual hallucinations were primarily related to “obsessive thoughts”, “derealization” and “fear of losing control” while auditory hallucinations were related to “thought controlled (by others)” and “derealization”. Alienation was a basic experience during all hallucinatory moments. Thought pathology was also intrinsically related to perceptual psychotic pathology. However, VHs were more threatening (“fear of losing control”) when they occurred, while paranoia (“thoughts controlled”) characterized moments of AHs.

The most intriguing finding came from the observation that subjects reported being more hungry when hallucinating audibly — not visually. A first interpretation would be that hunger was time-related and that the frequency of hallucinations also varied during the day. However, in our analysis of the social context of psychotic experiences (Chapter 9) we did not find a within-day pattern for AHs. Hunger can also be considered a micro-stressor that triggers the hallucinatory experience. Because schizophrenic patients are poor assessors of their own (visceral) experiences (Frith, 1992), mislabeling is possible. If this hypothesis holds, other experiences reflecting tension should co-occur. We found, however, no relation between AHs and items such as “tension” or “anxiety” (see, however, the next section). A final exploration brings us to neurology and neuropsychology. Hunger is triggered by activation of nerve centers in the hypothalamus — it is stimulated by the lateral (“feeding” center) and inhibited by the medial hypothalamic nucleus (“satiation” center). AHs are traditionally associated with stimulation of the verbal cortex — Broca’s and Wernicke’s Area. The specificity of this cortical stimulation for AHs has been confirmed but also challenged in empirical studies. SPET and PET studies have demonstrated higher cerebral blood flow during AHs in a large number of cortical as well as brainstem areas including the hippocampus and amygdala (David, 1994). This brings us back to the limbic system where feeling hungry originates and from where, more generally, emotions emerge. We hypothesize that AHs, can also be considered emotions and look forward to further empirical multidisciplinary studies combining prospective monitoring (ESM) and neurologic as well as neuropsychologic assessment techniques to assess this issue.

Our final discussion point concerns the use of psychotropic substances. It is sometimes assumed that these substances are used by the patients to control psychotic symptoms. Our study showed that irrespective of the type of psychotropic substance significantly fewer substances were used prior to the hallucinatory experience. Can we conclude from this that — with the exception of the consumption of coffee in the schizophrenic subjects — alcohol and medication were effective protective factors? This result was based on prospective ESM data. We compared moments directly following the consumption of a substance with unrelated moments for each subject. The results cannot be caused by the most ill subjects taking, for instance, more medication. We can also assume that AHs led to social withdrawal, which makes psychotropic substances less available. This could explain the lower alcohol consumption prior to

reports of hallucinations. It does, not however, explain the effects related to medication because most of the subjects were on a tight medication schedule that was unrelated to moment-to-moment symptomatology. Interesting future research would use prospective monitoring (ESM) in a quasi-experimental design in which medication use and alcohol and coffee consumption is under experimental control.

Dynamic analysis of hallucinations over time

In the final set of analyses we have attempted to describe the dynamic process of the emergence, occurrence and intensity of hallucinations in daily life. What happens to our subjects emotionally and contextually before, during and after a hallucinatory experience. To do so, we defined hallucinatory moments and mapped out the last moment before, the first during, the last report still part of an hallucinatory episode and the first report after the episode. We compared these with baseline assessments. An extensive discussion on the coding of the different phases of the hallucinatory episode may be found in Appendix 5. The analyses were performed using random regression multilevel techniques (see Analysis section of Chapter 9).

Basic Model	Model	Auditory Hallucinations			Visual Hallucinations		
Intercept (fixed)	1F	11808.43	—		3938.55	—	
Intercept (random)	1R	8098.10	3710.33	***	3392.02	546.53	***
+ Hallucinatory event	1R 1F	5190.97	2907.13	***	1883.88	1508.14	***
+ Periods (before, first IN, IN, last IN, after)	1R 5F	5165.38	25.59	***	1882.99	0.89	n.s.
Mood (including interactions)	Model	Auditory Hallucinations			Visual Hallucinations		
Anxiety	1R 11F	5103.10	62.28	***	1724.65	158.34	***
Cheerful	1R 11F	5123.63	41.75	***	1783.43	99.56	***
Lonely	1R 11F	5135.31	30.07	***	1848.69	34.30	***
Satisfied	1R 11F	—	—		1871.84	11.15	*
Demotivation	1R 11F	5145.10	20.28	**	1861.25	21.74	***
Context (including interactions)	Model	Auditory Hallucinations			Visual Hallucinations		
Activities (nothing/ work/ leisure)	1R 17F	5032.14	133.24	***	1769.33	113.66	***
Persons (Alone/ Not Alone)	1R 11F	5144.68	20.70	**	1875.44	7.55	n.s.
		(Subjects: 78; Beeps: 3232)			(Subjects: 38; Beeps: 1654)		

•: p<0.10; *: p<0.05; **: p<0.01; ***: p<0.001.

Table 10.2 Multilevel random regression results for hallucinatory episodes (χ^2 -tests).

Pattern of hallucination intensity — The intensity of AHs and VHs at specific moments in time was modeled using the event (occurrence/ no occurrence) and the different phases within the episode (last before, first during, during, last during, first after). The weighting of the between-subject differences in the model had an important effect on the estimations. Traditional beep level analyses would overestimate intensity ratings as well as the shape (e.g. VHs) of the process due to the small number of subjects who hallucinate almost constantly. The multilevel random regression models control for this (see the difference between the raw data and the modeled process in Figure 10.4). As could be expected, the information that an hallucinatory moment was detected enhanced the prediction of hallucinatory intensity significantly (AH: $\chi^2_{(1)}=2907.13$, $p<0.001$; VH: $\chi^2_{(1)}=1508.14$, $p<0.001$). Overall, the intensity of VHs ($\bar{x}$ = 3.37) was lower than AHs ($\bar{x}$ = 4.21). In contrast, differences between the phases of the hallucinatory episode were marginal. Allowing differential weights for the phases of the episode added to the prediction of AH ($\chi^2_{(4)}=25.59$, $p<0.001$) but not VH-inten-

sity ($\chi^2_{(4)}=0.89$, n.s.). A marginally lower intensity of AHs was observed for the first report of the episode. The intensity reached a peak during the episode but dropped at the end. In contrast, the intensity of modeled VHs was unaffected by the episode's phase. The models of Figure 10.4 are the starting models on which we built our next analyses. Although we will not reiterate this constantly we included only the variables in the equations that resulted in a significantly better prediction of the empirical data.

Figure 10.4 Models of auditory and visual hallucination episodes: raw data (triangle-lines with confidence intervals) compared to multilevel random regression solutions including differential estimates of the phases within the episode (circles).

Concurrent emotions — Emotions have a marginal place in the diagnosis of schizophrenia. It is often assumed that emotional reactivity is non-existent and that subjects suffering from schizophrenia have a “blunted” or “inappropriate” affect. In Table 10.2 we can see that concurrent emotions — such as “anxiety”, “cheerfulness”, “loneliness”, “satisfaction” and “demotivation” — have a significant and differential role in the enhanced prediction of the hallucinatory intensity — but not their occurrence as we have discussed in the previous section. The addition of different emotions resulted in enhanced predictive power for both AH and VHs — more pronounced for VHs. In both situations the prediction based on “anxiety” resulted in the best fit of the original data. Figure 10.5 shows the multilevel random regression solution for “anxiety”.

- *AHs* — The addition of anxiety did not alter the shape of the curves for AHs over the different phases of the hallucinatory experience but its influence was nevertheless significant ($\chi^2_{(1)}=14.28$, $p<0.001$). Moreover, when we allowed a differential estimate of the interaction of anxiety with different phases of the AH episode the effect was very significant ($\chi^2_{(5)}=48.00$, $p<0.001$). This final model had an R^2 of 0.886. Intriguingly, higher levels of AH intensity were predicted for the lowest levels of anxiety *during* the episodes. Also of interest was the anticipatory anxiety before the first report of an AH. The elevation was more pronounced for high anxiety and was intensity related. It showed that AHs were often preceded by elevated anxiety levels.

- VHs** — Looking at VHs we saw a different pattern. The addition of anxiety resulted in a marginally better prediction of the shape of the curves over the different phases of the AH experience ($\chi^2_{(1)}=6.39$, $p<0.05$) and a dramatic effect when we allowed a differential influence of anxiety intensity for the different phases of the VH episode ($\chi^2_{(5)}=151.95$, $p<0.001$). While AHs were only marginally related to anxiety, VH intensity reflected extreme variability when accounting for the intensity of anxiety over the different phases of an episode. The relation to anxiety intensity was very pronounced at the first beep of the hallucination episode, less at intermediate beeps and most extremely at the last beep of a hallucination period. As could be expected considering the dynamic pattern in Figure 10.2, the final model for VH had an R^2 of 0.751, slightly lower than for AH. In contrast to AH, we found no anticipatory anxiety for VH.

Figure 10.5 Auditory and visual hallucinations modeled over time: prediction of hallucination intensity by anxiety levels.

Concurrent context — The influence of environmental cues on schizophrenic symptomatology is unclear from a theoretical point of view. Some authors assume that schizophrenics are unresponsive to environmental characteristics because they are “blunted” to everything. Other theories — such the frontal lobe dysfunction hypothesis — predict that complex environments are more confusing for these patients and will lead to more symptomatology. Yet others assume that restricted, concentration stimulating situations will minimize symptomatology. As a consequence, different environments should result in different levels of symptomatology. In the next set of analyses we assessed the influence of “being alone” and different activities (doing nothing, work and leisure) on the intensity of hallucinatory experiences in daily life.

- SOCIAL CONTEXT** — Social context did not result in differences in the intensity of hallucinatory experiences (AH: $\chi^2_{(1)}=3.534$, $p<0.07$; VH: $\chi^2_{(1)}=0.532$, n.s.). The differential effect of social context on the prediction of the model during hallucinations was small for AHs ($\chi^2_{(1)}=15.14$, $p<0.001$ — $\Delta=0.11$ on a 7-point Likert scale) and non-significant for VHs ($\chi^2_{(1)}=1.93$, n.s. — $\Delta=0.05$). Traditional beep-level analyses (subgraph with error bars in Figure 10.6a) showed differences but

this was due to the lack of control for the between subject variability in the data. When we included the interaction terms between social context and phases of the hallucinatory episodes we found a significantly better estimation of the original data (AH: $\chi^2_{(1)}=20.70$, $p<0.001$ — $R^2 = 0.885$; VH: $\chi^2_{(1)}=7.55$, $p<0.01$ — $R^2 = 0.726$). The absolute effect was marginal but the data indicated — for both AH and VHs — that while hallucinatory intensity was higher at the beginning of an episode when a subject was alone, the intensity ratings switched at the last hallucinatory report of an episode. This indicated that when subjects remained in a social situation, the hallucinatory intensity tended to remain the same (even rise for VH) over time, while in a non-social environment the intensity dropped. Overall and in our sample — and although its effect was marginal — withdrawal seemed to be an effective coping strategy for hallucinations in schizophrenia.

Figure 10.6a Auditory and visual hallucinations modeled over time: prediction of hallucination intensity by social context (dashed graph is the raw data plot).

- ACTIVITIES** — Doing nothing, working or being involved in leisure activities, did not result in intensity differences for AHs ($\chi^2_{(2)}=3.744$, n.s.) but a marginal effect was found for VHs ($\chi^2_{(2)}=6.242$, $p<0.05$). When we allowed a differential effect of activities during the hallucination episode the estimation of the models improved significantly (AH: $\chi^2_{(4)}=90.254$, $p<0.001$; VH: $\chi^2_{(4)}=81.346$, $p<0.001$). Subjects who were not engaged in an activity (AH: $\Delta= +0.311$; VH: $\Delta= +0.291$) and to a lesser extent during leisure moments including TV watching (AH: $\Delta= +0.180$; VH: $\Delta= +0.079$) hallucinated more intensively. In contrast, during work the hallucination intensity was lower (AH: $\Delta= -0.492$; VH: $\Delta= +0.370$). In all situations the effects were smaller for VHs compared to AHs but the overall intensity of VHs was also lower. In a next analysis we assessed whether the activities had a differential effect on the phases of a hallucinatory episode. The overall improvement over the previous model was significant (AH: $\chi^2_{(8)}=42.83$, $p<0.001$; VH: $\chi^2_{(8)}=32.32$,

$p<0.001$). The resulting final models had an $R^2 = 0.889$ for AHs and $R^2 = 0.741$ for VHs. Once again the fit of the model for AHs was higher than for VHs. The effects were complex and different from the findings derived from the raw data (error bar plot in Figure 10.6b). We nevertheless distilled general trends. When subjects did nothing the intensity of hallucinations dropped slightly over the phases of the episode (more pronounced for VHs than for AHs). Under leisure activity conditions the overall intensity of hallucinations was similar to the “doing nothing” condition but, and in contrast to idling, over the phases of the episode the intensity worsened. Finally, when subjects were working the evolution of hallucinatory intensity showed a clear decrease for AHs and a less pronounced effect for VHs. For VHs, the assessments in the middle of a hallucinatory episode were not affected by the type of activity — a finding that we also observed for the concurrent intensity of anxiety in this situation.

Figure 10.6b Auditory and visual hallucinations modeled over time: prediction of hallucination intensity by activity (dashed graph is the raw data plot).

DISCUSSION

Hallucinations in daily life were most often observed in severely ill chronic schizophrenic patients and in a much lower frequency in the chronic depressive subjects, who also were less ill. Unexpectedly, compared to AHs more subjects had at least one moment of VHs during the sampling week. The overlap between both was more than 80%. Those figures can be caused by our rather unique assessment strategy which makes it difficult to compare our data with available prevalence data in the literature. The non-schizophrenic subjects who hallucinated were comparable to their schizophrenic counterparts but their number was too low to assess them separately. There were, however, indications that the non-schizophrenic hallucinators were more emo-

tionally reactive and responded differently to psychotropic drugs. More in-depth studies of different diagnostic samples of hallucinating subjects, such as patients with dissociative states and manic depression, are planned in the future.

We have computed that an hallucinatory episode lasts approximately 2 hours and 30 minutes for VHs and 1 hour longer for AHs. Hallucinating subjects hallucinated on almost one out of three beeps, slightly less for VHs. Patients who hallucinated more often were also more anxious overall, but within-subject comparisons of hallucinatory and non-hallucinatory moments did not show differences for concurrent "anxiety". For "loneliness", in contrast, both lonely subjects were hallucinators more often in this sample and lonely moments were related to a higher likelihood of hallucinations in daily life. Important differences were found for concurrent mental states of AH and VHs. AHs were often related to negative thoughts and depressive states, while VHs were experienced as more threatening.

The assessment of concurrent contextual cues during hallucinatory experiences resulted in different intensity estimates for the phases of the hallucinatory episode. In these current analyses we could demonstrate that most emotional items, except for satisfaction, were related to hallucinatory severity — more for VHs than for AHs — and differentially affected the phase-related intensity levels. The most prominent item was "anxiety". Overall, the prediction of hallucinatory episodes using antecedent conditions was poor. None of the contextual variables was elevated in the "last beep before" condition. Only for "anxiety" did we observe an anticipatory intensity raise above baseline levels (AHs). In contrast to the prediction of occurrence, the contextual influences on the course of an hallucinatory episode were important, both from an internal ("anxiety" in VHs) and from an external point of view (being alone and being involved in different activities for AH and VHs). This seems to indicate that the occurrence of AH and VHs is less under the subject's control, but that during the hallucination being engaged in different situations can alter the overall experience of intensity. Social withdrawal results in a decrease of hallucinatory intensity (AH>VH), while being in social environments slightly raises intensity levels (VH>AH). Even so, doing nothing (VH>AH) and, even more strongly, being engaged in work activities (AH>VH) led to decreases in intensity levels over time, while leisure activities resulted in increases in intensity levels of hallucinations (AH>VH). These findings are indicative of possible successful coping strategies. Doing nothing and being alone are non-complex/non focused situations. Work might be complex or not, but always demands concentration. Leisure activities, primarily watching TV, have a less typical characterization. We believe that in daily life both maximal engagement (as in work) and maximal disengagement (as in being alone and/or doing nothing) are coping situations for hallucinatory intensity. Being in the company of other persons or engaging in leisure activities is not.

Slade & Bentall Five Factor Model

Slade & Bentall (Slade, 1976; Slade & Bentall, 1988) have presented a model for AHs. Their central argument was that (1) stressful events lead to a (2) disturbed mood (internal arousal) that is evaluated against a predisposition resulting — or not — in suffi-

cient power to raise the hallucinatory tendency above a critical threshold. In this study we have not assessed the effect of stressful events in the antecedents of hallucinatory experiences⁶¹. We have, however, assessed the antecedent and concurrent emotions of hallucinatory experiences and found that raised anxiety (internal arousal) was found in the antecedence of a hallucinatory report. Our intriguing finding relating to “hunger” feelings in non-schizophrenic hallucinators can also be such a mislabeled internal arousal. If even “hunger” generates such an arousal in schizophrenic subjects, the pervasiveness of the illness can be understood. As a general indication we have confirmed the internal arousal hypothesis of the Slade & Bentall model. In addition we found some trends that the authors did not account for: the fact that our findings were more pronounced for visual compared to auditory hallucinations and for non-schizophrenic compared schizophrenic subjects. Also, our data show that mind-altering psychotropic substances — with the exception of coffee in schizophrenics — did not lead to hallucinatory experiences. The theory would have predicted more hallucinations in those situations.

Slade and Bentall also argued in their theory that hallucinations can only gain access to consciousness at the expense of individuals attending to external sources of stimulation. In our data we see that environmental cues were unimportant for the occurrence of hallucinations, but not for their sustained presence. We found that “work” was the most powerful coping strategy, leading to dramatic reductions in hallucinatory intensity over time. This is, of course, in line with the model presented by Slade & Bentall. In contrast to this, however, “doing nothing” and “withdrawal” also resulted in a decrease in hallucination intensity over time, while leisure activities did not. We think that in contradiction to the prediction of the Slade & Bentall model, but in line with the expectation from a frontal lobe dysfunction model, unfocused and unconcentrated situations can equally reduce hallucinatory intensity.

Another element from the Slade & Bentall model — (4) the “limited capacity channel” of consciousness — could not be assessed. In contrast, something can be said about the final step in the model — (5) reinforcement through mood state reduction. The actual reinforcing effect is something that evolves over longer periods of time and could not be assessed in this study. We were, however, able to assess the emotions after the hallucinatory experience. Our data indicate that anxiety is raised before the first reports of AH and returns to baseline when no hallucinations are reported anymore. This is, of course, speculative but when we assume, as the authors do, that anxiety reduction is rewarding, we might anticipate a positive feedback loop from the positive emotions leading to raising frequencies of hallucinatory experiences over time. Considering that schizophrenic subjects are poor raters of the effect of stressors on their mental state, hallucinations — being a strategy to give meaning to unlabeled arousal states — can result in a cathartic effect that reduces negative emotional states.

⁶¹ The current data collection process includes the assessment of this data in the beep level ESFs, but the number of cases that we had available was too small to make reliable assessments.

Individualized analysis

The trends presented in this study can create the illusion that schizophrenic patients experience hallucinations irrespective of the actual individual or specific moments in their life. Of course this is not true. Important between- and within subject variability exists in the data. The random regression multilevel analysis controls for those between-subject differences and, as a consequence, creates the “illusion” that a large amount of variability was explained by the model, while in fact, the random factor for the intercept makes up an important part of this variance. In the data in Figure 10.5 — modeling hallucinatory phases using anxiety levels — the explained variance for AH improved only from $R^2 = 0.72$ when we included estimated individual subject means (a random intercept) to 0.88 for the final model. The model-related gain for VH was more pronounced from $R^2 = 0.33$ to 0.73. With three quarters to one third of the variance explained by the between-subject differences, we urge the reader to rely more on $n=1$ assessments, especially when the data are used for clinical decision making related to a specific subject.

Figure 10.7 Auditory hallucinations by social context: assessments by subject ($n=24$).

Figure 10.7 illustrates this point. We (Delespaul & deVries, 1994) assessed the influence of concurrent social context on AHs. This relation was not significant ($F_{(46,1)} = 0.91$, n.s.) and consequently we concluded that the social context is unimportant to understand concurrent hallucinations. However, when we left out the subjects who showed no variability on AHs ($n=23$; 49%), only 17% ($n=4$) of the residual subsample had no difference in AH intensity in both social conditions. Half of the sample had more intense hallucinations when alone, while 33%, in contrast, had more AH when in the company of other persons. As a consequence, the validity of our modeled relations, while important for understanding the illness process, and even explaining an important part of the variance ($R^2 = 0.88$), fits only a small subgroup of subjects (“member validity”, Mehan & Woods, 1975). We therefore advise using individualized assessments when drawing clinical conclusions using ESM results. Of course, enough data is often not available on an individual subject to allow the nuanced

results we found in this study. However, for clinical decision making less reliable data can be used because the interventions can be evaluated and conclusions adapted in a second assessment period (Chapter 12).

Chapter 11: Assessing “Quality of Life” in Schizophrenia ⁶²

Information about the differential effects of interventions in psychiatry are greatly needed in an era of decreasing medical budgets and increased emphasis on targeting financial resources to interventions with proven effectiveness. In order to account for the legacy of what psychiatry has developed as an expertise over the previous century we have to defend ourselves against the extreme utilitarian arguments that come with the wave of new interventions. We do not want to argue against the assumption that treatments in psychiatry should be selected on empirical arguments and expected effects should be weighed against costs, for resources are limited and treatment choices should be optimized. We do, however, warn against the tendency to limit interventions that are paid for by insurance companies to those with the highest group effectiveness. What is effective anyway?

The intervention that is effective for the majority of subjects is not the only effective intervention. The treatment of choice for large subgroups of patients can be different from the one for the group as a whole. Unfortunately, the necessary information for making differential treatment choices is only sparsely available. Because the most successful patients for different interventions are an overlapping group, the subjects that might gain the most from alternative, less universally successful, treatment options are the non-responders. They quickly outrun the treatment options available in their insurance packages and only the subjects whose families can afford to pay for less straightforward solutions will receive further treatment.

Except for the manifest ethical problems related to a mental health system that only takes care of riskless patients, the basic research data that would warrant such a strategy do not exist. First, our current diagnostic systems do not allow us to make reliable predictions of individual therapeutic successes. Popper argued that given the degree of accuracy required in our prediction, we should be able to specify the degree of accuracy we need in our initial conditions (Popper, 1956). Kellert (1993) wrote a valuable treatise on predictions in complex systems. He defined “predictive hopelessness” as a condition in which even the best possible prediction fails to provide any information about the behavior we are seeking to predict: the predictions are random and meaningless. The “predictively worthwhile time” tells us the time it takes for the process to reach predictive hopelessness. How far in the future can we project our prediction? He concludes that most systems require impossible accuracy in the initial conditions for useful prediction. Prigogine (1980) defined principles to enhance our predictive power. He stated that the actual state of a system should always be understood from its past trajectory. Furthermore, predictions can be enhanced by continually including

⁶² This chapter entitled “Measuring Quality of life using Sampled Experience” was presented in collaboration with Marten deVries at the VIIth European Meeting of Social Psychiatry and Psychiatric Epidemiology in Vienna, Austria in April 1994. The research was made possible through a research grant from Solvay-Duphar.

information of the evolution of the system over time into the process. When monitored over time, the “predictively worthwhile time” stretches itself out over a longer period. Our diagnostic systems in psychiatry miss the accuracy needed for these long time predictions and although the criterion for a therapeutic success is broad, the “predictively worthwhile time” is short and the state of “predictive hopelessness” quickly reached. As a consequence, the prediction of therapeutic outcomes using the current diagnostic system is a hopeless endeavor and there is a great need in psychiatry⁶³ to enhance diagnostic accuracy.

Secondly, depending upon the purpose of a medical evaluation study the outcomes are restricted to a reduced set of indicators. The optimal treatment strategy for schizophrenia from the point of view of federal budget control is different from the one that relieves positive symptoms effectively or minimizes family burden. With highly divergent criteria, the selection of the optimal treatment on empirical grounds can easily result in inconclusiveness. This is unfortunate and does not lead to clear guidelines for decision making. We, therefore, need to standardize our evaluation protocols and agree upon the salient outcome variables. The outcome measures that are used for the evaluation of treatment interventions in psychiatry are: changes in the clinical picture, the detection of possible side effects — primarily in clinical trials of neuroleptics — and improvements in “Quality of Life” (QoL). Based on the definition of agreed upon outcome variables we can gain more information on the relative merits of protocolized interventions — traditional as well as new ones.

The costs of medical evaluation research are rising. It will therefore become increasingly difficult to collect the requested information for optimal decision making in therapy selection. We need to reconsider our study designs and assessment strategies in order to obtain this information more effectively. The available assessment techniques for subject characterization and outcome determination, such as standardized questionnaires and structured interview techniques, lack the sensitivity that is necessary for detecting the value of different intervention strategies in psychiatry at low subject cost. This is true for the evaluation of well-defined interventions such as the prescription of neuroleptics, but even more so for the evaluation of psychotherapy and social psychiatric interventions. Improvements in our assessment instrumentarium are highly needed. We have to focus on the development of assessment instruments that allow early detection of therapeutic effects and/or are sensitive enough to evaluate long term changes with a smaller number of subjects. In this chapter we will discuss the assessment of “Quality of Life”, which covers an important outcome area for the differentiation between interventions that have equal clinical outcomes.

⁶³ We do not believe that in psychiatry predictions would be possible with 100% accuracy were we able to enhance our assessments. The best we can and should do is to optimize the reliability of our assessment techniques, which will optimize our predictions.

DEFINING "QUALITY OF LIFE"

WHO definition

The World Health Organization (WHO) definition of health is the source of outcome assessments in medical treatment evaluation research beyond the mere reduction of symptomatology. Originally in 1948 and again in the Charter of Alma Ata in 1978 the WHO defined its goals as the "promotion of the attainment of the highest possible health for all people". Health was defined positively — the maximal attainment of somatic, mental and social well-being — not merely as the absence of illness. According to the WHO the absence of illness (negative definition) is a prerequisite for the attainment of general well-being (positive definition) and is, therefore, assumed in the WHO definition of "health".

We think that the WHO definition of health is valuable and innovative in somatic medicine but problematic in mental health. For somatic illness "health" means primarily the absence of somatic symptoms. In addition to the somatic treatment objective the WHO definition of health requests also mental and social well-being, a relevant addition. In mental health, however, the "illness" is the psychiatric symptomatology. Cure is the alleviation of those mental health symptoms. Somatic health, although often underestimated, is less of an issue and when relevant a prerequisite for mental health (Honig, 1989). Social well-being is, therefore, the only addition to fulfill the WHO definition of health in the field of mental health. We may know how to assess somatic and mental illness and the absence of it, but we do not know how to measure social well-being (Guadagnoli & Mor, 1990). As a consequence the assessment of "Quality of Life" — being the unmatched part of the traditional (negative) definitions of health — overlaps with "depression" or psychologic well-being in somatic research, but its specificity is not well-defined in mental health.

The WHO definition of "Quality of Life" is a political statement that kicked off the growing interest in the concept but played a marginal role in its actual operationalization. The pharmaceutical industry has popularized the concept, because it yielded economic incentives. They need to define and propagate new outcome measures that allow differentiation between competing compounds that are marketed with higher consumer costs without real difference in clinical effectiveness.

Current definition of QoL

In contrast to the political WHO definition of QoL, current operationalizations emphasize the day-to-day comings and goings of a free individual. Spilker (1990) argues that QoL must be viewed on a number of levels. A first, general level can be defined as "an individual's overall satisfaction with life and one's general sense of personal well-being" (Shumaker et al., 1990). It is referred to as Global Clinical Impression and is, for instance, assessed by thermometer scales. In the field of medical evaluation research QoL represents the functional effect of an illness and its consequent therapy upon a patient, as perceived by the patient. A second level of QoL divides the global impression in broad domains that contribute to the overall effect. Most of the work on

general assessments of QoL has focused on those domains, whose number and specification differ from author to author. The set according to Schipper et al. (1990) is "physical and occupational function", "psychological state", "social interaction" and "somatic sensation". Spilker (1990) skipped "somatic sensation" (pain) from his "domains of QoL"-list and added "economic status and functioning". The third and last level of QoL contains the disease or function specific operationalizations (test and scales) within each domain which were defined in a general or aspecific way at level 2. The diversity of measurement alternatives of QoL for specific domains, diseases and functions at the three different levels makes it very unlikely that QoL can be assessed as a single concept. The conclusion "subjects of condition A had a better QoL" should therefore be made and appreciated with caution.

QoL: objective or subjective concept?

An intuitive approach to a general QoL concept such as the one described above is meaningful and appealing. It relates to a condition of well-being, feeling happy and satisfied. The problems experienced in the operationalization of the concept are in contradiction to this intuitive meaning that almost automatically emerges in everyone when thinking about QoL. The reason, we think, is that those operationalizations try to objectify what is intrinsically a subjective experience. There is no external objective reference for "quality of life". QoL is subjective — it is what a subject calls his or her QoL. The intuitive concept of QoL matches the first level described by Spilker (1990) and is best measured using rather unsophisticated assessment tools such as a visual analog scale or a so-called thermometer.

A problem is that such a parameter is not sensitive enough to be used in clinical research that compares between individuals or within individuals at different moments in time. Although fluctuations in mood, thought and behavior are of prime concern in psychiatry, most psychiatric and traditional QoL assessments ignore momentary or state dependent changes in mental status and attempt to measure stable psychologic traits and global well-being. Attempts to create objective indices are paramount.

Daily life experience and QoL

In a review of QoL assessment instruments Lehman & Burns (1990) assert that we have no general agreed upon standard in the field of psychiatry and chronic mental health. They argue that apart from the necessity to standardize the available assessment strategies, basic work on the concept itself is highly needed. One of these conceptual problems is the time frame for QoL assessments: is QoL a retrospective global appreciation of well-being by the subject with reference to the previous weeks or months, or a concept that can be assessed by moment-to-moment, prospective experiences? Another such problem is whether the building blocks of QoL are the daily life experiences of subjects or, in contrast, the general feeling that subjects have about their life. In other words, can we feel depressed from moment-to-moment and still be satisfied with our life? We assert that it is in the daily life contact and interaction of a subject with his environment that psychopathology occurs. It is even more in the context of the subject's daily life that the impairments by the illness are experienced.

Therefore we believe that the assessment of daily life functioning must be central in the assessment of QoL of an individual. Of great potential relevance for QoL are the daily life experiences of positive mental states, rewarding situations and the satisfying use of productive and leisure time (see Grønmo, 1982). In its most elementary form, QoL is related to the ability of subjects to select and perform those activities that are most valued by them, resulting in the experience of optimal mental state when performed. The principal aim of the research is to identify measures that validly and reliably evaluate quality of life aspects of patients' actual experience as it is measured from moment-to-moment in the natural flow of daily life. This has traditionally presented methodologic problems. In this study we will use the Experience Sampling Method (ESM) to develop and exploratively evaluate new, concept-driven, QoL operationalizations based on aggregations of sampled momentary functioning and experiencing.

METHOD

Patients

This study is performed on a sample of 30 chronic schizophrenic subjects, 20 men and 10 women, aged between 21 and 64 (mean 36.6 ± 10.3 years). 43% had a regular occupation, including household chores and sheltered work, but only one subject had a regular job. 13% of the subjects lived alone, 7% with their own family and 37% still lived with their parents. Finally, 43% of the subjects were institutionalized at the time of the study. The sample was highly heterogeneous in illness severity as measured by the Brief Psychiatric Rating Scale (normalized BPRS 24.05 ± 12.96). All had been in care for an average of 10 years, ranging from 2 to 30 years. The social functioning was marginal (3.56 ± 1.12 on a 7-point scale ranging from complete helplessness to normal functioning).

The Experience Sampling Method (ESM)

The subjects were assessed using the Experience Sampling Method (ESM: deVries, 1987, 1992) a time sampling of daily life experience, activities and social interactions. ESM as we use it maps the daily life of subjects over the course of 1 week. A digital watch beeps randomly, 10 times a day for 6 days. The subjects fill out a questionnaire as soon as possible after the beep. Items that were repeatedly rated by the subjects were assessments of mood and psychopathology and descriptions of the place, persons present and activities performed at the moment of the beep. This takes an average of 2 minutes. Answers given after more than 15 minutes were skipped.

Understandingly, the ESM assessment is a complex endeavor and sampling chronic mental patients is not easy. A skilled psychiatric nurse introduced the method and checked on the subjects regularly, sometimes even on a daily basis. During a briefing the subjects were assessed with traditional instruments. We collected demographic data and rated the level of social functioning and illness severity. Those variables were used as global indicators of adaptation and related to ESM-QoL assessments.

ESM - operationalizations of daily life QoL

How can we conceptualize an assessment of QoL in daily life? Before presenting the selection of relevant models we should mention that we do not believe that momentary QoL exists. A subject can feel well at a specific moment in time and this moment can contribute to the subject’s QoL. An assessment of daily life QoL, however, is an aggregation over a set of moments that can be considered representative of the subject’s actual life situation. We describe some models of daily life QoL and indicate how they can be operationalized using ESM data.

“Time budget” based operationalizations of daily life QoL

- “Productivity” model: The “productivity” model of QoL assumes that subjects who engage more often in productive activities are better off. The larger this figure, the better the subject’s QoL. According to this model, an individual who is able to work has a higher QoL compared to people that are unfitted to work. Sophisticated versions of this model weight the list of possible activities on economic grounds to include non-remunerated productive activities in the formula.

$$QoL' = \frac{FREQ(Productive\ Activities)}{FREQ(Productive\ Activities)+FREQ(Unproductive\ Activities)}$$

The ESM operationalization of the “productivity” model computes the activity time budget to obtain the amount of time (% of valid observations) the subject engages in productive activities (work, study and household chores).

Figure 11.1a QoL assessment using sampled observations of productive and unproductive activities.

Prospective (ESM) assessments of sampled activities over a week do not match the retrospective self-ratings of those activities by the patient at the end of the week. Therefore, retrospective ratings are not cost-effective or reliable alternatives to ESM assessments for determining frequencies of activities in daily life. The retrospective assessments of the activity frequency for unproductive activities — brooding, naps, TV watching or listening to the radio — show an important under-rating. Chronic mental patients do these activities a lot. The proportion of time rated as productive activities will thus be overrated.

- “Diversity” model: We assume that QoL is related to the diversity in activities a subject performs, the places he frequents, or persons he interacts with. The larger this figure, the better the subject’s QoL. Concretely, a patient who has contact with a small number of persons (only family or only mental health professionals), who

is home all the time or watches TV all day long can be assumed to have a lower QoL than a comparable subject with a more diverse activity pattern.

$$'QoL' \approx \frac{\# \text{ Observed Different Contexts}}{\# \text{ Possible Contexts}}$$

Contexts: using activity, place and person assessment codes

The ESM operationalization of the “diversity” model of QoL counts the number of different kinds of contexts (activities, places and persons) a patient engages in. Small numbers are related to lower QoL. Of course this figure is related to the coding system that is used. We used the combined codes defined in Appendix 5.

Figure 11.1b QoL assessment using the spectrum of sampled contexts: low differentiation is low QoL; high differentiation is high QoL.

- “*Variability*” model: The “variability” model uses the differential pattern in time use by code. Subjects with a balanced time use pattern — who spent a balanced amount of time in different places or doing different things — are better off, compared to subjects with highly skewed time use patterns. The standard deviation in the frequencies of context codes for a specific subject is the operationalization we use in ESM studies. QoL is inversely related to time budget variability. When a subject is alone most of the time and at a specific other place for the remaining small proportion of time, the variability will be high. In the example of Figure 11.1b the Low QoL situation has a s.d.=57.6 compared to s.d.=11.3 in the High QoL condition. It is impossible to compute this QoL parameter when only one code is rated; for instance, when a subject is alone all the time. QoL will then be poor.

$$'QoL' \approx \frac{1}{SD_{FREQ(Code)}}$$

- “*Stability*” model: QoL is inversely related to the moment-to-moment stability in the activities a subject performs and the places or persons he frequents. The larger this figure, the better the subject’s QoL. For instance, patients who have contact with a small number of persons (only family or mental health professionals) are better off when they meet those people on a daily basis compared to patients who are always alone except for 1 day on which they have all contacts. When there is no moment-to-moment variability in the subject’s life we assume a lower QoL.

$$'QoL' \approx \frac{FREQ(Code[time_i] \neq Code[time_{i+1}])}{N_{[time_i / time_{i+1}] - Transitions}}$$

The “stability” operationalization of ESM-QoL is based on an assessment of change in activities or social and physical environment over time. We computed the probability of change in context by comparing the observation at one moment in time, with the next. Both could be the same situation (stability), or a transition from one context to another (temporal variability). Those figures were computed using the ESM time budget categories and codes (places, persons and activities — Appendix 5). The more variable a subject’s life is, the higher his QoL.

Figure 11.1c QoL assessment using the moment-to-moment transition in context: low change is low QoL; high change is high QoL.

In addition to the context based ESM-QoL parameters we also assessed the “diversity” and “variability” of thoughts. We believe that these are also QoL indicators.

“Experience in context” based operationalizations of daily life QoL

Almost by definition QoL assessments should assess the cognitions and emotions of a subject in daily life. The previous time budgets-related assessments of QoL are rigid and presumptive in this respect. The QoL operationalization based on the “productivity” model does not account for the experienced burden that can be related to work or other productive activities. Also, in the “diversity”, “variability” and “stability” models we have not accounted for the positive and negative feelings that are related to different activities and situations and should be weighed in an assessment of QoL. Frequently being alone is not a problem for everyone, just because it is assumed to be a problematic situation in our culture. It is a problem for those subjects who feel depressed when they are alone. This nuance of subjective evaluation is missing in the time budget assessments of daily life QoL. Two models of ESM-QoL include the concurrent mental state weights of the context and activities by the subjects.

- **“Flow” model:** Csikszentmihalyi developed a theory of optimal experience (“Flow”) which offers an interesting model for QoL (Csikszentmihalyi & Selega-Csikszentmihalyi, 1988; Csikszentmihalyi, 1991). The theory predicts that subjects pursue optimal experience by engaging preferentially in challenging activities for which they are skilled. The model predicts also specific emotions in moments defined by the perceived balance of challenge and skills at a certain time. High levels of challenge and high skills lead to “enjoyment” but coupled with low skills those highly challenging moments result in “anxiety” and “confusion”. Low challenge, coupled with high levels of skills result in “boredom” and “relaxation” but with low skills levels we see “stagnation”. According to the “Flow” model high QoL will be found in those individuals who are able to organize their lives in the

pursuit of high challenge and high skills moments and who experience the most positive emotions in those situations.

$$\text{'QoL'} \approx \text{MOOD}_{(\text{High Challenge/High Skills})} - \text{MOOD}_{(\text{Low Challenge/Low Skills})}$$

The assessment of the Flow theory is typically done using ESM ratings of the subject's own experiences and activities in the natural flow of daily life. The subjects repeatedly describe their current activity at the moment of the beep and rate the amount of challenge they perceive and the amount of skills they have for the activity. To allow comparisons between subjects the data are weighted (high scores are positive mental states) and normalized by subject⁶⁴. As a result of this normalization, positive scores reflect activities for which the perceived challenge or available skills level is above the subject's weekly average. Negative scores are below average. The empirical fact that the balance of challenge and skills organizes the emotional world of most individuals, does not imply that the pursuit of challenging and skillful situations results in an optimal mental state for all. Therefore, our ESM QoL measure does not use the frequency of activities in the High/High situation⁶⁵. QoL is assessed by computing normalized mood and pathology scores averaged over the beeps for both contrasting situations and computing the difference between them. High Challenge/High Skills situations are compared to Low Challenge/Low Skills. A positive mean difference score reflects a high QoL.

Figure 11.2a QoL-"FLOW" assessment using mean mood and pathology in contrasting situations defined on the balance of perceived challenge in and skills for the current activities: higher QoL are reflected by a positive mental state in high challenge and skills moments (upper right quadrant) compared to Low/Low moments (lower left quadrant).

⁶⁴ Items for which the standard deviation for a specific subject is lower than 0.75 are excluded from the analysis because inclusion would lead to figures with difference scores that are larger than the discrimination of the original scale.

⁶⁵ The frequency of moments in the high/high situation is contaminated by the normalization that is used to define those moments. Assuming that both items (Challenge and Skill) are normally distributed, each quadrant will contain 25% of the moments.

The subjects who were primarily used to develop the “Flow” theory were artists, mountain climbers, solo-transatlantic sailors and high achievers in schools. In many respects these subjects are different from chronic mental patients. We do not believe that chronic mental patients organize their lives in the pursuit of top experiences. The driving force towards QoL in patients with chronic mental illness is escaping boredom and negative feelings. We defined the “Escape”-index to reflect this. We also used the challenge and skills balance to define moments, but due to the skewness of the response distribution for challenge and skills in this population we used a cut-off score of $z=+0.40$. This procedure generates a large Low/Low area in which primarily additional data from the other quadrants with negative feelings — High/Low and Low/High — are included. The Low/Low situations are once again contrasted with the High/High moments and mean difference scores are used to compute the index. QoL is high when subjects feel good when they have escaped “stagnation”.

Figure 11.2b QoL-“ESCAPE” assessment using mean mood and pathology in contrasting situations defined on the balance of perceived challenge in and skills for the current activities: higher QoL are reflected by a positive mental state in high challenge and skills moments (upper right gray areas) compared to Low/Low moments (lower left gray area).

- “*High Frequency Behavior*” model: The next model is based on the “operant learning” assumption: high frequency behaviors are also preferred behaviors. In normal daily life people are able to select most of their activities. Optimal experience or high QoL will be found in those subjects who engage most often in the most rewarding activities — activities associated with positive mood states.

$$‘QoL’ \approx \frac{MOOD_{(High\ frequency\ situations)}}{MOOD_{(Low\ frequency\ situations)}}$$

The ESM QoL operationalization uses the same rationale as in the “Flow” model. Averaged normalized mood and pathology scores are computed for contrasting situations. The moments are defined using the frequency of the activities for each subject over the sampling week. Activity frequencies are rank-ordered from high

to low. A first operationalization compares the activities that the subject performs most often (up to 50% (“split-half”) or 33% (“contrasting groups”)) with a contrasting set of low frequency activities. High QoL will be found in the subjects who feel best when doing those activities they often do.

Figure 11.2c QoL assessment using mean mood and pathology difference scores in contrasting situations based on the frequency distribution of activities: a positive mental state in the more frequent activities reflects a higher QoL.

The “Experience in Context” based daily life QoL indices were computed for weighted mean mood and pathology scales, as well as for three specific items, selected because of their assumed relevance for QoL: “cheerfulness” (indicator of depression); “relaxation” (stress); and “de-motivation” to rate momentary motivation.⁶⁶

Analysis

To conduct this exploratory study on daily life QoL assessment we have defined a set of ESM operationalizations based on time use data and experience. 28 different indices of ESM QoL were generated using time budget information: the “*productivity*” model resulted in one index; the time budgets for *context* — activities, places and persons — generated ESM QoL operationalizations for the “*diversity*”, “*variability*” and “*stability*” models (3x3, a total of 9); applied to the six different coding modalities for *thoughts* (Appendix 5) the same three models yield 18 indexes. We have five operationalizations of experience in the ESM QoL indices — themes are average MOOD, average PATHOLOGY and the individual items “cheerful”, “relaxed” and “demotivated”. Using the “Flow” and “Escape”-models this results in 10 indices, while another 10 indices are generated using the high and low frequency activities with the “split-half” and “contrasting groups” models.

There is no “gold standard” to validate our daily life criteria of QoL. Furthermore, because we used available data in this exploratory study no agreed upon assessment of QoL was available for our subjects (e.g. Heinrichs et al., 1984; Bigelow et al., 1982; Lehman, 1983a,b, 1988). We will, therefore, study how cross-sectional assessments of

⁶⁶ A relevant QoL model that I will not discuss here compares the subject’s ideal activity pattern and contrasts this with his actual performance. Subjects who prefer to be with their family, should be able to do so. QoL is reflected in the realization of this aim. Subjects who have managed to organize their life in line with their goals, have a high QoL, especially when they feel good with it.

psychopathology and social functioning relate to the ESM QoL indicators. We expect that psychopathology influences social functioning, but QoL measures should reflect a relative illness-independence.

We will use a meta-analytical technique to assess the five groups of ESM-QoL measures — a total of 48 different QoL indices⁶⁷. We will assess the significance of the proportion of significant correlations against the random proportion defined by the α -error term (see Rosenthal, 1984). We realize that our ESM QoL indicators may be intercorrelated, possibly resulting in “self-generated” validity (Feldman & Lynch, 1988). We will, therefore, primarily focus on the assessment of the different dimensions in our ESM-QoL indicators using factor and cluster analyses.

RESULTS

Cross-sectional assessments of illness severity and social functioning

Our cross-sectional indicators of social functioning were highly interrelated (6/6 significant at 1%, $z=3.43$, $p<0.001$), as well as related to a general assessment of social impairment (4/4 significant at 1%, $z=3.43$, $p<0.001$). This was what we expected.

How does this social interaction relate to psychopathology? Illness severity was related to the four social functioning dimensions (3/4 significant at 5%, $z=2.02$, $p<0.022$), and inversely to the adequacy of the social adaptation ($r=-0.546$; $p<0.01$). Traditional cross-sectional indicators of illness severity and social adaptation were interrelated. Both dimensions are not independent. This is, of course, because mental illness reduces social adaptation.

Patterns in the ESM-QoL assessments

From a content point of view all our daily life QoL indices have an a priori relation to the broad concept of “Quality of Life”. We will now assess if QoL — as assessed by our ESM-QoL indices — can be considered a uni- or multidimensional concept. We will first analyze if the different operationalizations within a specific set hang together and secondly if significant differences exist between the sets.

Within set analyses

Our first subset assesses QoL using the contextual time budgets — Who?, What? and Where? The set is highly interrelated (10/36 significant at 1%, $z=3.24$, $p<0.001$). Closer inspection shows that correlations are defined more by the themes (persons, places and activities) than by the models (diversity, variability and stability): 8/9 compared

⁶⁷ The number of pairwise correlations in each meta-analysis can be computed as $\binom{n}{k} = \frac{n!}{(n-k)!k!}$, with n = the number of variables and $k=2$ (for pairs) the equation becomes $n(n-1)/2$. For instance, with 3 themes and 3 QoL models the total number of variables is 9 and the amount of correlations to be tested becomes 36.

to 6/27 ($z = 2.70$, $p < 0.003$). Comparing over themes between models we observed a trend for “diversity” which yields the most homogeneous QoL indicators (2/3 significant at 5%; $z = 1.58$; $p < 0.06$). Subjects with a differentiated social network also have diversified activity patterns and frequent more places. No significance was found for the “variability” (1/3 at 5%; n.s.) and “stability” (0/3 at 5%; n.s.) models.

For the thought assessments — 6 themes and 3 models — we see the same pattern. The correlations within (18/18 at 5%, $z = 5.71$, $p < 0.001$) and between the themes (40/135 at 5%, $z = 5.35$, $p < 0.001$) were significant, but more so within each theme than between ($z = 12.11$, $p < 0.001$). In contrast to the assessment of QoL using context time budgets, the assessment of thought-related QoL with the “diversity” model yields no significance (2/15, n.s.), but the “variability” (9/15, $z = 3.22$, $p < 0.001$) and “stability” (4/15, $z = 1.63$, $p < 0.05$) models do.

“Flow” theory-related indices of QoL were homogeneous (13/45 at 1%, $z = 3.71$, $p < 0.001$). The “Flow”- and “Escape”-QoL indices were correlated for MOOD ($r = 0.824$, $p < 0.01$) but not for psycho-PATHOLOGY ($r = 0.349$, n.s.). The MOOD and PATHOLOGY indices were more correlated for “Flow” than for “Escape”-QoL ($r = 0.713$ compared to $r = 0.502$). Inspecting the individual items we unexpectedly saw that “cheerfulness” and “relaxation” were related to PATHOLOGY but not to MOOD “Flow”-QoL. “Demotivation”, on the other hand, was related to both. All three individual items were related to MOOD “Escape”-QoL, but only “demotivation” was related to PATHOLOGY “Escape”-QoL. The findings support our assumption that “Flow”- and “Escape”-QoL operationalizations are different, especially for psychopathology-related QoL. Moreover, the specific patterns found for individual items warn us against an uncritical use of averaged mood and pathology scales to assess QoL.

The last set of QoL indices is based on the assumption that the best mental state should be observed during high frequency activities (“HF-Activity”-QoL). The split-half (50/50%) and contrasted group (33/33%) operationalizations are highly correlated (5/5 at 1%, $z = 3.13$, $p < 0.001$). Therefore, we will restrict the analyses to the split-half “HF-Activity”-QoL. Five out of 10 correlations were significant at 1% ($z = 2.51$, $p < 0.006$). MOOD and PATHOLOGY “HF-Activity”-QoL were interrelated ($r = 0.507$, $p < 0.05$) but, when divided at the item level, “cheerfulness”, “relaxation” and “demotivation” were related to MOOD- but not to PATHOLOGY “HF-Activity”-QoL.

Between set analyses

Not one of the other sets of ESM-QoL was significantly related to the index based on the “productivity” model. Equally so, both the “Flow/Escape”-QoL and the “HF-Activity”-QoL were independent of each other and other sets of ESM-QoL indices.

The meta-analysis of both sets of time budget-related QoL indices — of context (places, persons and activities) and thoughts — showed that both were interrelated (20/162 at 1%, $z = 4.09$, $p < 0.001$). Because the same computational rules were used in both sets of indices, we assessed whether the correlations between indices computed with more similar procedures (the “diversity”, “variability” and “stability” models) were higher. This was not the case. The method specific correlations for “diversity” (2/18 at 1%, n.s.) and “variability” (0/18 at 1%, n.s.) were not significant. In contrast, “stability” (3/18 at 1%, $z = 1.66$, $p < 0.049$) was. Overall and unexpectedly we found a trend

favoring non-method-specific correlations ($5/54 < 15/108$, $z=-1.106$, $p<0.14$). The significance between context and thought related QoL indices is therefore not spurious. Subdividing the set of correlations shows that the highest relations were found in subsets defined by content irrespective of the computational model. “Activity” related QoL was related to QoL indices based on the thought themes “on-line” and “content” ($13/18$ at 1% , $z=4.44$, $p<0.001$) — significantly higher than baseline ($z=11.74$, $p<0.0001$). In contrast, the QoL indices based on the other thought themes (“goofs”, “relation”, “time” and “pos./neg. value”) are related to “places” ($7/36$ at 1% , $z=2.58$, $p<0.005$), also significantly more than baseline ($z=2.04$, $p<0.021$). It is difficult to interpret these results but it is definitely intriguing that all significant correlations were exclusively found in those two subsets.

ESM-QoL indices and assessments of illness severity and social functioning

Lacking a “gold standard” as well as cross-sectional assessments of QoL, how does ESM-QoL relate to illness severity and social functioning?

Cross-sectional illness severity and ESM-QoL

The ESM QoL indicators are independent of illness severity. The exception is that the thoughts of the most ill subjects are less “on-line” (“diversity”: $r=-0.463$; $p<0.05$) and are less differentiated in “time perspective” (“diversity”: $r=-0.526$; “variability”: $r=0.531$; and “stability”: $r=0.548$, $p<0.01$).

ESM-QoL Indices	Psychopathology	Social Functioning	
	Severity	Severity	Dimensions (1-4)
Productive Activities	1%: 0/ 1; n.s.	1%: 1/ 1; n.s.	1%: 2/ 4; n.s.
(#=1)	5%: 0/ 1; n.s.	5%: 1/ 1; n.s.	5%: 4/ 4; $z=2.69$ **
TB's Context	1%: 1/ 9; n.s.	1%: 6/ 9; $z=2.94$ **	1%: 17/36; $z=4.58$ ***
(#=9)	5%: 2/ 9; n.s.	5%: 7/ 9; $z=3.13$ ***	5%: 22/36; $z=5.06$ ***
TB's Thoughts	1%: 3/18; $z=1.66$ *	1%: 4/18; $z=1.99$ *	1%: 9/76; $z=2.73$ **
(#=18)	5%: 5/18; $z=1.85$ *	5%: 7/18; $z=2.46$ **	5%: 25/76; $z=4.39$ ***
Flow/Escape	1%: 0/10; n.s.	1%: 0/10; n.s.	1%: 0/40; n.s.
(#=10)	5%: 0/10; n.s.	5%: 0/10; n.s.	5%: 3/40; n.s.
High Freq. Activities	1%: 0/10; n.s.	1%: 0/10; n.s.	1%: 0/40; n.s.
(#=10)	5%: 0/10; n.s.	5%: 0/10; n.s.	5%: 0/40; n.s.

* $p<0.05$; ** $p<0.01$; *** $p<0.001$

Table 11.1 Meta-analysis: interrelation of ESM-QoL indices and cross-sectional assessments of illness-severity and social functioning.

Cross-sectional social functioning and ESM-QoL

The relation between the ESM-QoL indices and subject characteristics of social functioning is more pronounced. The ESM-QoL index based on “productive activities” was related to social functioning, both for global social impairment and for the specific dimensions (work, social relations and self-care). Highly significant relations were also found for the relations between contextual and thought-related time budgets for the level of social impairment and the dimensions of social functioning. In contrast, no significance was found for the ESM-QoL indices that integrate activities and their experiential appreciation (“Flow/Escape”-QoL and “HF Activities”-QoL). This might indicate that these ESM-QoL are independent of illness severity and social functioning,

as we would expect for QoL assessments. Of course, a prerequisite for this conclusion is that our indices are related to traditional assessments of QoL.

ESM-QoL assessments: an overview

From the above analyses we selected the most promising and salient ESM-QoL indices — time budget QoL models for context and thoughts, MOOD- and PATHOLOGY “HF-Activity”- and “Flow/Escape”-QoL⁶⁸. Together with the cross-sectional assessments of illness severity (BPRS) and dimensions of social functioning (SFCL) we conducted a factor analysis with varimax rotation to analyze how the different indices co-vary and relate to the cross-sectional assessments. Six factors were extracted, accounting for 83% of the variance (Table 11.2).

Factor labels	Eigenvalue	Items (first entry cross-sectional s, second entry QoL)
I Social Functioning	4.48 (22.4%)	<ul style="list-style-type: none"> • SFCL (dimensions 1-4); • TB “diversity” QoL for Persons and Places;
II Psychopathology	3.80 (19.0%)	<ul style="list-style-type: none"> • BPRS Illness Severity; • TB QoL for Activity & Thought (content/form)
III Optimal Experience MOOD	3.12 (15.6%)	<ul style="list-style-type: none"> • none • MOOD “Flow/Escape” QoL , time of thoughts
IV Optimal Organization MOOD	1.99 (9.9%)	<ul style="list-style-type: none"> • SFCL (restricted network); • MOOD “HF activity”, TB “diversity” QoL for Persons;
V Optimal Experience PATHOLOGY	1.68 (8.4%)	<ul style="list-style-type: none"> • none • MOOD and PATHOLOGY “Flow” QoL , TB “diversity” for Activities;
VI Optimal Organization PATHOLOGY	1.54 (7.7%)	<ul style="list-style-type: none"> • none • PATHOLOGY “HF activity”, “Escape” QoL and TB QoL Thoughts (pos/neg)

Table 11.2 Factor analysis of indicators of illness severity and social functioning (first bullet) and the most important ESM QoL indicators (second bullet).

Figure 11.3 Visual representation of the factor structure of the most important ESM QoL indicators.

⁶⁸ Assessments based on individual mood and pathology items showed unexpected divergence and were, therefore, even more promising than conglomerate MOOD and PATHOLOGY ESM QoL indices. However, because we have no evidence for selecting the most appropriate items, and since including them all would make it difficult to assess the effects, considering the number of subjects, we restricted this analysis to summary indices.

In this factor analysis cross-sectional subject characteristics define the first two factors and have a specific relation to daily life ESM-QoL. The first of these factors is labeled “social functioning” and is related to ESM-QoL indices of Persons and Places. The second is “illness severity” and correlates with ESM-QoL indices of Activities and Thoughts. The other factors are independent of cross-sectional subject characteristics of illness severity and social functioning and can be considered true QoL factors. The varimax rotation differentiates between factors that are primarily defined using ESM-QoL indices with mood or psychopathology items of daily life experience, and the Flow/Escapes or High Frequency Activities models of QoL. We labeled the first factor Optimal Experience (OE-MOOD and OE-PATHOLOGY) and the latter Optimal Organization (OO-MOOD and OO-PATHOLOGY). Clearly, daily life QoL indicators are not homogeneous. Figure 11.3 illustrates how ESM QoL indices are related to cross-sectional subject characteristics of illness severity and social impairment or are independent.

INDIVIDUAL CASES AND QUALITY OF LIFE

To illustrate the relevance of those daily life QoL indices for concrete subjects we present two case histories of subjects with different ESM-QoL profiles. In the appendix to this chapter we have included an overview of reference data on ESM-QoL. Knowledge of the distribution of each index can be helpful to interpret the position of an individual subject within our sample of chronic schizophrenics.

Case 1

S. is a 21-year-old, subchronic schizophrenic woman — disorganized subtype — with a co-morbid condition of cannabis abuse. She is single, has never been married and lives in a sheltered psychiatric environment. She first consulted for mental problems at the age of 16 and was admitted to a mental health hospital when she was 18. She has been hospitalized more than 5 times. Her most salient BPRS symptoms were “hostility”, “suspicion”, “unusual thought content”, “blunted affect” and “tension”. The level of social functioning was marginal (mean SFCL of 3.20).

S. was involved in productive activities 35% of the time. The assessment of QoL using ESM shows a diversity in the “persons” she frequents and “activities” she performs. Also, her thought content was differentiated — primarily involving persons and objects. We saw no thought pathology and her thoughts were clear (100%) and focused on her current activities most of the time (52%).

The “Flow”-QoL indices were computed contrasting 18 high challenging, high skills observations and 18 in the opposite condition. She felt better than average in the high/high situations (0.23). The single mood-items that were most positively rated then were “feeling active”, “concentrated”, and not “lonely”, “sad” and “guilty”. In contrast, she experienced more psychopathology in the high challenge/high skills situations (-0.31) and the feelings “unsure” and “angry/irritated” — which she selected as her idiographic complaints in the briefing session — were elevated. In addition, the item “thoughts influenced by others” was extremely high. This subject has a highly differentiated activity pattern — no single activity has a high frequency. “Sleep, nap, rest” (13%), “household chores” (13%) and “TV, radio” (11%) sum up the high frequency third; while a large set of activities from “child care” (2%), “regular meals” (2%),... up to “volunteer work” (11%), sum up the low frequency third. In contrast to the Flow QoL index results, we found no mood difference between high and low frequency activities (-0.05), but did for psychopathology (0.34).

This schizophrenic woman — although seriously ill according to cross-sectional psychopathologic ratings — has a rather high QoL. Arguments for this conclusion are the differentiation of the time budgets, the positive mood in high challenging and skill

activities and the balanced organization of her activity pattern in such a way that psychopathology only affects her in lower frequency activities. These are probably also challenging to her, resulting in more pathology but positive mood states.

Case 2

S. is a 27-year-old, chronic schizophrenic man — residual subtype. He has never been married and lives alone. He first consulted for mental problems at the age of 18 and has been admitted once to a mental health hospital. He is currently under ambulatory treatment and takes Haldol. His most salient BPRS symptoms were “anxiety”, “depression”, “suspicion” and “hallucinations”. Social functioning was high ($\bar{X}$ SFCL = 4.60). The assessment of QoL using ESM showed a “diversity” for persons and activities. The time involved in productive activities, however, was below the group mean (15%, compared to 21.4 ± 15.0). The thought contents were normally differentiated but we found a high proportion of self-referenced thoughts (52%). Thought pathology was low. He was “on-line” 78% of the time and his thoughts were always clear.

The “Flow” QoL index could not be computed due to small number of observations in the contrasting conditions. In contrast, “Escape”-QoL could. The difference was low for moods (-0.02 , population: 0.21 ± 0.28) and psychopathology (-0.08 : reference 0.07 ± 0.32). The single mood-items that were most prominent in making a difference between high/high and low/low situations were “cheerful” and “angry”. No single items were prominent for psychopathology. This subject had a highly differentiated activity pattern and no single activity category had a high frequency. “Personal mental” (15%), “household chores” (15%) and “conversation” (12%) summed up the high frequency third; while “taking a walk” (4%), “personal care” (4%),... up to “TV, radio” (8%), were combined in the low frequency third. We found a poor QoL as rated by mood scales in high and low frequency activity conditions (-0.22 : reference group 0.09 ± 0.21), as well as for psychopathology (-0.60 , compared to 0.06 ± 0.26).

In summary, this subject is suffering from schizophrenia and his general level of social adaptation is fair. However, his daily life QoL was low as we concluded from the low frequency of productive activities and the below average figures for the “Escape” and “HF-Activity” QoL indices.

CONCLUSION

We have developed QoL indices using an a priori intuitive approach to the concept. We applied those intuitive operationalizations to ESM observations of activities and experience in daily life of chronic schizophrenics. The analyses in this study as well as both case histories illustrate that the assessment of QoL is a complex endeavor. Our first conclusion is that the ESM-QoL indices are effective descriptors of the daily life functioning. Secondly and contrary to our hope and expectation, the ESM-QoL indices are not uni- but multidimensional. As a consequence, the daily life adaptational structure reflecting the first level of QoL (“Global Impression” — Spilker, 1990) cannot be assessed in a more objective and sensitive way using sampled experiences in daily life. The unidimensional assessment of QoL originates in a globalized retrospective assessment. Finally, our third conclusion is that some ESM-QoL indices are independent of illness severity and social impairment. This is important because, from a conceptual point of view, QoL reflects the subjective experience of the patient in response to the illness severity and social impairment. In our study the assessment of different ESM-QoL indices shows divergent results. We do not think this reflects an unreliability of the indicators. Because we found that the ESM QoL indices are not unidimensional, we will need to assess whether different indices or index combinations can be equally applied over different subpopulations. Further studies will have to explore this multi-

dimensionality and demonstrate how the ESM-QoL indicators predict real life differences in the appreciation of daily life functioning. Furthermore, more research should be conducted to optimize the operationalization of salient ESM-QoL indices. We are especially interested in the selection of relevant emotions to include in the equations, since we have to consider the divergence found in the relation between specific and global mood-related QoL assessments.

APPENDIX: PRELIMINARY REFERENCE INFORMATION OF ESM-QOL

Item	$\bar{x}$	s.d.	Min	Max	Distribution (Model)
<i>Time budgets: context</i>					
Who?	4.90	1.56	3.00	8.00
What?	11.47	2.99	1.00	14.00
Where?	3.83	1.80	1.00	7.00
Productive activities) (work, study, household chores)	21.39	15.04	0.00	57.13
<i>Time budgets: thoughts</i>					
Content categories	13.57	3.03	7.00	20.00
Goofs	2.33	1.40	1.00	6.00
Relation	4.73	1.08	3.00	7.00
Time	2.97	0.62	2.00	4.00
Positive/ negative formulation	2.47	0.73	1.00	3.00
<i>Flow</i>					
Flow[Mood]	0.28	0.41	-0.69	1.04
Flow[Pathology]	-0.19	0.94	-3.55	0.80
Escape[Mood]	0.21	0.28	-0.35	0.79
Escape[Pathology]	0.07	0.32	-0.36	0.69
<i>High frequency behaviors (HFTB)</i>					
Mood	0.09	0.21	-0.35	0.61
Pathology	0.06	0.26	-0.54	0.52

Table 11A.1 Basic statistics of the time budget parameters (n=30 chronic schizophrenics).

Part 4: Experience Sampling in Daily Clinical Practice

.....

Chapter 12: ESM in psychiatric rehabilitation ⁶⁹

Our ultimate goal in psychiatry is to help suffering individuals. How can ESM contribute to the treatments in mental health care? Several authors have applied ESM as an assessment strategy in clinical situations or added it to therapy as a self-help strategy⁷⁰. These applications cover a range of activities from patient empowerment to case management strategies and dyadic therapeutic approaches. All require a different level of intensity and involvement of both the individual and the therapist.

The most important contribution that we can expect from sampling daily life experience is that it generates useful and treatment-relevant information that was not available to mental health professionals through the use of traditional assessment instruments such as clinical interviews, observations and standardized questionnaires. A case⁷¹ illustrates the power of daily life experience for treatment.

Jan is a 42-year-old man, who had lived for almost 20 years on a residential ward of the Maastricht local mental hospital. He is a paranoid schizophrenic, who had not been symptom-free for the last 10 years. He had an oral neurolepticum at a mega-dose (100 mg. Haloperidol daily) which controlled his symptoms only marginally. Diminishing the amount of medication, even marginally (<10%), blind to himself and the nursing staff (substituting the neuroleptic liquor with an equal amount of water), resulted very quickly in a worsening of the symptomatology. Jan became restless, thought that he was responsible for peace on earth and eventually became aggressive and assaultive to other patients as a consequence of paranoid thoughts that accused fellow patients of being responsible for an emerging third world war.

We asked Jan if he could rate the intensity of his most prominent psychotic symptom — the voices he heard in his head. He agreed to rate them on a scale from 0 to 10: with 0 = no voices; 1 = the voices can be heard but are indiscernible in the background; 2-9 = the voices rise progressively in intensity; and 10 = the voices are so prominent that nothing can be done except attend to them. Observations were collected over a period of 2 weeks, for 2 hours each, between 8:00 and 22:00. In a total of 98 observations, Jan never rated a 0: he was constantly hallucinating. The highest intensity, an invalidating score of 10, was rated in 9% of the observations. No differences were found for the different days of the week, but we found differences for the hours of the day. Invalidating levels of hallucinations were found between 14:00 and 18:00. Neither Jan nor the nursing staff were aware of this daily rhythm in psychotic symptomatology. The sampling of daily life experience highlighted this trend. In this period of high symptom intensity, each day after lunch, Jan took a nap. The ratings showed that the afternoon nap could be an unconsciously adopted coping strategy. It

⁶⁹ Part of this chapter was written as a paragraph for: deVries, M. W. & Delespaul, Ph. A. E. G. (1993). Zelfobservatie en Intensieve Tijdsampling in de Psychiatrie. In C. A. L. Hoogduin, P. Schnabel, W. Vandereycken, K. van der Velden, & F. C. Verhulst (Eds.), *Jaarboek Psychiatrie en Psychotherapie* (pp. 94-108). Meppel: Van Loghum Slaterus.

⁷⁰ ESM as a self-help strategy (Donner, 1992), increasing the opportunities for optimal experience (Massimini et al., 1987), facilitating activities that achieve optimal experience (Delle Fave & Massimini, 1992), evaluating responses to therapeutic interventions for use in individualized case management (van der Poel & Delespaul, 1992), and examining the role of self-awareness in psychopathology and therapy (Figurski, 1992).

⁷¹ First published as: Delespaul, Ph. A. E. G. & deVries, M. W. (1991). Gedragsdiagnostiek met behulp van de Experience Sampling Methode. *Tijdschrift voor Klinische Psychologie*, 21(1), 45-59.

was effective 60% of the time, leading to a status-quo or drop in the symptom intensity. For 40% of the time, however, the symptom intensity remained high.

We had assumed that the mega-dose of Haloperidol would result in constant high blood concentration levels. Therefore, medication was administered in one daily dose at 17:30. Using the observations of the daily fluctuations in symptom intensity, we engaged in a renewed attempt to reduce the amount of neuroleptic medication. We shifted the administration time from 17:30 to 12:30. In addition, Jan was trained to become sensitive to his symptom fluctuations and to use his coping strategies more effectively. We advised him to reduce the environmental stimulation — by taking a nap or going for a walk — for periods of a maximum 2 hours, in response to experienced rises in symptomatology. We managed to reduce the medication by 25% in the next month and even further over the next 2 years. Jan now lives in a sheltered home in the center of the city.

Often in clinical practice, we are presented with a problem for which we have no immediate answer. In the case of Jan we were convinced that the high amount of neuroleptic drugs should be diminished. However, previous trials reducing the medication had disturbed Jan’s mental equilibrium and increased the danger. Insufficient information often leaves us in chaos and contradictory information prevents us from adequate and purposeful implementation of interventions. We are convinced that a more precise assessment of the moment-to-moment variability in mental state — understood through its relation with concurrent changes in internal and external variables — yields a rich source of information for clinical decision making. Traditional assessment strategies reduce the individual to the shared characteristics of their reference group. As a consequence, treatment relevance will be reduced to the prediction of outcome based on empirical studies of protocolized interventions. These have only marginal generability to the individual subject and the specific dynamics of his symptomatology. The assessment of daily life experience, in contrast, can result in an individualized treatment strategy. By accounting for the individual variability and the concrete contextual embedding in the subject’s daily life, the best cost-effective interventions can be designed.

ESM IN CLINICAL PRACTICE

How does ESM contribute to the process of change?

deVries (1992) argued that regardless of the specific ideologic or theoretic position of the therapist’s approach, certain common ingredients are involved in the therapeutic use of intensive monitoring techniques such as ESM:

1. The effect on the *patient*: Actively carrying out the self-evaluations from hour to hour affects the subject. As a consequence of the constant focusing on mental state and activities, a growing consciousness of behavior and experiences develops. In an individual who is motivated to change, this potential increase in self-awareness can be a powerful catalyst. For such people, the monitoring through the ESM procedure alone functions as a learning experience from which spontaneous shifts in consciousness and behavior can occur.
2. The effect on the *therapist*: In ESM new previously undisclosed information becomes available for therapy (Delespaul & Kerpentier, 1990). The clinician can take a directive role and actively advocate alterations in the patient's lifestyle, environment and social relationships, based on ESM data about the occurrence of symptoms or well-being. Simple advice may be given so that negative situations are avoided and positive situations are pursued (Delle Fave & Massimini, 1992; Chapter 13). The interpretive feedback by the clinician based on the accumulated self-reports is a valuable assessment tool for the case manager. ESM information can help organize the mental health worker's care planning, improve his or her view of the patient and focus the selection of treatment approaches and specific interventions.
3. Effect on the *therapeutic relation*: Providing feedback on the data collected with ESM involves an interpersonal process in which the patient and the therapist construct and integrate a shared view of a patient's life and mental state. The resulting activation of alternative modes of thought, behavior and emotion, potentially promotes experiential learning, which leads to psychologic change and experienced well-being, as exemplified by Donner (1992) and Delle Fave & Massimini (1992). Within the context of clinical care, the briefing and debriefing processes of ESM, and the research alliance required becomes an integral part of the therapy. The private nature of the disclosed information has a powerful effect on the therapeutic relationship between the mental health worker and the client, in terms of transference, feelings of being understood and consensus.

In the process of therapy, information functions as a videotape of daily life that the therapist and patient project and view together. The increased information may break through the high level of dependency between client and therapist. Viewing the week together fosters mutual respect and partnership. The detailed dimensions of a person's experience, visualized from the plots and frequency counts, averaged data, situational correlations, time-budgets, variability, stability and patterns in their self-reports, provide the material for a script immediately available for clinical use.

The paradox: maximizing or minimizing ESM induced change

We have argued that ESM is a useful instrument to study psychopathology and is promising in evaluating the effectiveness of psychotherapeutic interventions for groups of subjects. The repeated sampling of experience can monitor the effect of clinical interventions on variables not typically considered targets of treatment. Time-budget allocations, percentage of alone time and of time spent in social interactions, activi-

ties, work, leisure, but also the flexibility of emotional reactions, as well as the range of mental state reactions to life situations, are useful for evaluation.

We now introduce ESM as a strategy to induce change in its application in the treatment of individual patients. In ongoing therapy, ESM provides concrete milestones or endpoints for therapies that should have an effect outside the relationship between therapist and client. The window on the real world of the patient provides the therapist with a better perspective than the evaluation of the relationship with the client from the office alone. ESM introduces, more concretely than is typically done, aspects of a person's life that are often hidden from the therapist's observation. This is therapeutically beneficial and may be accommodated within a variety of theoretical frames of reference from the psychodynamic to behavioral approaches. The door is open for creative applications. One can use ESM to optimize interventions but also to manipulate the reactivity of the instrument with therapeutic purpose (inducing change).

Is the concurrent use of ESM as a therapeutic instrument as well as for evaluation problematic? Of course, if an assessment instrument changes the assessed situation — when it is reactive — it is not very useful for assessing the effects of interventions reliably. From our list of therapeutic effects of ESM, only the influence on the person — his or her possible shift in consciousness and behavior over time by doing ESM — is independent of the therapist's feedback. However, the therapeutic effects that are a consequence of the feedback by the therapist occur after the sampling period and, therefore, can not influence the quality of the data during the collection period. Only the hypotheses that are generated by the subjects themselves during the sampling period, can change their cognitions, emotions and behaviors at that moment. This occurs at the expense of reliability (Delespaul, 1992)

The reactive influence of ESM on emotions and behaviors is similar to "self-monitoring", a technique that combines evaluative potential with a therapeutic effect (Nelson, 1977; Bornstein et al., 1986). The reactive or therapeutic effects of self-monitoring are far from uniform (Nelson et al., 1980, 1982) and not all individuals benefit from its use (Zimmerman & Levitt, 1975). The literature on self-monitoring in behavior assessment gives indications of how to maximize the therapeutic effect of the self-monitoring intervention and, consequently, how to minimize reactivity when self-monitoring is used evaluatively (McFall, 1977; Nelson, 1977). Some of the salient aspects are procedure related, some are subject related. The results are not uniform, but general trends can be distilled from the empirical literature.

The accuracy of the recording is not related to higher reactivity and therefore higher therapeutic effect. Behavioral change was observed with assessments of low interrater reliability and even after the subjects stopped systematic recordings (Bornstein et al., 1978). Therefore, we conclude that attempts to enhance reliability of the self-recording do not automatically lead to more behavioral change. Reactivity can be maximized using more salient self-recording devices — which focuses the subject continually on the recording process — and by including extensive training in the briefing session and instructions during the sampling week. Behavior change can also be enhanced using obtrusive self-recording devices. Obtrusiveness is sometimes hidden in the sampling process itself. Frederiksen et al. (1975) found significantly

greater reactivity for continuous self-monitoring, compared to daily or weekly self-monitoring. Furthermore, they also demonstrated that those changes only became apparent after 3 weeks of continual monitoring in a “smoke habit”-study. Before this time, neither of the procedures created behavioral changes. The sampling process in ESM lasts typically for only 1 week, which is too short to lead to reactive behavior changes. Nelson (1977; Nelson & Hayes, 1986) suggests that reactivity occurs when the subject perceives and is sufficiently aware of engaging in the target behavior, even though the subject may not record response occurrences on the monitoring devices or does so unreliably. Thus the reactive effects of self-monitoring may partially result from a general antecedent form of control. If this is the case, the entire self-monitoring process serves as a discriminative stimulus for attending to and modifying the target behavior of interest (Bornstein et al., 1986). Contrary to self-monitoring in behavioral assessment, in ESM the target behaviors, cognitions and emotions are defocused. In ESM, the monitoring is not triggered by the occurrence of the event, but is contingent on a time schedule. Moreover, the random time-sampling procedure we use, minimizes the continuous self-awareness of the individuals, making it even less probable that the monitoring process itself will act as a discriminatory stimulus. In addition — and in contrast to traditional monitoring in behavior assessment, such as counting the number of cigarettes smoked — in ESM the target of the assessment is not well-defined. Here, we assess a multitude of cognitions, emotions and behaviors, as well as an open-ended multidimensional description of context. The random sampling procedure, as well as the instructions to report immediately after the beep, minimizes the influence of assumptions that can corrupt the quality of the reports. Reactivity, therefore, is minimized.

Is it realistic to assume that subjects are able to draw adequate conclusions from the observations of their own emotions, cognitions and behaviors and the context in which they occur, and consequently change their behavior accordingly? In real life, neither mental states nor salient environments are well-defined and as a consequence their impact can not be easily determined in the continuous flow of experience. Previously (Chapter 2) we have argued that nominal states (presence/absence) in emotions and psychopathology are generated by retrospective assessments (“I had a panic attack yesterday”) while prospective assessments yield ordinal assessments (“my anxiety reaches a specific level now”). Equally so, the salience of a context can seem clear for an external assessor (e.g. being alone/ not alone), while for a specific subject living his own life from moment-to-moment and context to context the detection of salient environmental cues is not easy, especially if we also accept that the same context does not always result in the same experience (see case 2 in Chapter 8). Kosko (1993) argued that it is virtually impossible to assess the simultaneous influence of more than one independent variable on a dependent one. Prediction becomes even worse when the true relations are non-linear (progressively more of the independent variable does not lead to more of the dependent variable) or when the influencing variables are not bivalent but fuzzy (when even under optimally controlled conditions the true relation between the independent and dependent variable is unstable, undetermined, or intrinsically probabilistic). In those real life situations, it is not realistic to expect that subjects who engage in their personal

preferred situations will irrespectively feel best then. While some subjects might be able to assess their mental state in one molar dichotomized situation contrasted to all the other ones, normally those aggregations yield poor results. If it was that easy, subjects would probably long ago have found solutions to their problems, experimented with alternative coping strategies and assessed their respective impact adequately. In summary, the complexities of daily life are mirrored in the ESM assessments and this minimizes reactivity in all subjects.

Subject variables also influence reactivity. Motivation, change expectations, and self control skills, for instance, have been described as influencing behavioral change in self-monitoring. Subjects must be motivated for the self-monitoring procedure, they should know which behavioral, cognitive or emotional changes are expected, and possess the necessary self-control skills to achieve this expected change. This holds for every subject, ranging from normal individuals to schizophrenics.

The influence of self-monitoring on psychotic symptoms such as auditory hallucination (Rutner & Bugle, 1969) and paranoid ideation (Williams, 1976) has been documented. The chances that schizophrenic subjects will be reactive to the self-monitoring procedures is, however, less likely compared to normal subjects. Frith (1992) describes how schizophrenic subjects are unable to self-monitor their behaviors, emotions and cognitions. We have demonstrated repeatedly that this deficit is not caused by an inability to recognize their own concurrent cognitions and emotions, or to describe their current behavior or context. The problem is the assimilation of contextual cues in an interpretative summary. Schizophrenic subjects are less able than normal subjects to scan concurrent environmental cues, to allow interpretations of their own mental state in relation to the context, and to aggregate memories of previous observations into an understanding of their actual cognitions, emotions and behaviors. Lacking the necessary skills to integrate the meaning of context into an adequate understanding of their own mental state, schizophrenic subjects will be less able to learn from the self-monitoring experience and to change according to their expectations.

When self-monitoring is used in therapy, we can maximize its potentiality to induce change by carefully planning its operationalization and application. One should define the target behavior as clearly as possible — preferably as present/absent. It should be presented in a questionnaire in which the behavioral assessment is the major question, possibly with the addition of a small set of scales that directly target the most salient hypothesis about the occurrence of problematic behavior. One should ascertain that the aim of the monitoring procedure is behavioral change and that the subjects is motivated to do so. Furthermore we advise using the most reactive time schedule (fixed interval, for instance, each hour) and to use a cumbersome and intrusive sampling device. In some situations one might choose to include self-instructions of coping, arising from the subject's assessment of his or her mental state.

When the focus of the use of ESM is the collection of systematized evaluations, even in clinical practice, the opposite strategy should be pursued. The reactivity of the self-monitoring procedure can be minimized through the means of a strategic implementation. We would strengthen the variability in the symptom behavior, the contextual

diversity in which it occurs, the large number of concurrent emotions. We would also avoid the subject continually monitoring himself but restrict this to the sampling moment alone, using non-intrusive signaling devices.

In the application of ESM we often choose a scientific position in which we minimize reactivity, even when behavioral change is our target, as in personalizing rehabilitative interventions. We are convinced that the gains of generating reliable evaluative information on the contextual determinants of mental states outweigh the possible benefits of change induced by the monitoring process itself, even in therapy. The strength of ESM in treatment is not its possibility to induce change directly, but the therapeutic alliance between patient and mental health professional that it helps to develop. In this contact, the therapist is not the specialist on the patient's behavior, emotions and cognitions. The subject is the expert, because the information shared in the diagnostic process originates from the patient's own life. The therapist-assessor is only an expert in aggregation; the subject's daily life dictates the contingencies.

ESM AND BEHAVIORAL ASSESSMENT ⁷²

The assessment phase is a crucial moment in each therapeutic process but is made most explicit in behavioral therapy. Over the course of the therapeutic intervention, the behavioral therapist needs valid data to chart the problematic behaviors of the patient, deficits as well as excesses, to plan interventions and to evaluate their effect (Nelson & Hayes, 1986). This process is fundamentally an idiographic⁷³ process. It applies to a specific subject with his or her specific problems at a specific moment in time. Behavioral therapists are not interested in means and global trends reflecting subject characteristics. In behavioral assessment, the contrast between the situations in which symptoms occur or do not occur is important. Averaged symptomatology irrespective of context is not interesting, but variability is. A high score on the Zung is less valuable than the knowledge of differences in cognitions and emotions in specific situations. Equally so, the illness severity reflected by the mean BPRS score is less interesting than the pattern of symptoms and the specific contexts in which high intensities of symptoms are found.

The necessary data for generating a holistic theory, topographic or functional analysis are usually collected in a clinical interview. This method can be useful for the topographic analysis that reconstructs the individual history or the genesis of symptomatology. Sometimes and sparingly — usually in an attempt to evaluate constituent elements of the holistic theory — the assessment material is enriched with results of

⁷² The argumentation presented here was originally presented as: Delespaul, Ph. A. E. G., & deVries, M. W. (1988). Experience Sampling Method: A Useful Tool in Behavioral Assessment. In *International Behavior Therapy Congress, September 1988*, . Edinburgh; and Delespaul, Ph. A. E. G., & deVries, M. W. (1991). Gedragsdiagnostiek met behulp van de Experience Sampling Methode. *Tijdschrift voor Klinische Psychologie*, 21(1), 45-59.

⁷³ De Boeck (1988) argued convincingly that the term "idiographic" was not well-chosen, because the ultimate goal is to derive general trends ("laws") that apply on a specific individual and not on a group. The terminology "idiographic" and "nomothetic" relate to the informational value of "raw data" in its unique occurrence versus "summarized laws", not to individuals in contrast to groups.

standardized questionnaires. For the functional analysis — which forms the core of the behavioral assessment and is directly related to the planning of strategies for change — clinical interviews and standardized questionnaires are not very adequate tools. The clients have to remember their feelings and behaviors, as well as the context in which they occurred. We invite them to focus their attention also on those elements that seem unimportant to them. Inevitably, the story will be incomplete and colored (Lamiell, 1981; Mischel, 1968). This unreliability is no problem in psychoanalytic therapy, since there, the personal interpretation of events forms the basis and content of therapy. In behavioral therapy, however, the personal interpretation of events is less relevant and a meticulous description of symptoms in the actual context in which they occur is crucial to plan interventions adequately.

Because retrospective methods yield less accurate data, they are unreliable and are replaced by prospective methods, such as monitoring and diaries, for the process of behavioral assessment. Both techniques have their strengths and weaknesses. In monitoring, the target behavior should be well-defined in advance. Not infrequently, we realize later on that no information was collected on other behaviors that are functionally identical to the central symptom, and that an assessment of frequency, intensity and context of those symptoms, was also necessary to plan a more purposeful intervention (Evans, 1985, 1986). Meanwhile, time elapses. Therefore, monitoring is an adequate technique to assess baseline values to evaluate interventions, but it is overly restrictive in a preliminary explorative diagnostic phase. In contrast to monitoring, diaries are open and unstructured. As a consequence, we have no control over what is written down and what is omitted. Moreover, due to the interval between the events and the report, usually at the end of the day at the best, the subjects will color their reports using their own assumptions (see Fiske, 1971). Those assumptions are not very relevant because it is due to this kind of misinterpretations that the patients were forced to seek treatment. Summarizing, diaries yield no quantitative data and cannot be used for baseline assessments, but the descriptive power of diaries makes them ideally suited in the primary, explorative phase of therapy-related assessment.

What do behavioral therapists need?

Behavioral therapists would like to be there when the subject “performs” the problematic behavior ([R]esponse), in order to describe its determinants. They want to chart the context ([S]timulus) before the target behavior occurs and to know its effect upon the environment ([C]onsequence). Finally, they want to assess the intensity of the relation between context and behavior ([o]rganism) and between behaviors and their consequences ([k]ontingency). Answers to those questions are pursued in the functional analysis (Kanfer & Phillips, 1970) and translated directly into behavioral interventions. Ideally, the behavioral diagnostic process should assess the different modalities of the problematic response (Evans, 1986): the target behavior itself (CAR — “Conditioned Avoidance Response”), the concurrent physiologic reactions (CER — “Conditioned Emotional Response”) and the cognitions (COV — “Covered Operants”). Unfortunately, not all the different modes of expression can be observed using external observational techniques. In contrast, self-assessment is the only way

to assess cognitions. Furthermore, self-assessment is the least complex method to assess emotions and physiologic reactions (e.g. palpitations of the heart) and are effective to assess behaviors with minimized reactivity in social living systems. Moreover, self-assessments allow observations in situations and places where a professional assessor is unable to go. Those gains outweigh the disadvantages of the unreliability of self-observations.

This unreliability of self-observations can further be minimized by:

- direct observation and *immediate report*: to minimize inaccuracies in the reports, the time interval between behaviors and reports should be minimized. Even intervals as short as 15 minutes already yield inaccuracies;
- observe in the *natural environment*: while it is possible to generate relevant information from the observation of behaviors in an artificial situation (e.g. in role playing), the most informative situation is the direct observation of the target behavior with its naturally occurring contingencies;
- observe *continually or repeatedly*: behaviors are embedded in behavioral chains. When we are able to observe behavioral sequences over time, antecedent moments can teach us what caused the problematic situation to emerge. Lagged observations after the target moment can show its consequences and help us understand to interpret the repeated occurrence of the target situation (Barrett et al., 1986).

ESM is a random, repeated assessment of target cognitions, emotions and behaviors that is ideally suited to meet the requirements of behavioral assessment.

ESM SOFTWARE TO ASSESS DAILY LIFE VARIABILITY

The data generated by the Experience Sampling process is complex and the assessment of the interrelations between collected bits of data is not easy. We have therefore developed the Experience Sampling Database Management System (ESDbms)⁷⁴ with a double motivation in mind. The ESDbms offers all the data management tools for the use of ESM in clinical practice. The ESDbms is also an archive of available ESM observations. ESM is a new assessment technique and no reference material is available for most of the ESM variables. The availability of this information is important to interpret the data of new subjects. Therefore, in addition to its clinical use, the ESDbms archives ESM data and allows the continual update of look-up tables of reference material on ESM variables, using the growing set of subjects. This project is called the "Basic properties of illness and well-being — Developing profiles of experience".

The ESDbms has a modular structure. It contains a design-module with the definitions of ESM questionnaires and sampling schemes; a data-module with the data archive; and an analysis-module that focusses on ESM clinical usefulness.

⁷⁴ The ESDbms program — currently at version 3.20 — has been written collaboratively by Philippe Delespaul (IPSER Foundation, Maastricht) and Peter Emonds (ECS International B.V., Maastricht) for Apple Macintosh computers with 4 Mb internal memory and a hard disk. The application uses the 4th Dimension® relational database shell.

Overview of the ESDbms modules.

Design-module

ESM questionnaires (ESFs)— The ESFs have a standard format. The researcher or clinical assessor can select questions within this structure in order to customize the questionnaire to match his own needs in the assessment of an individual subject or a group. Open-ended questions, which are coded by the research assistance (see Appendix 5), multiple choice and (7 point) Likert-type scales are used when appropriate. We advise including the next set of questions in each ESF:

- an open-ended question that paraphrases the content of the *thought* at the moment of the beep: thoughts are coded on content, form, pathology and time perspective;
- a set of Likert-type scales, selected from a list of 80 items, assesses the *mental status*: e.g. I feel cheerful, happy, angry/irritated, lonely, anxious, sick, tired. The clinician can customize his selection for each subject but we discourage the omission of specific moods such as: anxious, depressive (sad, (not) cheerful), tired, lonely and relaxed;
- a smaller set of diagnostic group related Likert scales on *psychopathology*: e.g. “I feel suspicious” and “I hear voices” in schizophrenia or “palpitations of the heart” for anxiety subjects. These psychopathology related item sets for anxiety, depression, stress, pain, migraine, schizophrenia and addiction have been developed using expert panels or focus groups of patients;
- Likert scales of *idiographic complaints*: we ask the subjects in the briefing session to formulate two of their most salient complaints and to rate the intensity of this symptomatology repeatedly in the ESM week. This serves a double purpose: it allows the assessment of aspecific as well as the most relevant symptomatology.
- assessment of *current activity*: an open-ended question describes the content and the subjective appreciation of the activity is rated on a set of Likert-type scales: e.g. I’m “active”, “skilled”, “challenged”, and “I prefer doing something else”;
- assessment of *current environment*: open-ended questions describe the place, time and the persons involved in the interaction. The subject additionally gives a personal appreciation of the social context on a set of Likert-type scales: e.g. “I like this company”, “I prefer being alone”, “we are actively involved”;
- assessment of *between beep events*
 - In a multiple choice format we present a list of *mind-altering psychotropic substances* such as coffee, tobacco, drugs and medication. Those are selected from a set of 20 relevant options. Consumption between beeps is assessed;
 - An open-ended question assesses *micro life events or hassles* and the subjects rate the effect on them — pleasant/unpleasant and elicited coping — on a set of Likert-type scales.

It is possible to restrict the item set to a very small selection. This seems to reduce the burden on the subject and to facilitate the analyses and interpretations for the assessor. Our experience in Maastricht shows, however, that:

- in contrast to open-ended questions, the reduction of the number of Likert scales does not lead to higher response rates or avoid subject drop-out;
- a larger set of items in the preliminary assessment phase allows the appreciation of unanticipated influences on the subject's mental state and, as a consequence, a richer diagnostic process;
- a large set of items, in contrast to a set that only matches the hypothesis, guarantees a more reliable — less reactive — assessment.

We therefore advise projecting a rich data collection, which allows a descriptive explorative hermeneutic diagnostic process.

Sampling — The beep that cues the repeated assessments in ESM is emitted using pre-programmed signaling devices. The ESDbms can generate, store and download schedules of beeping times in the programmable signaling devices. The archived schedules can be used to compute the delay between the beep and the actual response, and allow the selection of the reports that fall within the ESM “reliability window” — 5 minutes before and 15 minutes after the beep. The ESDbms software includes routines that allow the generation of different time schedules (see Delespaul, 1992).

The choice of the sampling schedule has consequences for the quality of the collected observational material as well as for the statistical analyses. We distinguish two different schedule types and variations on each group are possible. Random time schedules are useful for selecting a representative sample of moments. In addition, random schemes make it virtually impossible for the subject to anticipate the beep and changing behaviors accordingly becomes meaningless. As a consequence, the subject “forgets” that he is involved in a sampling process, only to be reminded of it by the sound of the beep. The second type of sampling schemes has a fixed between beep time interval. Those schedules are ideally suited for analyses of emotions, cognitions and behaviors over time. In fixed interval schedules, data collected one moment before a target event are always equally remote in time. This facilitates summarizing computations. A fixed interval beep, however, can easily be predicted by the subject and this leads to reactivity.

The parameters that can be defined in the sampling schedule generation routine of the ESDbms allow the computation of virtually every possible schedule. We should, however, avoid oversampling the subjects because of the burden this creates. As a rule of the thumb, 6 to 7 days with a sample intensity of 10 beeps each day is manageable. If the research questions warrant longer sampling periods we advise reducing the intensity to six per day — less is often not necessary, especially when a random sampling schedule is used. For research questions that need more intensive sampling schemes, for instance 20 beeps each day, the sampling should be reduced to 2 to 3 days.

The archive

The archive contains software routines to create and edit, to search and sort, to import and export and to summarize ESM datarecords. The archive of ESM data structures the raw observational material in a standardized format, irrespective of the actual questions in the questionnaires. As a consequence, the archive facilitates the study of “basic properties” of the daily life reality of illness and well-being and offers a

reference for “diagnosis related profiles” of experience. We hope to be able to re-define our nosologic systems using those basic properties and profiles in the future.

The “Personalized Rehabilitation” analyses

The analyses of ESM data from individual subjects for clinical use are less problematic than between group comparisons (see Chapter 7). All the data come from the same individual and, therefore, we do not have to worry about differential variability between the reports of different subjects. The serial or time-related dependencies in the data might still be a problem. However, because we do not pursue scientific truth but a salient hypothesis to act upon in treatment selection, this is less of a problem. The conclusions from the assessment process can be updated using a second ESM assessment period when the first effects of therapy are expected.

Standard protocol — In collaboration with local clinicians in the psychiatric hospital and the ambulatory care facilities we have defined and selected those ESM variables that are the most powerful in describing clinically relevant information. The next variables are useful for clinical decision making:

- *TIME BUDGETS*: differences in time use describe salient aspects of a culture but also of differences between psychopathologic groups where it reflects global adaptation. Time budgets illustrate the spread of the social network, the differentiation of activities and the places subjects attend. We have illustrated that the time budgets computed using sampled experience are different from those rated by mental health professionals for specific patients (Delespaul & Kerpentier, 1990);
- *MENTAL STATE DATA*: the mean scores for the different subscales reflect the severity of the illness and between subject differences in sampled experience match the results obtained from cross-sectional questionnaires (deVries et al., 1988). Sampled mental state information can, therefore, be useful to fine-tune diagnostic decisions;
- *MENTAL STATE IN CONTEXT*: the concurrent sampling of mental state and context allows the assessment of optimal or most sickening environments. Relating mental state differences to other contextual changes over time allows the assessment of coping strategies and its effectiveness (deVries & Delespaul, 1989);
- *GLOBAL ADAPTATION*: the software also generates global parameters of quality of life by combining and summarizing time use and mental state in a single figure.

The last case in the next paragraph entitled “diagnostic fuzziness” illustrates the actual state of the standard analysis protocol that is currently used for the planning of “personalized rehabilitation” strategies (see also the computer print-out in Appendix 7).

Custom queries — The life of each patient has a different history and is different. A standard protocol will never be able to map all the salient relations. A standard protocol cannot contain enough detail to match the diversity of a the mental state-environment interaction of a specific patient. Therefore, the ESDbms offers the possibility to formulate custom queries. We can evaluate, for instance, if a specific subject perceives alone time differently in the morning and in the evening. The interactive query option is necessary for clinical decision making.

DISCUSSION

Until recently it has been very difficult to collect structured and ecologically valid cognitive, emotional and behavioral observations. This made it difficult to draw conclusions on the frequency and intensity of symptoms in the natural context and, as a consequence, to learn how mental health problems emerge and evolve over time.

Microcomputer-based devices can help us to collect self-reports reliably and in the subject's own environment where the naturally occurring contingencies affect them. The miniaturized beeping devices allow the subjects to lead normal lives during the assessment period. Their lives will only briefly be interrupted to write down the reports on the ES forms. We hope that technologic developments will allow even less disturbing assessments in the near future. Microcomputer technology can further help us archive and analyze the large amounts of data that are generated in these repeated measurement strategies. This makes it possible to avoid restricting the set of working hypotheses for therapy too early in the diagnostic process. On the contrary, we are able to assess the interactions of a broad set of variables. Such a large number of variables is problematic in scientific research due to the large between subject variability involved. In clinical diagnostics, however, it generates a stimulating set of data that can result in hypotheses which will be weighted in a hermeneutic process and checked against the concurrently collected contextual information.

ESM promises interesting applications for clinical practice. Until recently, therapy related knowledge in psychiatry was acquired through diagnostic strategies that assessed subjects through a reductionistic tube. Starting with a clinical interview or a standardized questionnaire, we generate a summarizing result, reducing the wealth of information that is available on the patient. This result is then checked against our empirical or theoretic knowledge of the reference group of the subject. The strategy assumes that we can generalize the observations made in research subjects who share characteristics with the individual patient in our clinical practice. The research information is used to make predictions about the cognitions, emotions and behaviors of our clinical patients as well as about the etiology and expected outcome of their symptoms. ESM allows us to avoid those reductionistic inferences because it uses the data on the symptom behavior of individual subjects in their own environment. We need no tricks to know or predict symptomatology because the data for the individual subject are available and only marginal informational gain can be obtained from the reference subject group.

The primary objective of science is to generate laws that allow predictions. Variability within and between subjects is hereby considered noise. Computer power allows us to go beyond this traditional scientific reductionism and has led to a shift in scientific practice that originated in physics and has recently found its way into psychiatry. Observing the mental world of subjects with its naturally occurring variability is often a chaotic experience. Therefore, the regression into the safe harbor of reductionism is understandable and often necessary. Nevertheless, we should acknowledge that the

determinism that leads to general laws is often not very helpful for the understanding of the daily varying clinical world. Assessment techniques such as ESM can force us to open our eyes to the variability in the subjective experience and yield the promise of higher therapeutic relevance.

Chapter 13: Case Studies in “Personalized Rehabilitation”

In this chapter we will describe how ESM is currently used as a diagnostic tool in clinical practice. The sampling of experience, psychopathology and behavior in the natural occurring contexts of daily life is promising because it generates the information we need to customize treatment selection. ESM can be incorporated in a dynamic process of recurrent periods of assessment, intervention and evaluation leading to further optimization of the rehabilitative practice over time.

Standard intervention protocols for all subjects, even those treatments with the highest outcomes in evaluation research, are not always the best options for specific patients. Individualizing treatments to a subject's own goals and adapting them to the patient's own strengths and weaknesses has been heralded in psychiatry as the premium strategy for chronic mental health. The diagnostic tools that allow such a process are, however, lacking. This leaves us with a credo to which we pay a lot of lip service but are unable to realize. Our goal is to apply ESM as a diagnostic tool that facilitates the decision making process in order to individualize intervention strategies. Contrary to traditional assessments in psychiatry, the ESM process uses only the subject's own observations. The case material presented in this chapter will demonstrate how ESM material can be applied in a clinically meaningful way, often leading to unexpected non-standard insights and interventions.

In the first case, a schizophrenic patient is followed over a period of almost 1 year. We pay special attention to the effects of interventions in the field of work rehabilitation and illustrate how ESM allowed us to evaluate the unexpected effects of an intervention, and helped us to customize the future interventions over later periods of time. In the second case we used ESM to get a feeling of the activities of a chronic schizophrenic patients in situations that were unavailable for observation by the clinician. The concurrent assessment of experience in different situations enabled us to make predictions of the possible evolution of mental state over time and to formulate advises to optimize coping and long term community tenure.

The last case illustrates the limitations of the traditional nosologic diagnostic system in a seriously ill patient with no clear-cut diagnosis. In this case ESM helped us to formulate a tentative diagnosis of “adult autism” (Asperger syndrome). We were also able to demonstrate successful and unsuccessful coping options of which the patient was unaware and which allowed us to target interventions. In addition to this interesting diagnostic content, we also present the standard output of the computer program we use to generate the ESM material for decision making in clinical practice.

CASE 1: "AGAINST CULTURAL EXPECTATIONS"⁷⁵

"The challenge of psychiatric research over the next decade seems clear: the systematic description of mental disorders and behavior with a high degree of situational and temporal detail."

M.W. deVries (1987)

The deinstitutionalization of mental hospitals and the movement towards a community integrated mental health care system is a resilient worldwide trend in the organization of medical services. From the patients' point of view, the central objective of this process is to optimize the functioning of individuals in their own living environment. This differs from the period prior to deinstitutionalization, when the traditional psychiatric hospital was the standard living environment for chronic psychiatric patients. In those stable settings, systematic treatment plans as well as relatively standard interventions were developed and carried out. Deinstitutionalization, however, added complexity and reintroduced the patient's own environment into treatment planning. As a consequence, mental health services had to incorporate the realities of community life into treatment programs. Personalized rehabilitation strategies are an attempt to tailor rehabilitation to individual needs. Developing treatment programs in such a complex context is not easy, particularly since the data required to plan treatments, the actual interactions of the patients with their living environments, are generally hidden from observation by mental health professionals. Personalized rehabilitation extends traditional treatment approaches that attempted to correct skill deficits through training, by incorporating information about the patient's interactions with an active and changing environment in the rehabilitation process. In the personal rehabilitation process as well as in deinstitutionalization in general, psychiatry has become more dependent on the existing environmental diversity of social relationships and changing socio-cultural goals. Psychiatry, by the nature of its own development, has been challenged to develop assessment strategies and interventions that promote and support individual adaptive strategies (Lieberman, 1988; Watts & Bennett, 1983). Leading researchers in the field of psychopathology highlight the importance of adding the dimensions of time and space to the description of mental phenomena (Zubin, 1989; Ciompi, 1989; Strauss, 1989). Research concerning the variability in illness course of depression and schizophrenia (Brown & Harris, 1978; deVries & Delespaul, 1989) has brought to our attention the host of assets and protective factors that influence the shift from health to sickness and back. The natural daily life context in which coping and adaptation occur are the target of personalized rehabilitation strategies. Today the clinical interview is the standard tool for gaining information about the quality of a person's life. We feel, however, that interview data or questionnaires for do not

⁷⁵ This section was written collaboratively and has been published: van der Poel, E. G. T., & Delespaul, Ph. A. E. G. (1992). Applicability of ESM in Personalized Rehabilitation. In M. W. deVries (Eds.), *The Experience of Psychopathology*. (pp. 290-303). Cambridge: Cambridge University Press. Eith thanks to Anette Hoffman for conducting the intensive interviews.

elucidate sufficient situational detail to genuinely add to our understanding of disordered behavior. Assessment methods such as the Experience Sampling Method (ESM) go a step further and allow the collection of concurrent information on subjective experience and psychopathology with the description of the context in which they occur.

We thus propose to supplement the interview that is traditionally applied in the assessment of the person-environment interactions with ESM. In this chapter we will illustrate the advantages of ESM over interviewing techniques for collecting information required for implementing personal rehabilitation in a case study.

The evaluation of a year in the life of a chronic mental patient

The rehabilitation process of John, a schizophrenic patient, is traced over the course of 1 year during which time we evaluated the effect of different therapeutic interventions on the subject's mental state. We asked if the choices made in the rehabilitation process fulfilled the goals of: 1) stabilizing his mental state; 2) enhancing his self-confidence; and 3) optimizing his quality of life?

Assessment techniques

Intensive interviews and the Experience Sampling Method were used to study decision-making in the rehabilitation process. We repeatedly gathered intensive clinical interviews. Over the course of 1 year a total of 37 interviews were conducted every 2 weeks. The interviews focused on current mental state and experiences between interview periods. The aim was to assess patterns in the course of the disease and rehabilitation process as reflected in the experience of activities and environmentally induced changes. The interviewers and the patient developed a strong research alliance. Although information was fed back to the treatment staff once, no direct therapeutic responsibilities were assumed.

The second method was the Experience Sampling Method (ESM), described in more detail by Csikszentmihalyi & Larson (1987) for normals and Delespaul & deVries (1987) for schizophrenic subjects. Subjects carried a small signaling device that beeps 10 times each day for 6 consecutive days. The signal is the well known alarm of a digital watch which is commonplace and does not create excessive commotion in the subject's environment. At the signal, the subject is asked to fill out a small booklet containing a set of mental state Likert-type self-assessments and open-ended questions structured in subsets of items that gather information about thoughts, mood, pathology, activities, the number and kind of people present in the environment and the place he is at that moment. ESM is ideally suited to assess the data used in treatment planning because ESM data is not distorted by retrospective recall. We further attempt ecologic validity in that the data is generated in the natural context of the patient's live, which is usually an inaccessible context for mental health professionals. In the present case study, the subject was sampled during three periods of 1 week each, spread over a year.

Subject

John is a 45-year-old male, the fourth of six children from a working class Dutch family that was dominated by a strict matriarch. Since childhood John has led an isolated life marked by few friends and little contact with his brothers and sisters. John was married at the age of 20. He did not want children, but his wife insisted as soon as they were financially and materially prepared. After 10 years of marriage, they had a daughter. John was first employed as a piper-fitter in a chemical plant, followed 5 years later by a job in a porcelain factory. Before his psychiatric problems began, he had been promoted to a low level management position in this firm. He was also active in the union, being particularly concerned with health conditions at work.

After the birth of his daughter, 7 years prior to the study, his mental health career began. Following an acute period of paranoia related to conflicts over his union membership and involvement, he presented to the crisis intervention team at the local mental health center. Subsequently, he lost his job, was divorced and took up residence in a caravan on a secluded lot. He felt disappointed and lived isolated from all social contact, including his family. His behavior was driven by paranoid delusions that resulted in frequent violent and aggressive outbursts toward the mayor, doctors and social workers he visited. After being threatened and robbed one night in his caravan, he refused to return. Following the advice of his priest he moved in with his mother, father, eldest brother and youngest sister. There he was overwhelmed by his dominating brother, whom he attacked with an ax, leading to a trial conviction, and was assigned to a psychiatric hospital for 1 year.

John remained in the hospital for 3 years and complied with an oral Haloperidol medication regime. Since his initial period of aggressive behavior, he has today transformed into a sub-assertive behavior style, marked by paranoid delusions and persistent fears of decompensating again. He ruminates over why he has lost everything he owned: his family, his work and his health. Accordingly he has developed the delusional rational of: "I became too important. Therefore some people (organizations) set the whole thing up to get me down and they can do it again whenever they like". He described the goal of his sub-assertiveness as a coping strategy to keep him out of the "limelight", to stay "unimportant". The role of chronic mental patient became for John a safe identity. He was thus strongly motivated to keep contact with his doctors and the psychiatric hospital. Six months prior to the study he was discharged and began living on his own. John maintained regular supportive contacts with a day treatment program, but had no social contact outside his hospital acquaintances and all contact with his family was lost. During the interviews, John presented himself as a shy, submissive person, who experienced all events in his life as if they were happening to him; events he did not understand or was unable to influence. Although most of his earlier life had been marked by a "go getter" attitude, today he seeks the safety and security of the hospital and the industrial therapy activities. Over the last years, John had experienced all changes as frightening deviations from the daily routine. His DSM-III-R (APA, 1987) diagnosis was schizophrenia, paranoid type, in remission.

The therapeutic process described in this case study unfolds over the course of a year, beginning 6 months after his release from the hospital. We will pay particular attention to his interactions with the rehabilitation unit that formed the essential part of his life. For the whole period John lived in a small, one-person apartment, and attended day-care at the day-hospital. The year can be divided into three distinctive phases related to his rehabilitation process, characterized by the type of day-care provided. During the *FIRST* period, he participated in an *industrial therapy* program on the hospital grounds. This consisted of low-skill-activities such as packing products, which were therapeutically meant to structure the day and provide social interaction opportunities. The low skill activity was below John's abilities. Since the staff felt that the activity did not enhance his self-esteem, they experimented with tasks requiring more skill. His "promotion" started 6 weeks before the *SECOND* ESM assessment. John was appointed as the *driver of a small bus*, used to transport patients to and from the day-care center. John's driver's license rendered him qualified for the job. In addition, it provided him the prestige of being in charge of an expensive new automobile. In the 6 months between the first and the second phase other aspects of his life also changed: he was assigned to a new psychiatrist and regained contact with his daughter. He, however, felt lonely and additional social contacts were organized. The *THIRD* period was scheduled 3 months after this intervention as a follow-up evaluation of the long term effects of the *bus driver job and social contacts* on his life situation. In this study we compared the relative merits of the clinical interview and ESM in providing a clear picture of John's life.

Questions and variable selection

1. Given the severity of his illness John was stabilized clinically at an acceptable level when he was discharged from the hospital. Our research question was therefore: was this stability preserved after the interventions that were planned to enhance his self-worth and optimize his quality of life? Measures to determine this were derived from the clinical interviews and from the pathology and mood items, which were repeatedly assessed using the ESM. The following markers of psychopathology — experienced suspiciousness, thought intrusion, unreal feeling, feeling out of control — and mood items such as feeling happy, sad and anxious were selected from the ESM measures.
2. John's primary coping strategy at discharge was to remain unimportant; regaining self-worth was reported as threatening. His reluctance to assume meaningful social roles had obstructed John's rehabilitation process. His appointment to the bus driving job for which he was skilled was meant to help with this problem. But did bus-driving enhanced his self-worth? The variables selected for this assessment were expressions of enthusiasm in the clinical interviews and the ESM mental state items "relaxed", "powerless" and "self-assured"; activity assessments as "being active" and feeling "skilled" and motivational measures such as "concentration", "prefer doing something else" and "like doing" the activity.
3. Since total symptom alleviation is usually not a realistic goal in the treatment of schizophrenia, the ultimate goal of rehabilitation should focus on the quality of life. Was John's quality of life enhanced over the year? The ESM data used in this eva-

luation were the frequency distributions of places, persons and activities, and the responses to Likert items of “satisfaction” and “loneliness”.

Results

Assessment of mental state

John’s mental health history and initial clinical interviews confirmed that he was well-adjusted, although not symptom free, throughout the first assessment period. The data in Table 13A.1 from period 1 confirms this picture⁷⁶: low scores on most of the symptom dimensions were found with the exception of suspiciousness and anxiety. Because the subject’s mental state had been stable for quite some time, we assumed the bus driving intervention would not affect his symptoms. In subsequent clinical interviews, in contrast, decompensation due to overstimulation of his new work situation seemed eminent. On ESM, however, the postulated negative effect of the bus driving on his mental state was not confirmed. ESM data in fact showed no significant decrease in mental functioning and a gradual improvement of the mental health from period 1 to 3 was actually observed. One important and highly significant exception was that John gradually felt more “unreal” between measurements 2 and 3, a difficult to interpret finding.

Items		Period 1 (n=66)	Period 2 (n=56)	Period 3 (n=48)	Test F(2,166)	t-Test		
						1/2	1/3	2/3
Mood	happy	3.08 (0.27)	2.95 (1.41)	2.63 (1.02)	n.s.	-	(*)	-
	sad	1.21 (0.54)	1.14 (0.65)	1.00 (0.00)	n.s.	-	*	-
	anxious	2.80 (0.67)	2.65 (1.12)	2.88 (1.10)	n.s.	-	-	-
Pathology	suspiciousness	3.30 (0.68)	3.10 (0.37)	3.00 (0.21)	5.80 **	*	**	-
	thought intrusion	1.56 (1.07)	1.36 (0.84)	1.13 (0.64)	3.34 *	-	**	-
	feeling unreal	1.59 (0.80)	1.00 (0.00)	2.38 (0.70)	60.81 ***	**	(**)	(**)
	fear of losing control	1.33 (0.66)	1.00 (0.00)	1.00 (0.00)	11.47 ***	**	**	-

* p<0.05; ** p<0.01; *** p<0.001 (“0”: unexpected findings /worsening).

Table 13A.1 Mental state changes over time (1 = not at all; 7 = very).

Assessment of self-worth

One of John’s major problems was his low self-esteem and marked sub-assertiveness. These were related to the paranoid delusional thoughts that dominated his life and his related need to remain unimportant and stay out of trouble. In the industrial therapy activities in the day-care center during period one, John was relaxed and accepted his situation. Yet, in the interviews John continually complained that he had lost everything, that his life had lost any meaning. Therefore, the clinical decision makers decided to search for an intervention that would provide more meaning and offered him

⁷⁶ Before analyzing the ESM data we have to make adjustments for the response tendencies over the three periods in the Likert data. The high end of the scale was almost never used. When it was used, it was used for assessing positive experiences (being active, stimulating environment,...). Only 15% of the scales had at least one score above 5. At the other end of the scale the score of 1 was not used for two items (ideographic complaint and suspiciousness). Both are negatively formulated and have a direct relation with pathology. 92% of the scales were answered using the lowest position. Accordingly we considered a range of 1 to 5 to be the active range in the Likert scales and regard 1 = “not at all” and 5 = “very”. More information on interpreting ESM in Larson & Delespaul (1992).

the bus-driving job. ESM data (Table 13A.2), however showed a significant drop in his skill self-assessment between periods 1 and 2. The patient was a skilled driver, but he did not experience himself as skilled. For example, during the second period the self-assessments of being skilled were significantly lower on weekdays when he drove the bus as compared to weekends (2.89 vs. 3.50), while the reverse relation held in the previous period (3.84 vs. 2.86). Some mood scores such as (not being) relaxed, self-assurance and feelings of powerlessness did change in a negative direction between periods 1 and 2. In period 3, powerlessness and self-assurance did not recover.

Items		Period 1 (n=66)	Period 2 (n=56)	Period 3 (n=48)	Test F(2,166)	t-Test		
						1/2	1/3	2/3
<i>Mood</i>	relaxed	3.15 (0.61)	1.68 (1.05)	2.56 (0.71)	50.78 ***	**	**	**
	powerless	2.82 (0.68)	4.83 (0.43)	4.68 (0.77)	176.70 ***	**	**	-
	self-assured	2.51 (0.77)	1.15 (0.51)	1.02 (0.13)	133.41 ***	**	**	-
<i>Activity</i>	skilled	3.51 (0.94)	3.06 (0.73)	3.10 (0.91)	4.69 *	**	*	-
	concentration	3.65 (0.84)	3.10 (0.47)	3.21 (1.48)	7.35 **	**	**	-
	rather something else	1.69 (1.13)	2.46 (1.28)	1.78 (1.17)	6.34 **	**	-	**
	like doing it	3.43 (0.97)	3.19 (0.82)	3.50 (1.29)	n.s.	-	-	-

* p<0.05; ** p<0.01; *** p<0.001.

Table 13A.2 Activity assessment over time (1 = not at all; 7 = very).

Over all measurement periods, John became less self-assured. We had expected higher scores since John would gain self-confidence from managing a highly valued and prestigious activity for which he was skilled. Clearly, this intervention failed to reconstruct his global self-confidence. Instead of gaining strength the patient became overwhelmed with feelings of “powerlessness”. The transition from period 1 to period 2 is the most significant one. Over all the periods the mean powerless scores on working days, compared to weekends, were the highest ($F_{(5,164)} = 89.49$; $p<0.001$). No systematic morning and afternoon differences were found, but an interaction effect with periods was observed. During the bus-driving periods (2 and 3), the highest experience of powerlessness was observed in the afternoon, a no-work timeblock. In period 1 no difference were found between parts of the day. Interestingly, concurrent activities showed no significant relationship to mental state. The ESM data then showed that the link between feelings of powerlessness with the activities organized by the rehabilitation staff was not direct. It could be that the effect of work is not experienced concurrent with the activity, but that the accumulated effect is experienced later when at home and alone.

Assessment of quality of life

Enhancing the quality of life is the primary motivation for rehabilitation work. We used two criteria: an objective criterion based on the level of differentiation of the frequency distributions of the places, activity and social context at time of the beep (time-budgets); and a subjective criterion measured by asking someone how he feels. Did John live a restricted life? Considering the time budgets computed, he clearly did (Table 13A.3). The subject is alone and home almost all the time (see also Table 13A.4). Compared to computed activity frequencies, the ranges of social context and place were more restricted. These variables are situationally linked in that the people one meets and the places one goes are often a consequence of the living situation.

Activities, however, tend to be more personal choices less related to the living environment, especially when someone lives alone.

Items	Period 1 (n=66)	Period 2 (n=56)	Period 3 (n=48)	Log-linear model	
				Periods	Codes
Where?	Home	65	70	$\chi^2_{(42)}=1108.39;***$	$\chi^2_{(30)}=47.70;*$
	Network/Family	0	0		
	Work/Therapy	23	23		
	Other	12	7		
What?	Nothing	2	9	$\chi^2_{(63)}=363.71;***$	$\chi^2_{(44)}=54.38; \text{ n.s.}$
	Work/Therapy	12	9		
	Socializing	5	4		
	Household Chores	25	18		
	Leisure	37	38		
	Other	20	22		
Who?	Alone	75	80	$\chi^2_{(18)}=1622.26;***$	$\chi^2_{(14)}=26.02;*$
	Family	0	0		
	Friends	0	0		
	Colleagues	23	18		
	Strangers	2	2		

* p<0.05; ** p<0.01; *** p<0.001.

Table 13A.3 Time Budget changes (%) as an objective measure of quality of life.

Over the three assessment periods the subject’s living environment became even more restricted. The subject “lost” 25% of his, already reduced, opportunities for social interaction between the first and the second ES assessment period. Alerted by this ESM data and by John’s comments during the interviews, the case managers decided to adapt the daily routine and extend the opportunities for social contact. This had no effect on the time budgets. In the 3 months between periods 2 and 3 John lost an additional 50% of social time, leaving only 11% social interaction opportunities. This was statistically significant⁷⁷. No significant effect of the periods was found for the activity codes. Considering this negative development we have to conclude that John’s living arrangements had a deteriorating effect on the quality of his life, but John compensated for this by a differentiated activity spectrum, a differentiation that was not lost over time.

Situation	Period 1	Period 2	Period 3	Overall “loneliness”
alone/home	43 (65%)	39 (70%)	35 (74%)	1.57
alone/not home	7 (11%)	6 (11%)	7 (15%)	1.45
not alone/not home	16 (24%)	11 (19%)	5 (11%)	1.44

Table 13A.4 Changes in John’s environment over time (no data available on “not alone/not home”).

How did John perceive this living situation? We restrict ourselves to “satisfaction” and “loneliness” variable. Table 13A.5 shows a progressive deterioration of satisfaction from period 1 to period 2 and, also significant, from period 2 to period 3. Interestingly, the subject complained of loneliness in the repeated clinical interviews. This complaint is understandable considering the global assessment of his living situation. However, the experience of loneliness was almost totally unreported in the ESM reports in periods 1 to 2 and remained at this low level in period 3. Broken down by

⁷⁷ Using log linear model fitting techniques.

environmental situation (as in Table 13A.4), no differences were found in the loneliness scores.

Items		Period 1 (n=66)	Period 2 (n=56)	Period 3 (n=48)	Test F(2,166)	t-Test		
						1/2	1/3	2/3
<i>Mood (feeling...)</i>	satisfied	4.50 (1.07)	3.55 (1.51)	3.15 (0.36)	22.78 ***	**	**	*
	lonely	2.26 (0.88)	1.02 (0.14)	1.11 (0.45)	76.12 ***	**	**	-

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Table 13A.5 Changes in Satisfaction and Loneliness (1 = not at all; 7 = very).

Discussion

We have studied a paranoid schizophrenic patient living alone in a small flat after discharge from the hospital using interviews and the ESM. The working hypothesis of the clinical staff was that the patient attempted to avoid drawing attention to himself, assuming that this would lead to self-destruction. The rehabilitation effort was, therefore, not focused on psychopathology but on recovering self-worth and developing new activities. The medium was meaningful “work”. Considering the possible strain on self-worth of the patient and according to Gruenberg’s Social Breakdown theory (1967) the clinical staff anticipated the possibility that this task would overstimulate him. They were, therefore, prepared for negative consequences after their initial intervention. The uneasiness reported by the patient at the second measurement was interpreted promptly as an eminent decompensation, requiring admission to hospital. However, in response to John’s loss of social meeting opportunities, a simpler intervention was chosen: a small change in his day-program providing additional social hours with other patients over lunch.

A closer look at the subjective experience reflected in the ESM showed that, contrary to interview information, a decompensation was not imminent at period 2. Moreover, his repeat measures of psychopathology showed that the subject gained in health over the different assessment periods after his discharge. Nevertheless, a negative appraisal of his life situation was also observed in that the subject became less satisfied and happy. These negative feelings were not associated with the bus driving intervention, nor could the loss of self-confidence and the dramatic rise in feelings of powerlessness be related to it. It is of course possible that the feelings of powerlessness are not concurrent with activities but are delayed to a time when the subject gets the chance to evaluate his activities, when at home or alone. Moreover, the feelings of powerlessness were higher during working days but not during the bus driving activity per se, nor were they present in the morning when the bus driving took place. We do, however, think that the influence of the driving job did cause part of the negative observations at period 2, but this influence was temporary and he recovered in period 3. This is supported by the motivational data (“rather doing something else”) and “relaxed”.

To further describe the negative changes reported at moment 2 we investigated the feelings of loneliness. In contrast, between periods 1 and 2 the assessment on loneliness improved significantly, even though a rather dramatic change in time-budgets took place, resulting in even more time spent alone. The subsequent intervention made by the treating staff aimed at broadening his social opportunities, had no effect at

time 3. Considering the fact that the subject was only provided with opportunities for interaction with patients and former patients, our interpretation is that a “normal” extinction curve of interactions with this former colleagues after discharge is observed. No compensation for this loss was created within the patient’s life. John’s ambulatory life in a small one-person flat became a lonely life and the small changes in the daily routine initiated by the staff had only marginal effects. The intervention was not optimal, since it did not create enough opportunities to broaden his network.

Clinically, it is interesting to note that, when comparing the answers on the Likert type scales, improvement over time could primarily be observed on those items measuring positive symptoms of schizophrenia (suspiciousness, thought intrusion,...) while negative symptoms assessments (tension, feeling unreal,...) deteriorated. This was probably less related to the rehabilitation interventions than to the living situation (alone in a flat). In terms of setting effects the data is similar to that of Hamilton et al. (1989) who found correlations between negative symptoms and smaller social networks. The question remains whether inherent psychopathology, or the difficult-to-cope-with living situation, is responsible for this trend.

What advantage has the ESM offered over the interview in this particular case? While ESM did not always generate straightforward answers, it did provide a closer look at John’s actual experience, which is especially important when aiming to tailor the rehabilitational process to the subject’s individual needs. Neither the clinical working hypothesis, that the patient would decompensate as a consequence of the rehabilitation intervention, nor the predictions based on the overstimulation of “new activities” proved correct. The self-perception of skills, as compared to objective skills proved most important for the rehabilitation process. The patient’s narrative seemed confounded by the distortions resulting from the problems with accurate retrospective recall of experience, John’s communication of dependency needs and personal goals for therapy as protection. Those may have misled the staff. Rehabilitation interventions therefore tended to be based on projections of the therapist’s institutional goals and the patient’s unconscious needs, clearly another world from the patient’s actual living environment. ESM helps us to break out of this mould by visualizing the patient’s world better in a more cost effective way. This in turn, provides better treatment opportunities such as increasing “out of hospital” social network which in turn may be evaluated. We do not argue that the intensive involvement of personal counseling contacts should be diminished and substituted by the ESM assessment technique but this report highlights the importance of supplemental information about real life experience for adequately assessing and tailoring individual rehabilitation interventions.

CASE 2: "THE CHALLENGE OF PARENTHOOD" ⁷⁸

In the psychiatric care of chronic patients different organization principles are used. Qualitatively, good care has a number of characteristics. First, care has to be offered on a long-term basis and with continuity. This implies that repeated diagnoses, as well as frequently changing care relations and treatment settings, should be avoided. Care should be offered in the natural environment. This means that within the scope of the treatment the restrictions to liberties and abnormalities of the culture of origin and the natural environment of the patient should be kept to a minimum. Furthermore, care should be provided with a minimized intensity. The skills and behaviors controlled by the subject should thus be determined by diagnostics. In therapy they should be stimulated and they should be prevented from being lost. This is "measured care", but besides the restrictions in intensity, "measured care" also stands for restrictions in the scope. Qualitative care is adapted to the individual and connected with and limited to the needs experienced by the patient and his family.

The case manager should check quality by providing care that meets the above-mentioned characteristics as much as possible. In addition, he/she has to take into account the dynamic character of pathology and the fact that over- as well as understimulation can have negative consequences for the patient. The directive theory here is the "vulnerability-stress" theory, which states that each individual has a different, but stable, vulnerability. This explains why different persons do not tend equally to psychiatric problems in comparable situations. The vulnerability-stress theory defines vulnerability as something biologic and thus stable, whereas stress — especially environmental stress and life events — is the variable psychologic factor. It should be noted that stress is always "experienced stress" and therefore subjective.

Care consists of altering the vulnerability threshold through medication. Psychotherapy and training in coping behavior can reduce the negative reactions to events. Environmental stress should be analyzed and subdivided into a controllable and an uncontrollable component. The living circumstances, e.g. to be employed or not, living alone or with a family, can cause stress. Certain choices can be made to minimize the environmental stress. Subsequently, a policy can be outlined to plan life events on moments of low environmental stress. In this way, the risk for decompensation is kept to a minimum. Of course this is not possible for all life events. Self-chosen and thus controllable life events are, for example, choosing to live alone, starting a new study, whereas events such as being dismissed for economic reasons or losing someone in a car accident are always unexpected. A good overview of the vulnerability-stress theory is given by Falloon et al. (1993).

⁷⁸ This article published as: Delespaul, P.A.E.G., deVries, M.W. & Radstake, D.W.S. (1995). *Personlijke Rehabilitatie: Een 'Experience Sampling Casus'*. In G. Pieters & J. Peuskens (Eds.), *Rehabilitatie van de Chronische Psychiatrische Patiënt. Op weg naar een Gemeenschapspsychiatrie*, (pp. 215-234). Leuven: Grant. The authors wish to thank the team of the SPD (Social Psychiatric Service) in Maastricht, Truda Driessen, Helga Nauta and Frieda van Goethem for their support in the article, as well as Carolien Dijkman for translating this text into English.

There is, however, a problem with the application of the vulnerability-stress theory in practice. No single parameter from the model is known for a defined individual: neither the vulnerability level, the genetic material and the protection by psychopharmaceutics, nor the subjectively experienced environmental stress and the effect of life events. As a consequence, the presence of vulnerability and stress in a retrospective interpretation is based on the observation that someone has decompensated. As a basis for a scientifically reliable personal care, however, we want to be able to predict and control the effects of vulnerability and stress.

For the planning of qualitatively reliable rehabilitation programs for chronic mental patients we do have principles, but few practical diagnostic instruments. During the last 20 years, the interest in psychiatry has been especially focused on the classification systems for mental phenomena (the DSM and ICD classifications). In the meantime, the entire social scientific world, excepting psychiatry, focused more and more on elaborate phenomenologic descriptions. In ecological terms, these descriptions remain close to the subject in the natural context and include the promise to increase our knowledge. In psychiatry, we need instruments to observe the patient and his subjective experience in the direct interaction with his environment. The frame of reference, a group to which the subject is compared and which can predict his behavior, is less important. The observation material we collect is valuable because it is a direct description of feelings and reactions, not in hypothetical but in real situations experienced by the patient.

In this chapter we want to illustrate the use of the Experience Sampling Method, a diagnostic instrument that makes it possible to create an elaborate and valid description of the daily life of a patient, in the resocialization of a chronic mental patient.

The Experience Sampling Method (ESM)

The Experience Sampling Method, which means “gathering experiences”, has been developed in the United States and Europe. It is a method for repeated, randomized self-observations in the natural context (deVries, 1987, 1992). By combining these elements in one instrument, the shortcomings of other observation techniques of daily life — such as diaries, end-of-the-day evaluations, or event sampling — can be overcome. Applied in psychiatry, the ESM is a valuable technique, offering very detailed descriptions of the daily life of mental patients and the variability of symptoms over time and in different situations. Reliability and validity have been proved. These studies reveal a stronger sensitivity than traditional diagnostic instruments such as clinical interviews and questionnaires. This sensitivity makes it possible to monitor medication and accurately outline the effects of changes in the environment on mood. Moreover, in the ESM the influence of response styles is minimized (deVries & Delespaal, 1992).

The ESM has been standardized for application in the Maastricht setting. This means that all subjects were signaled to fill out a questionnaire during six consecutive days, at unexpected moments, 10 times a day. This questionnaire includes about thirty 7-

point scales, measuring mood, psychopathology, motivation, and assessments of social interactions. In addition, the open-ended questions give an impression of the thoughts, the place, the activity, and the persons in the environment at the moment of the signal. Because of the repeated measures, we obtain a fairly complete impression of the daily life of the subject during the sampling week.

The use of the ESM has been illustrated repeatedly in the frame of psychiatric care (Donner, 1992; Massimini et al., 1992; Delle Fave & Massimini, 1992; van der Poel & Delespaul, 1992; Delespaul & deVries, 1991). deVries (1992b) summarizes the possible therapeutic effects of the ESM as follows:

1. The effect on the subject when he assesses his personal experiences and behavior in his daily life. Increasing of the consciousness about mood and behavior.
2. The effect of the more or less interpretative feedback to the mental health worker, based on the collected self-assessments.
3. The effect of the ESM information on the planning of the care process, the mental health worker's perception of the patient, the choice of treatment alternatives and the specific interventions.
4. The directly derivable changes in living style: "case management".
5. The effect of this information on the therapeutic relation between mental health worker and the client, a tool for "empathy" — the feeling of being understood, ...

In the therapeutic process, the information gathered with ESM can be seen as a film of the daily life played by the patient together with the mental health workers. The great qualitative value of the information, through the recognizable embedment in the personal experiences, breaks through the dependence of the patient vis-à-vis his caretaker. By discussing the ESM-week together, mutual understanding will increase. The detailed dimensions and nuances in experiences, visualized through graphs, tables and patterns in reactions on environment, offer material that can be directly used in the clinical setting.

Peter, or "The challenge of parenthood"

Peter is a 36-year-old man, diagnosed with chronic disorganized schizophrenia, with acute exacerbations. The severity of the pathology based on the mean BPRS item scores is 2.83. Anxiety, feelings of guilt and hostility are intensified (score > 3), as well as the psychotic items such as distrust, unusual thought content and hallucinations. We also saw a certain tenseness and a locomotive hyperactivity. His social functioning is reasonably high: 4.6 on a scale from 1 to 7. The ambulatory supervision is done weekly by a psychiatrist of the Social Psychiatric Service. In this period, the patient takes Haldol 50 mg decanoate/ 2 weeks and orally an additional 5 mg/day. To diminish the side-effects he takes a daily dose of 8 mg Akineton. In the evening he takes 5 mg Valium.

He lives together with his 28-year-old girlfriend, also a schizophrenic woman, but of the residual type with especially negative symptoms. They have a 5-month-old baby and they are living in an apartment which is too small for the three of them. After the delivery, Hilde became psychotic and had to be hospitalized. The child was placed in

a foster home but after the clinical treatment both parents stated that they wanted to bring up the child themselves. Since the mental health workers were rather skeptical about the success of this initiative, the initial policy was to discourage the plan. Both parents, however, held to their plan and consequently the mental health workers decided to support them and set up coaching. Over 2 months the parents were trained to handle their child: knowledge, practical skills concerning care, feeding, ..., and emotional skills were trained in a Children's Day-Care Center.

The monitoring with ESM was requested for by the mental health workers to gain better insight into the contact of the parents with the child outside the training setting and to plan additional rehabilitation strategies for the entire system. Only Peter was prepared to participate in the study. We followed him from Friday till Wednesday.

Results

An important issue preceding the interpretations of the data is the assessment of the quality of the data. The data we received from Peter were better than expected. Table 13B.1 gives an indication of the reliability of the data on the beep level. Of the total number of beeps to which it was theoretically possible to react (n=60), only 4 had been missed. There was no systematic pattern of the moments on which beeps were missed. They were missed in the morning as well as in the evening, but it was remarkable that not the first but the second beep of the day was missed, and not the 10th but the 8th or the 9th beep. The three missed beeps on the last day could be an indication of fatigue symptoms. With an intensity of 10 beeps per day, the maximum number of sampling days is 6. Peter's reaction to the signal was remarkably quick. No single beep had to be kept out of analysis because the response came too late — usually the maximum interval is 15 minutes. The mean interval between the signal and the response is 4.9 minutes, with a range between 3 and 9 minutes. To assess the value of these intervals, one should realize that this includes the time necessary to find a quiet place to fill in the ESM booklet, as well as the time for actually filling in the questionnaire (this needs less than 3 minutes). Peter immediately interrupted his activity at the moment of the beep to fill in the ESM questionnaire.

Day	No. of beeps missed	Delay of response	Normal day	Influenced mood	Changed plans
1 (Friday)	0	5.3' (4' - 7')	1	6	3
2 (Saturday)	-1 (B2)	4.9' (4' - 6')	2	6	5
3 (Sunday)	0	5.1' (4' - 7')	2	6	2
4 (Monday)	0	4.9' (3' - 9')	1	n/a	n/a
5 (Tuesday)	0	5.1' (4' - 7')	n/a	n/a	n/a
6 (Wednesday)	-3 (B2,8,9)	4.0' (3' - 6')	3	4	4
	-4 ⇒ 56 valid	4.9' (3' - 9')			

Table 13B.1 Quality and reactivity of the ESM data on beep level.

Peter's response to the reactivity questions shows that the gathering of the data did not happen without any effort. These questions are assessed at the end of the day on a 7-point Likert scale from 1 (= "not at all") to 7 (= "very much"). Peter indicated that the sampling days were more or less normal days. Nevertheless, he adapted his activities under the influence of the ESM method (mean 3.5) and ESM also affected his mood during the day (mean 5.5). These results are understandable, given the response

quality mentioned above. Although the instrument is liable to reactivity, this information demonstrates that the data have been gathered in a substantially qualitative manner. We are convinced that this reactivity has little impact on the validity.

Assessment of the day structure

An important parameter to assess the coping of chronic mental patients with their problems is whether they are able to bring some regularity into their lives. This can be explained by two reasons. The first one is rather pragmatic: a chronic patient living a regular life also has an instrument to monitor his symptoms. Deviations from this regularity may indicate that there is something wrong, maybe an imminent decompensation. When signals are recognized, coping mechanisms can be put into action, e.g. by asking for additional medication, retiring, or doing things together with others (Lieberman, 1988). There is a second reason for leading a regular life. In addition to internal — i.e. biological — “clocks”, daily fluctuations are also synchronized by external events, the “social Zeitgebers”. Deviations in the daily rhythm can activate psychopathologic symptoms. The most important social Zeitgebers are the sleeping period, meals, and the starting and end time of all sorts of regular activities (Monk et al., 1990, 1991). In the ESM, a number of questions have been included at the beginning and the end of the day, in order to describe the day structure. We will first discuss the questions related to sleep and quality of sleep.

Day	Wake up	To bed	Sleep latency	Times wake up	Sleep quality	Breakfast	Lunch	Dinner
1 (Friday)	n/a	n/a	n/a	n/a	n/a	9:30	13:30	18:00
2 (Saturday)	7:25	22:33	n/a	0	4	9:30	12:30	16:00
3 (Sunday)	7:16	21:55	10'	0	5	8:30	13:45	21:00
4 (Monday)	7:15	22:10	15'	0	6	10:00	14:00	19:30
5 (Tuesday)	6:33	n/a	30'	1	5	n/a	n/a	n/a
6 (Wednesday)	7:34	22:30	5'	0	5	7:30	13:00	18:30
Minimum	6:33	21:55	5'	0	4	7:30	12:30	16:00
Maximum	7:34	22:33	30'	1	6	10:00	14:00	21:00
Δ (max-min)	1:01	0:38	25'	1	2	2:30	1:30	5:00

Table 13B.2 Sleep assessments and day structure.

Peter slept well during the entire period of the study. He fell asleep easily, he almost never woke up during the night, and he positively evaluated his sleep quality. The pattern of being asleep/awake was regular. He went to bed between 22:00 and 22:30 and got up about 7:30. There is no difference for weekend or weekdays. This regular life is a good basis for the care for the baby.

The regularity of the pattern of being asleep/awake does not at all apply for the meals. Breakfast is not related to the time of getting up ⁷⁹. The differences are striking: up to 2 hours and 45 minutes. Lunch and dinner are usually not related to events, but to the time of the day, i.e. 12:00 and 18:00. What events do disturb the regularity in Peter's life? This can probably be explained by the activities of the other family members. Sampling of his girlfriend Hilde could give an answer to this.

⁷⁹ The attentive reader may have noticed that a number of the times Peter wrote down for breakfast were before the time he got up. This can be explained by the fact that the times for the meals were written down at the end of the day, globally, rounded up in blocks of 15 minutes. On the other hand, the time of getting up is written down at exactly that minute.

Assessment of time budgets

The subjects participating in the ESM study are repeatedly (10 times per day, 6 days) answering the questions “Where am I?”, “With whom?”, and “What am I doing?”. On the basis of the answers we can determine their time budgets. Practically spoken, time budgets are assessments of the time subjects spend in certain situations, based on the frequency distribution of the observations. In this ESM study we only sample between 7:30 and 22:30, not during the night. Time budgets show intercultural differences (Szalai et al., 1972), as well as differences between pathologic groups (deVries et al., 1988) and within a subject in different phases of illness and treatment history (van der Poel & Delespaul, 1992). Time budget patterns are significant and sensitive to contextual changes. In Peter’s case the clinical importance of these data becomes clear.

Where? — Peter spends a lot of time at home, 71% of the time. The health care center takes second place (12%). He also spends time with his family and in public places (bars etc.). In real time this means a presence at home of 10 hours and 39 minutes between 7:30 and 22:30. The distribution of the remaining time over other categories is more or less equal: between 1 and 1.1/2 hour. This time distribution is not so unusual for chronic mental patients who are treated in an ambulatory setting. They usually spend a lot of time at home (76%). The remaining time, however, is often less differentiated and strongly dependent of the way their lives are organized — network, daily activity program, work. When they go out, they usually go to well-defined places. In Peter’s case there is relatively much variation.

Figure 13B.1a Time budget: places.

Who? — Who is present in his environment at various moments? He is alone only 21% of the time. This could be a consequence of the living situation in a small apartment, which becomes evident from the frequencies of the following categories. The time spent in the vicinity of his wife is 26%, whereas he spends 23% of the time together with wife and child. It is striking that the moments together with the child occur on 3 days for 88%, i.e. during the weekend and the following Monday. The data reveal that the living situation causes intensive interactions: Peter, Hilde and the child are together almost constantly. Because of the occasional irritations (high expressed emotion?), an interaction of 7.5 hours per day — more than 50 hours per week — can

be a risk factor. This supports the request for a larger apartment with more privacy, especially now that they have a baby. Peter's time budget pattern is extremely deviating from the pattern of the average schizophrenic, who is alone 50% of the time; spends 30% of the time with his family; and 13% with friends or other acquaintances.

Figure 13B.1b Time budget: persons.

What? — Next, we will discuss the activities. The most frequent activity is “daydreaming”. Peter spends a salient amount of time with reflecting, in an abstract way, about his life and the world situation. The other activities are, however, active. Household chores occupy 21% of the time, of which 70% consists of actively caring for the child. Spread over the 6 days of this study, this is 2 hours and 15 minutes per day. These data give us the answer which we hoped for. We can rely upon the fact that the child is being cared for. We can not tell yet if this is true for the longer term. It is also not clear how the care is realized, nor how husband and wife relieve each other. We will try to answer these questions.

Figure 13B.1c Time budget: activities.

Schizophrenics always spend a lot of time daydreaming/doing nothing (17%), but for Peter the percentage is extremely high (33%), at the expense of his social activities. Peter's time budget also strongly deviates with regard to household chores. For the other categories the deviations were minimal compared to the average schizophrenic.

Assessment of the mental state

In this paragraph we discuss the answers on the 7-point scales — a 1-score means that the item applies “not at all” and a 7-score the item applies “very much” at that moment. There is a number of possible checks to verify whether the meaning of the scale is completely understood. In Peter’s case all different scores of the scales, including the extreme values, are used. By studying two items of which the answers have an anticipatory day pattern, we can convince ourselves of the reliability of the answers. We can see that the hungry feeling is low at the first beep, then increases and remains high until the 7th beep of the day (approximately 6.00). The hungry feeling is completely gone then, but reincreases afterwards. The usual fluctuation in the hungry feeling, which in Peter’s case is not found, is the one around 24:00. In the same way, we can make an analysis based on the day pattern of tiredness, which is stable and slightly increased (score 3) from the morning until 19:00. From 19:00 to 23:30 feeling tired increases to 5. From these data we can conclude that balanced answers have been given.

Peter is feeling quite well, he is cheerful — not sad. In contrast, he feels lonely, anxious, slightly irritable and discouraged. In the next paragraph we will try to explain this mood pattern. In the ESM protocols, specific items have been added for each pathologic group. For Peter, the module for psychotic patients was applicable. From these items, two are striking: “I can’t let go of my thoughts” and “My thoughts are influenced by others”. This was to be expected because of the clear cognitive preoccupations described as activities. The other psychotic items are less striking. The mean score is 3 with an explicit presence of distrust and hallucinations — verbally as well as visually. Finally we discuss the answers on the assessments of the social interactions. These tend to a positive evaluation of the contacts — the wish to be alone is seldom present — whereas the contacts often have an active component.

Items		n	Min/Max	$\bar{X}$	s.d.
Mood	I feel cheerful	56	1-7	4.07	1.54
	I feel lonely	55	1-7	3.98	1.41
	I feel relaxed	55	1-5	2.76	1.20
	I feel anxious	55	1-7	4.49	1.41
	I feel irritated	54	1-6	3.37	1.40
	I feel sad	54	1-6	2.98	1.35
Activities	I prefer doing something else	55	1-7	3.53	2.04
Pathology (psychotic)	I feel suspicious	55	1-7	3.05	1.84
	I can't express my thoughts	56	1-7	2.91	1.62
	I can't get rid of my thoughts	56	1-7	4.96	1.31
	My thoughts influenced by others	56	1-7	5.07	1.54
	I feel unreal	56	1-5	2.59	1.40
	I hear "voices"	56	1-7	3.20	2.04
	I have "visions"	56	1-6	2.71	1.80
	I fear losing control	56	1-6	2.70	1.13
Complaints	Most important personal complaint	55	1-5	2.56	1.57
	Other personal complaint	55	1-6	2.96	1.64
Social interactions	I like this company	44	1-7	4.82	1.21
	I prefer being alone now	44	1-6	2.16	1.43
	We were interacting	44	1-7	5.32	1.61

Table 13B.3 Mental state.

Mental state in context

The most interesting discussion of the data material is related to the interpretable differences in mood, pathology, and the assessment of social interactions in different contexts. Because we are interested in the mutual relations between Peter and Hilde and the resulting interactions with relation to the child, we will use the social context as the basis for our analyses. The mental state is determined by means of the items from Table 13B.3, in such a way that a higher score is linked to a positive mood, absence of psychotic psychopathology and a positive assessment of the social interactions. A subtraction of the global mean (determined for all beeps) from the mean per context gives a score between plus and minus 1 that can be respectively interpreted as: Peter feels better or worse than usually.

Figure 13B.2a shows that Peter felt best when he was alone with the child. This situation, however, only occurs twice. Thus the assessment of this situation has a large measurement error. More generally we can state that, for a positive experience, the child and the group of friends play an important role. The situations in which he is alone or with his wife are less reinforcing. This is also proven by the irritations and the corresponding feelings of guilt which were described as a response to the slowness of his wife (negative symptom).

Figure 13B.2a Mood in different social contexts.

In the experience of psychopathology, we see the same pattern as for moods (Figure 13B.2b). Psychopathology — low figures are negative mental states and therefore reflect the presence of psychopathology — is mostly experienced in the alone-situation or when Peter is with his wife. Being only with his child is a neutral condition for psychopathology — it was very positive for mood. The observations during childcare when the spouses are together with their child, as well as the external contacts with friends provide the highest reductions in psychopathology.

There was no assessment of social interactions when Peter was alone (Figure 13B.2c). In the appreciation of the social interactions the positive assessments clearly shift

from the family situation to the external relations: friends and acquaintances. These situations are the most reinforcing.

Figure 13B.2. Pathology in different social contexts.

Figure 13B.2c Self-assessment of the interactions in various social contexts.

Discussion

Personal rehabilitation means fitting in care as much as possible with the daily life and experience of the patient. In this context, diagnostics is not used to select people for standardized interventions, but to gather information, enabling the health care worker to plan person-based interventions. In Peter’s case, including Hilde and the child, the rehabilitation process has already achieved important goals. The observed social and psychological situation is “good enough”. The ESM data made clear that the parents were caring for the child. The following three questions have our special attention for the further development of the rehabilitative process:

- What can be done to consolidate the day structure? The diversity of the times of the meals indicate little structure in “household” routine;
- Is there a way to guarantee, in the long term, the reinforcing value of the social interactions? Given the fact that frequently being together with his wife causes irri-

tation and feelings of guilt, it is important to create more space in the mutual relationship between the spouses. Therefore, two objectives are formulated:

- is it possible to create a better balance in the time budget at home, where the interaction with the child is more satisfying than the interaction with Hilde?
- can a larger apartment ameliorate the situation: in the spatial sense of a better division of the territory (rooms), as well as in relation to the activities around the care of the child (who does what and when?)
- Finally, it is important to follow this realization for the longer term. After all, the education of the child is a matter of years. Will the enthusiasm of the parents last?

From the above analyses we can derive a number of interventions, such as:

- an educational intervention about the necessity of structure and daily routine in life. In this interview both Peter and Hilde should be involved. If they both recognize the problem, the developed skills in coping with problems can be applied. As a homework task, for a month they could each write down in a diary when he/she went to bed, got up and had been eating. This information could be used in weekly follow-up interviews, to evaluate the proposed solutions or adapt the strategies. The goal is a "good enough" structure. After 3 months, a new ESM period will follow. The data of this period could be analyzed on this aspect.

An alternative is to bring a better structure in the lives of Peter and Hilde by means of a therapeutic activity program. The advantage would be that health care workers (e.g. the work supervisor) could continually monitor them and signal emerging relapses in which the health of the child could be jeopardized. Because both husband and wife are not motivated for these kind of activities this has not proved to be very successful. There is more motivation for the care of the child, an activity which they perform quite well.

- Their relationship should be influenced in order to avoid tension and irritations. On the short term, few intrapsychic changes are to be expected, so our attention should focus on the amount of time they spend together. They live in a small apartment, which, in the longer term, increases the risk for accumulating conflicts. This could be changed and improved by moving into a new apartment, provided that in this living situation the mutual territories are well-defined each having privacy and specific household tasks. To get a better insight in these interactions, a sampling of Hilde is desirable. Unfortunately, she could not be convinced of the relevance of this. She thinks the whole ESM thing is childish.
- Peter is often alone, a situation he doesn't experience as reinforcing. Theoretically, there are several alternatives: more time with Hilde, more time with the child, and/or more time with friends. Preference is given to the latter two alternatives, which do not occur very often in Peter's life, but which he experiences as reinforcing. A possible change is to substitute the time he spends together with Hilde (not very reinforcing) for more time alone with the baby (specific tasks for Peter) as well as time with friends. Although this is only a marginal change, a development in this direction could be a reason for Peter to more enjoy his time spent with Hilde. These hypotheses, derived from the data, can lead to interventions which can be evaluated in a following ESM period.

Sampling after 3 months

Three months later Peter was sampled again — Hilde was still refusing — and again there was a satisfying response frequency (52 valid beeps with a mean response interval of 3,5 minutes). We see an improvement of the day structure. This does not mean that there was a fixed time for eating and/or going to bed, but there was a clear reference time for these activities. Lunch was at 12:30 (without exception) and dinner at 18:00 (with 2 exceptions). They always got up around the same time, i.e. between 7:15 and 8:00. They usually went to bed at 23:30 (4 times) or 21:45 (twice). The presence of a reference time, with the possibility to deviate, is a better situation compared to the total absence of routine during the first sampling period.

If we have a closer look at the time budget, we can see that the time alone as well as the time with Hilde without the child has equally decreased (-10%), whereas the time with Hilde together with the child, as well as the time with the child alone remained the same. We can also see an important expansion of the time with friends and acquaintances (+20%). The balance is an expansion of the time spent in the more reinforcing social environments. This is a positive result. We see a similar positive trend related to the activities. There was no change in the care for the child and a clear decrease in the daydreaming/"doing nothing special" category (-14%).

A look at the mental state, shows a significant improvement of psychopathology. The thoughts are less obsessional ($t = -1.73$; $p < 0.1$) and, in his experience, are less influenced by others ($t = -2.11$; $p < 0.04^*$). He hears voices less frequently ($t = -2.64$; $p < 0.01^{**}$). These are the most important positive symptoms and although the psychopathology remains manifest, the decrease is an important improvement. Furthermore, Peter feels more active in the second sampling period ($t = 2.26$; $p < 0.03^*$) and the desire to be alone occurs less often ($t = 2.11$; $p < 0.04^*$).

In total, the second sampling gives a consistent impression. There is more routine in the organization of Peter's life and he feels better. Conditions have been met to face the future optimistically.

CASE 3: DIAGNOSTIC FUZZINESS

To illustrate the diagnostic use of ESM in clinical practice, we will present a case in which we were diagnostically confused and ESM was used to generate new information with more richness in details than was available from of clinical interviews and behavioral observations. The case demonstrates the use of the standard analysis protocol of the ESDbms, of which the unabbreviated output is added in Appendix 7. To fine-tune the interpretations we computed some additional analyses in SPSS.

Case description

Bob is a 27-year-old man. He was referred to the closed ward of the mental hospital because of a vital depressive episode leading to repeated suicide attempts with pills, binge alcohol drinking, scratching and jumping before cars. He had a long history of mental illness, starting at age 3, and has been almost continually in residential care since age 16. Interestingly, he was able to graduate from high school, and his IQ was rated as above 110. Diagnostically, the picture was unclear. The depression that led to his current hospitalization from a sheltered living environment was assumed to be reactive upon a less well defined clinical state. Bob presented himself as a shy and socially avoidant individual, causing a lot of irritation by yelling, cursing and using abusive language. This seemed to be non-functional, occurring in almost all places and situations. Other verbalizations were one-word orders: “pills”, “food”, “open door”. Only exceptionally we were able to have a short conversation with him. The topics were related to selecting treatment objectives. Bob was never fully involved but made clear — using short replies — which options were acceptable to him. In addition, Bob was extremely anxious and walked along walls in a crippled and quasi-ritual way, as if the open space of the 2-meter-wide corridors was too threatening. No evidence was found for psychotic symptoms such as delusions or hallucinations. Another differential diagnoses that was considered was Gilles de la Tourette Syndrome but was discarded by two consecutive neurological assessments. The frustrating fact was that none of the diagnoses was satisfactory.

Cross-sectionally, using the BPRS, 7 items were rated severe (5 or more on a 7-point scale). The highest rating was found for “depression”, followed by “suicidality”, “emotional withdrawal” and “tension”. Still severe were “blunted affect”, “uncooperativeness” and “mannerism”. Generally speaking, we saw an elevation of the emotional cluster, as well as tension and bizarre behaviors.

At the moment of assessment, Bob was receiving Fluvoxamine and Seresta, which controlled the depressive episode in an atypical way, leaving a spectrum of asocial behaviors — including binge drinking — and a high consumption of pain medication in response to unspecified “headaches”. Because the treatment stagnated we were interested to know what caused those behaviors and to understand the diagnosis underneath. We decided to use the Experience Sampling Method to gain a better insight in Bob’s daily life experience.

Responses to the ESM methodology

He was introduced to the ESM method in a meeting, together with his parents. They are very involved in his treatment. We explained that we wanted him to respond on the ESM forms as quickly as possible after the beeps that will randomly signal him over the course of a week. Our purpose was to know and understand what he was thinking, doing and experiencing all day long.

The first attempt failed. Bob responded to 4 beeps the first day, a Friday, but his response rate dropped to 2 on Saturday and even 1 on Sunday. On Monday, a second briefing was given. This motivated Bob, however, for only 1 additional day. Over the whole week, Bob responded to 19 beeps, 14 of which were valid. Because demotivation and uncooperativeness were typical characteristics of his, we were not very confident about being able to get good assessments out of these data. However, when scanning through his answer forms, we were struck by the consciousness of his answers. When checking the context of the valid beeps, Bob was at home and alone in his room for 100% of the time. His shyness probably caused this skewness, which he confirmed. In addition — and unexpectedly — the pattern of activities was differentiated: 43% unpurposeful (“sleeping”, “nothing special” and “thinking”) but 57% of purposeful activities (“TV watching or listening to the radio”, “reading a book” and “hobbying”). His thoughts were even more intriguing: “about the batteries that I have ingested yesterday and that I have to shit”; “there’s nothing on TV”; “I’m hungry already and it is too early to eat”; “I’m feeling restless and this irritates me”; and “I enjoy this TV program — How beautiful are those nuclear explosions!”. We decided to give it a second trial, 3 weeks later. This second attempt gave us a full set of assessments (45 responses; 38 were valid⁸⁰).

ESM related observations

Thoughts

Our first observation was that Bob was constantly monitoring his mental state, using the broad range of emotional alternatives that were part of the ESM forms. This was in contrast to the reduced vocabulary of curses he used in his normal communication. To illustrate the striking contrast between both, we hereby give a verbatim description of the open-ended questions on 2 days (Table 13C.1a and 1b).

#	Description
1	08:39 MISSED
2	09:18 Response: 09:31
	Thought ... of the stuff in the newspaper. Disappearances in a cave.
	Activity Preparing for the ESM assessments
	< Events Reading the newspaper (rated slightly pleasant)
	Comment A little headache
3	11:57 MISSED

⁸⁰ Bob sometimes tried to guess the beeping moment after he finished answering the questionnaire. 6 beeps had a response time 1 or 2 minutes before the actual beeping time, while one was too early — more than 5 minutes — to be invalid. Bob’s “guesses” were so close that the actual answers probably came within 10 minutes and are therefore considered valid.

#	Description (cont'd)
4	13:22 <i>Response: 13:25</i> Thought I'm a little bored. I might consider going for a drink. Activity Just sitting in the living room Comment I'm bored, don't want to do anything. I could read, but don't want either. Would like a glass of beer.
5	14:27 <i>Response: 14:27</i> Thought To the things I saw on TV: Bosnia and Somalia. It makes me sick — I want a drink Activity Watching TV news Comment I feel bored/ don't want to do anything. Look and zap TV. I would like a glass of beer. Feel chagrined.
6	15:59 <i>MISSED</i>
10	B:21:49 <i>MISSED</i>
	<i>Day News</i> Israel: the Rabin government accords an agreement with the PLO to hand over Gaza and Jericho and <i>Headline</i> to engage in Palestinian self-government.

Table 13C.1a Verbatim responses to open-ended questions on Tuesday (day 1 of period 2).

#	Description
1	08:51 <i>MISSED</i>
2	09:58 <i>Response: 10:00</i> Thought I'm thinking about my life. It stinks. Shit. No future. Just vegetating in the mental hospital. Activity reading the newspaper Comment I'm getting fed up. Also, to be in a mental hospital.
3	10:56 <i>Response 11:00</i> Thought My life is senseless Activity Just sitting in the living room Comment I want a bottle of "vieux" (heavy alcohol)
4	13:04 <i>Response: 13:00</i> Thought My head ache irritates me. I'm wondering why so many people are on the ward today. Activity I'm lying in bed and listen to the news on the radio Comment My headaches. I don't want those strangers on the ward. I'm feeling tense.
5	14:02 <i>MISSED</i>
6	15:18 <i>Response: 15:17</i> (Guessed retrospectively and missed by 1 minute): <i>ACCEPTED</i> Thought The fact that I don't feel well, irritates me Activity I'm lying in bed and listen to a news commentary program on the radio < Event My headache became less intense (rated "0": not positive or negative)
7	B:17:28 <i>Response: 17:28</i> Thought I'm wondering what I will do the whole day (staring at the TV or going for a drink in the cafeteria). Activity Resting (I just got dinner) < Event Dinner (rated "0" — neutral) Comment I feel terrible. My abilities are at the lowest ebb. I also have diarrhea.
8	18:58 <i>Response: 19:01</i> Thought Dissatisfaction about the "oligophrenic" pocket-money regulation (DFL 5 /day). Irritated by headache. Activity I'm sitting in a chair listening to the radio Comment My head aches/ I'm feeling tense and cannot concentrate
9	20:48 <i>Response: 20:49</i> Thought About an article in a periodical about the weapons industry. Activity Watching TV on my bedroom while reading the periodical Comment Headache/ trouble concentrating
10	21:40 <i>Response: 21:42</i> Thought I'm reflecting about the documentary program on TV. I'm not motivated for anything Activity Watching documentary on TV Comment Growing dizziness — no headache anymore. Concentration is better. Sit in a chair just getting irritated
99	99:99 <i>Response at 00:01</i> (no response on a random beep) <i>SKIPPED</i> Thought I'm reflecting on Newsnight (BBC TV) about the news in the world Activity "Gazing" at the TV news for the whole evening < Event That I finally feel a little better (rated "2" positive) Comment No "headache"/ less dizzy/ concentration and tiredness is better/ feeling more lucid/ less tense
	<i>Day</i> I had a bad day today. I came to nothing. Headache/ tension/ feeling depressed, not feeling well. <i>Reflections</i> This makes me want to go for a drink. Now and then I'm irritated by the "oligophrenic" regulations for my pocket-money (in respect to alcohol). But if nothing helps, I (sometimes) don't know anything better (except going for a drink). <i>Day News</i> Bosnia: peace negotiations failed. <i>Headline</i>

Table 13C.1b Verbatim responses to open ended questions on Wednesday (day 2 of period 2).

These two days illustrate the interplay between illness, stress and coping, but also the role played by interests and personality in this process. Here, we did not see one-word comments, telegraphic style of speaking, major communicative deficits and lack of reflection on the world around him, nor on himself. From a statistical point of view, the thought data show a differentiation that is comparable to the one we found in the sample of chronic mental patients. Bob, however, shows less thought-related pathology (“goofs”), is more focused on current activities (78%) and is clearly more negativistic (100%) than his fellow patients.

Mental state

A first glance at the list of mood items shows that Bob was never “lonely”, “anxious” or “guilty”. No variability on negative mood items such as “anxiety” and “guilt” is often found in the ESM forms of normal subjects. Never feeling “lonely”, not even the slightest bit, is not. Even more exceptional is the complete lack of variability on “satisfaction” — a positive mood item — and the extreme low variability (s.d.<0.75) on additional scales such as feeling “unsure”, “relaxed”, “active” and “challenged”. The pattern of emotions reflects the affective blunting especially in relation to the social context and activities. In addition, the extreme scores on tension and the continual lack of satisfaction clearly illustrates Bob’s experienced burden. The blunted emotions do not fulfill a protective role. Items that show variability (and which can be used to analyze environmental reactivity) are primarily the assessments of cognition (“pleasant”, “clear” and “normal” thought, as well as “concentration”). In addition, the mood items “cheerful” (low mean, but high variability), its inverse “sadness” (high mean, high variability) and “angry/ irritation”, are promising. In the activity cluster, only “motivation” and “skilled” are useful. In addition, the social interaction is actively monitored by Bob (“pleasant company” and “prefer being alone”) but “doing things together” is infrequent.

Items		$\bar{X}$ (s.d.)	Range	Items		$\bar{X}$ (s.d.)	Range
Thought	pleasant	2.04 (1.44)	1-5	Social interaction (n=15)	pleasant company	3.43 (1.55)	1-5
	clear	6.65 (0.86)	4-7		prefer being alone	4.75 (1.53)	1-7
	normal	5.33 (2.44)	1-7		active together	1.36 (1.08)	1-5
	concentration	5.12 (1.18)	2-7		difficult to express	3.98 (2.68)	1-7
Mood	cheerful	1.49 (1.14)	1-6	Pathology (thoughts)	obsessive	2.39 (2.12)	1-7
	angry/irritated	4.57 (1.65)	1-7		(influenced by others)	no variability	
	sad	4.39 (1.73)	1-7	Idiographic	1. tension	4.92 (1.43)	1-7
	(unsure, lonely, relaxed, anxious, satisfied, guilty)	no variability		Complaint	2. negativism/sad	4.50 (1.55)	1-7
Activity	not motivated	2.29 (1.60)	1-6				
	skilled	5.71 (1.46)	2-7				
	(active, energy, challenge)	no variability					

Table 13C.2 Overview of mental state.

For the ESM sampling of Bob we used the standard booklet for schizophrenia, which is usually also selected for chronic mental patients. As could be expected, Bob reacted to none of the most typical schizophrenia items (suspicion, hallucinations both verbal and auditive, “fear losing control” and “feeling unreal”). In addition, Bob’s thoughts were rated “difficult to express” and “obsessive” and to a lesser degree “influenced by others”. In addition, the items that reflected his own problems the best — selected by

Bob during the briefing session — were “tension” and “negativism/ sadness”. The second idiographic complaint matched the ESM item “sadness” in answer pattern. However, the first one, “tension”, was not (negatively) correlated with “relaxation”. For Bob, tension as a reflection of his own complaints, was something different from “not feeling relaxed”. Tension was related to his headache.

Context

The contextual differentiation of the whole sampling period shows more diversity than the first week. Being alone (70.6%) and at the hospital ward (home: 80.8%) still got the highest frequencies, but there was more differentiation now (persons: family 25.5% and mental health professionals 3.9%; places: network 15.4%). The most unexpected observations were found for activities. While the nursing staff usually observed Bob in his bedroom, playing dead, hiding under his sheets, the ESM showed a picture of a subject watching TV, listening to the radio (38.5%) and reading the newspaper or other books (11.5%), with a special interest in news and commentary programs, which he critically, intellectually and even emotionally commented on. Summarizing, the differentiation in activities, already observed in the first week, was replicated in the second period, but the larger set of observations allowed us to assess the cognitive and emotional appreciation of those activities.

Mental state in context

Time effects — The protocol overviews “days of the week” and “week/weekend” are used to select eventful days and to focus our attention on those moments. As a general trend, Tuesday and Friday were the best days, while Wednesday was the worst. Bob was not aware of this. During the debriefing he described the week as “more troublesome compared to normal, especially on Thursday to Saturday — Sunday being better.” Only the Sunday effect — the most recent day at debriefing — was confirmed in ESM. We saw a clear difference on a set of variables on week compared to weekend days: the best moments were during the weekend (13 out of 17 items; $z=1.60$ $p<0.06$). Interestingly, Bob was less tired (1.6 versus 2.6, $F_{(1,49)}=8.80$, $p<0.01$) and less tense (4.2 versus 5.3, $F_{(1,50)}=7.50$, $p<0.01$) on weekend days. On the other hand, weekend days had the highest level of both positive and negative between beep events.

There were no specific trends for parts of the day. The significant correlations we found may be a random consequence of the large number of relations tested. Possibly relevant, however, is the higher intensity of positive events (4.0 compared to 1.4) and ratings of “pleasant company” (5.0 compared to 2.3) in the evening hours.

Context effects —

- *SOCIAL CONTEXT* — The pattern of mood and mental state differentiation related to social context is not straightforward. The table shows a lot of significant differences, but it is difficult to derive a general conclusion, partly because of the confusing headings that were given by the standard protocol. Recomputing the data and restricting it to account only for the alone/not alone difference, we found a trend towards a better mental state in the not-alone situation (13 out of 17,

$z=1.60$; $p<0.06$). Comparing the categories with the highest frequencies, the not-alone situation — being with family — strengthens this trend: being alone is a more negative situation (15 of the 17 relations, $z=2.41$; $p<0.01$). To reduce the complexities of the interactions we computed a mean mood (over 9 variables with a s.d. >0.75) and pathology score (5 items). Alone / not alone differences were significant for both parameters (mood: $F_{(1,44)}=5.47$, $p<0.02$; pathology: $F_{(1,48)}=4.96$, $p<0.03$). The differentiation based on the type of relation Bob had with the persons present, favoring his own family, was also confirmed (mood: $F_{(2,43)}=3.87$, $p<0.03$ / family effect: $t_{(29,10)}=3.08$, $p<0.01$; pathology: $F_{(2,47)}=2.45$, $p<0.10$ / family effect: $t_{(28,22)}=2.44$, $p<0.02$). However, when we recomputed the data to allow a differentiation based on the number of persons present, being in the company of other persons and especially family members was not irrespectively the best situation. We computed a variable using the number of persons in the environment, in an alone code ($n=36$), a small (1 or 2 persons, $n=10$) and a larger group (3 persons, $n=6$). In this last group, 5 out of 6 observations involved family members (3 times the third person was another relative, one time a stranger and one time unspecified), one time 3 mental health professionals. Results are shown in Figure 13C.1. Bob felt best in the situation where only two persons were present (8 out of 10 were his parents). The results for mood were significantly different ($F_{(2,44)}=3.37$, $p<0.04$), but the larger social environment (number of persons=3) took a middle position between two persons, the best situation, and alone, the worst. Only the difference between the extremes was significant ($t_{(17,23)}=2.73$, $p<0.01$). Considering pathology, the relation was even more pronounced ($F_{(2,48)}=5.90$, $p<0.01$). Pairwise comparisons confirm that the alone situation is the worst environment and with two persons is the best (different at $t_{(40,92)}=5.36$, $p<0.001$). Here the company of three persons matches the alone situation ($t_{(8,15)}=0.07$, n.s.) but is also significantly different from the best environment($t_{(6,25)}=-3.15$, $p<0.02$).

Figure 13C.1 Mental state in relation to the social context.

The difference between the three situations could not be explained by the presence of the parents. They were almost always present in the larger social environment of three persons. We assumed therefore that the mere fact that three persons is a

- more complex environment, caused this difference. This led to the hypothesis of a frontal lobe dysfunction, which was further explored using neuropsychologic testing. A crowded hospital ward maximizes this confusing and complex relational situation and might be the cause of the cursing and yelling that was the most salient observation of the hospital staff. Forcing him to remain in the livingroom of a crowded ward was a frustrating experience for him. The thought descriptions of the second day (Table 13C.1b) illustrate this even further. Although this relation is well-documented in the ESM, and we had some evidence from observational data that this was the case, Bob didn't know this. During the briefing session he recalled being able to describe the variability in his mental state, but also said "I'm never able to explain this variability" which was frustrating to him.
- *ACTIVITY CONTEXT* — The data did not show any relation with activities. We recoded the activity code into "doing nothing", "productive activities" and "leisure (including TV and radio)". The data were nonsignificant over the whole line (mood: $F_{(2,44)}=0.13$, n.s.; pathology: $F_{(2,48)}=0.40$, n.s.).
 - *PLACES* — The final analysis of context is related to the places. Persons and places are interrelated. So, the results found for persons reflect a little bit of what we found here. Bob did feel better, both for mood and pathology, when he was not on the hospital ward ("home" in his situation). The significances were $F_{(1,44)}=4.58$, $p<0.04$ for mood and $F_{(1,48)}=7.02$, $p<0.01$ for pathology. When we cross-tabulated places and persons we saw that all the non-home situations were with his parents and most of them are in the 2-person situation. Because ESM is an ecologically valid assessment, it is impossible to make an independent assessment of contextual elements. We think, however, that the interpretation presented in the paragraph on social context still seems to be the most likely interpretation.

Figure 13C.2 Mental state in relation to the places.

Coping attempts — The ESM also showed that binge drinking was to be considered as an attempt by Bob to self-medicate himself in response to diffuse feelings of restlessness. This was also confirmed by Bob in the debriefing interview. The pocket-money restrictions — allowing him a maximum of two alcohol consumptions in the pub — were experienced as a real boycott of his assumingly effective coping strategy. We checked on the effectiveness of the strategy by comparing ESM reports after alcohol

consumption with control reports. We expect that effective coping will lead to reduced pathology following an alcohol report. Results are shown in Figure 13C.3. The difference for mood was not significant ($F_{(1,45)}=1.92$, n.s.) but it was for pathology ($F_{(1,49)}=6.28$, $p<0.02$).

Figure 13C.3 Mental state in relation to the substance use.

To enhance our understanding of the dynamics of this assumed coping we checked on the last non-alcohol report before the alcohol consumption. This is the best possible approximation of the mental state before the alcohol consumption. The number of observations is low ($n=5$) but data show that pathology was at a low level then. There was no pathology before the alcohol consumption. This makes alcohol use as a coping strategy not very successful (3.04 versus 4.56; $t_{(4.93)}=-1.82$, $p<0.13$). The fact, as shown in the ESM ratings, that this drinking was unsuccessful in reducing the feelings of restlessness — possibly even worsening the mental state — came as a surprise to Bob. It is interesting to know, however, that this knowledge — when fed back to him — did not automatically lead to a behavior change. Bob was still able to conceal the procedures he had developed to raise money for his binge drinking. His pocket money was certainly insufficient for those expences. We suspected him of stealing but we found no reports of this. It is possible that this happened during the missed beeps in the middle of the day, and that reports were intentionally left out to conceal this fact.

“Quality of life”

The standard ESDbms output (Appendix 7) gives some indication of Bob’s “Quality of Life”. Unfortunately, he did not experience anything as challenging and therefore the item had no variability, making both the “Flow” and the “Escape”-indicators of QoL meaningless. The QoL assessment that uses the assumption that well-being will be experienced in those subjects that are able to spent most of their time in the most positive situation, resulted in a score of 0.23 for mood and 0.14 for pathology (using the contrasted activities operationalization). Both are above average. In comparison to

the reference group of chronic mental patients, Bob had a good level of adaptation. The activity with the highest frequency was "TV watching and listening to the radio", which was experienced as a lot more positive compared to a large number of different other activities such as "maintenance", "hobbies" or even "reading a book". From this point of view, his coping is effective. Changing his withdrawal on his bedroom, watching TV and forcing him to spend more time on the ward in the neighborhood of fellow patients might lead in a less satisfying situation.

Discussion

It was unexpected for us to find that Bob was able to reflect upon his own mental state and was able to monitor environmental cues. We derived this conclusion from the content of his thoughts in the ESM reports. It revealed a previously unexplored experiential world which we were able to assess using the ESM material in therapy sessions with him. Previously, Bob was uninvolved in the interviews with us in spite of intensive attempts over the months he was on the ward. In discussing his ESM reports Bob opened up to a level that allowed a collaborative engagement in his own treatment. He finally became a partner in his rehabilitation process. Even then, however, Bob was still a rather shy person who seldom spoke spontaneously, disclosed his inner experiences reluctantly and had to be prompted for responses. He spoke with a cynical undertone and a special blunted humor that excited sympathy.

Bob was reluctant in social contact but became relatively "open" while discussing the ESM material. In addition, he reported an environmental reactivity upon the number of persons present at the moment of the ESM beep. In contrast to the clinical interview, ESM assessments were made in intimacy and were therefore much stronger indicators of his social invalidation. Unconscious of the underlying social interpretation Bob responded differently from moment to moment in daily life. Until that moment the pervasive communication deficit had made it very difficult to make an appropriate assessment that could lead to insight in the diagnosis. ESM reports, however, reflected a variability in his mental state that could be related to social situations. This convinced us to look more into the diagnosis of adult autism and the "Asperger" syndrome. When assessed according to the "theory of mind" tasks (Happé & Frith, 1994) Bob functioned at a high level. This was confirmed in a neuropsychological assessment and consecutive training program. Social withdrawal combined with a higher than average intelligence and a relatively high level of functioning points to the diagnosis of "Asperger" syndrome (Frith, 1991). In contrast to Hurlburt et al. (1994) who studied "Asperger" subjects with the "thought" sampling technique related to ESM we, however, did not see a characteristic visualization of inner experience. Bob was extremely reflective and introspective in a verbal way. Further evidence for the "Asperger" diagnosis came from the family history — Bob's brother is diagnosed with autism — and from an in-depth developmental assessment involving both Bob and his parents.

After deciding upon the syndromal diagnosis the appropriate selection of therapeutic options was still a problem. As a consequence of the communication disorder, the nursing staff thought that Bob was constantly asking for headache medication and assumed he was abusive. In the ESM, however, we saw that tension and headaches

were salient experiences for him, and that he was able to rate the variability in intensity adequately. Verbalizations of cognitions showed that at moments of intensive headache, he was continually weighing pain levels. He realized that asking for medication at some moment would reduce his chances for additional medication later on the day when pain levels would raise even more. The ESM data told us to take Bob's headache reports seriously. Based on this data we decided to do a complete reassessment of the headache resulting in a nose operation that diminished the intensity of the headaches drastically.

ESM is an adequate assessment strategy when information on the subject's experience is crucial. The fact that the patient does not seem to be very communicative is often not a problem in ESM. As in the case of Bob, we were able to build a reliable research alliance in some very extreme cases. On the other hand, seemingly very communicative patients also refrained from participation. Prediction in advance was seldom possible.

Furthermore, the standard ESDbms analysis protocol (Appendix 7) gives a lot of interesting information but is not "smart" enough to allow interpretations that are readily available for clinical practice. The necessary additional analyses currently still take several hours of computations. We plan to develop a "smarter" program that advises on mental state summary measures and salient environmental cues. The ESDbms also needs an additional module that allows interactive assessments. This would enable the clinician to formulate additional "what if..." questions dynamically, customized to the patient's data. We think that the case of Bob illustrates the sensitivity of the ESM assessments and their possible use for the evaluation in clinical settings. To become useful in practice, however, the application needs to become more user-friendly.

Finally, interventions based on these assessments are not always possible. While a large range of alternative treatment strategies can emerge from ESM analyses, the implementation of these interventions in clinical practice does not always follow the assessments process. One of our future investments will be in developing the interface between ESM data and planned change.

NEGOTIATED MEDICAL CARE

The extensive cases presented in this chapter illustrate the power of ESM in clinical practice. In the first case we have demonstrated how a marginal change in daily life routine, applied with the best intentions, led to increased mental health problems. While the worsening of the psychiatric symptomatology was also clear from the clinical interviews, only ESM allowed us to relate the changes in the clinical picture to a growing intensity of loneliness. This conclusion was shared by the therapist and the patient and led to a marginal intervention that created no real objective change in the patient's time use pattern — he was still alone most of the time — but resulted in less emotional problems.

In the second case the mental health professionals were worried about the health of the baby child of two schizophrenic parents. Our problem was that we had no information on the child care routines of the parents when they were at home — a situation that is hidden from observation by the clinician. ESM showed sufficient caring and loving. Even more, because some stress was observed in the communication between the parents, we could stimulate the contacts of the man with his friends as a balancing coping strategy in the marriage relation.

In the third and last case we were confronted with a diagnostic puzzle. From the ESM observations we deduced that our patient was a smart and intellectually active person with a special invalidity in social relations — especially when they involved more than 2 persons. The picture matched the Asperger syndrome. We could engage in a discussion on coping strategies and demonstrate how the use of alcohol was less effective than the patient himself thought, to alleviate his restlessness and pain.

Assessing for “personalized rehabilitation”

The ESM is an important innovative assessment tool that allows customizing rehabilitative interventions. Older models that explain the emergence of psychiatric problems such as Gruenberg's “Social Breakdown Syndrome” (Gruenberg, 1967) and more recent models such as the “Vulnerability-Stress Model” (Zubin & Spring, 1977; Nuechterlein & Dawson, 1984) are all powerful descriptors of the course in schizophrenia, but miss the necessary diagnostic tools to predict the imminent relapse or recovery. At best, those models offer an “a posteriori” rationale that explains why the patient relapsed: the impact of the stress must have been too much for the patient. There is no way to anticipate how much stress a specific subject can bear at a certain moment in time. Therefore, opportunities for treatment are often missed. The result is that we play on the safe side and avoid inflicting our patients with treatment related stresses that might lead to the worsening of symptoms. The unfortunate result is that the patients live restricted lives in restricted (hospital) environments (Barton, 1959; Goffman, 1961). The development of diagnostic tools that customize interventions in

the care for chronic mental patients is warranted as a solution for the poor effectiveness of traditional treatment strategies (Bachrach, 1987).

We think that ESM can be such an assessment tool and can be applied to assess treatment needs as well as accurately describe therapeutic effects. Ideally, ESM assessment periods should be repeated every 2 to 6 months depending on the expected rate of change in the therapy. Such a procedure allows alternating circles of assessment, treatment implementation and evaluation, leading to dynamically updated optimized intervention strategies. Naturally, not everyone can or wants to be sampled. Schizophrenic patients, however, do not refuse more often than other subjects. Understandably, their motives for refusing are typical: "ESM is childish", "I am afraid that the watch might control me", or "the beeping device contracts cosmic rays". A good therapeutic relationship is an absolute condition for successful conclusion of the ESM-mediated rehabilitation process. Often, it is worthwhile to include the patient's network members in the assessment rationale. Sometimes, the family protectively took the watches and booklets away from the patients. If all the conditions are met ESM data can even be gathered in the acute phases of the treatment. We applied ESM with catatonic patients and in the separation room.

The promise: "negotiated medical care"

Detailed careful monitoring as part of the continuing dialogue between the professional, patient and care givers can help build a "collaborative ethos" in which patient and families regard themselves as partners in the management of schizophrenia (Smith & Birchwood, 1990; Drury, 1992, 1994). The motivational potentials of a shared responsibility in the treatment plan cannot be overestimated. There is a great demand for communication and assessment strategies that foster this shared responsibility.

ESM can be very valuable in bridging the gap of inequality between the clinician "who knows" and the patient "who does not know". In the traditional therapeutic relation the clinician is the specialist while the patient just waits to hear his verdict — the therapeutic plan. In ESM the patient is the specialist of his own life and the data that he has collected targets the interventions. The clinician just aggregates the data. Sometimes, the patient challenges the conclusions. However, because the face validity of ESM data is high, it is intuitively understood that the aggregation process changes nothing in the contingencies that are embedded in the data. By discussion and the involved questioning of the data, the patient becomes a partner in negotiating his treatment plan. As a consequence ESM offers a base for a true negotiated medical care.

The practice: Timing problems in chronic psychiatry

Treatment implementation in chronic mental health is subject to a good diagnostic process, but also to a careful monitoring and an appropriate timing. This is different from a treatment process for "new" patients or at first onset, where interventions follow almost naturally from the diagnostic process.

In chronic mental health the patient has an illness history that makes the choice of treatment strategies much more complex. We often know which strategy worked previously and which failed. But not infrequently we are left in doubt because interventions have resulted in partial effects and changes in strategy bear the risk of worsening the symptomatology.

The cases presented in this chapter are primarily illustrations of the theoretic assessment power of ESM. In reality, often the intervention of choice was postponed until the patient and therapist felt secure that implementing the strategy was the best thing to do. This could take a long time. We do, however, believe that it is important not to postpone interventions when changes are requested, for instance, in non- and partial responders. To do so, we have to generate the information that maximizes the confidence in the treatment effect. In the next years we will study how the ESM feedback protocol can be enhanced and the communication between patient and therapists optimized for the best possible treatment implementation for a specific patient at a specific moment in time.

Downloaded from <https://www.cambridge.org/core>. University of Cambridge, on 02 Jun 2018 at 10:00:00, subject to the Cambridge Core terms of use, available at <https://www.cambridge.org/core/terms>. <https://doi.org/10.1017/9781315325400.010>

Epilogue

Symptoms of Daily Life Adaptation

Richard Lamb once said that “chronic mental patients are eternally moving to nowhere”. When I think about this frustrating aspect of living with and treating patients suffering from schizophrenia the litho “Klimmen en Dalen” of M.C. Escher (1960) comes to mind. It is a beautiful illustration of an illusion in which two rows of unspecified “slave”-figures climb (“klimmen”) and descend (“dalen”) the same stairs, built in a square on top of a building. The climbing and descending figures never turn — always moving in the same direction. Nevertheless, each circle brings them back to their starting point. Despite all the energy spent, the investment is not productive. The image is appealing but illusory. Escher’s building can not be built because it makes a two-dimensional misrepresentation of a three-dimensional construction. An element that should be in the back of the picture (the end of the stair at the first floor) is brought to the front (where the stair started).

As with the Escher litho, the degenerative long-term course of chronic schizophrenia is an illusion created by our representations of the reality. In fact, chronic mental patients are ill part-time. Their illness is episodic. Often psychotic periods alternate with months or years without symptoms and most patients live at home. Why are we not able to recognize this long-term variability? We build our concepts of the illness on what we see in clinical practice. Successful cases do not recurrently show up for admission in the hospital but refractory patients do. Out of sight is out of mind. This makes it difficult to appreciate the long-term course of schizophrenia. Even less is known about its moment-to-moment variability. Analogous to Escher’s litho schizophrenic patients are skillful at climbing stairs but unsuccessful at reaching goals. We have to reckon that basic daily living skills remain unaffected. We believe that an in-depth understanding of the variability in emotions, thoughts and behaviors from moment to moment is important for the study of psychopathology and as a key to planned change.

VARYING SYMPTOMATOLOGY

For the most impaired chronic patients symptom intensity dramatically varies over the course of the day. Often patients do not recognize these changes in mental state. Only exceptionally are they able to use this variability constructively and optimize their daily coping. The clinicians are not of much help because their assessment strategies do not allow them to make this differentiation. They are blinded by the massive impact of the illness and are not able to detect its part-time character. We are even less able to use the non-ill parts constructively in therapy.

Those who work with chronic schizophrenic patients, as I have done for years, are confronted with a paradox. Marginal neuropsychologic deficiencies in concentration,

measurable in microsecond differences in reaction time, result in major adaptational disruptions. The anticipated behavioral deficits based on these tests, however, are unrelated to our actual behavioral observations. Concentration problems should make it more difficult for schizophrenic patients to assimilate the environmental and bodily information that they need for biking, but not influence grooming behavior. But in clinical practice we often observe skillful biking and poor self-care.

Even more intriguing is the lack of system in our observations. I remember a schizophrenic patient with major deficits in social skills and table manners who lived for years on a residential ward of our psychiatric hospital. On one of the rare occasions that his family visited him he was able — unexpectedly — to brew coffee and serve his guests in the most distinguished manner. He even responded to subtle non-verbal cues and presented his visitors a piece of pie at a socially appropriate moment. It was as if, in the right context, a whole lost behavioral repertoire was reactivated. Meanwhile, he was still psychotic and alluded in the conversation to the time he was fighting as a commander in the Dutch/British naval wars of the 17th century. The next day we were back to nowhere. The other patients on the ward were confronted with his poor table manners again and did not want to eat with him anymore. Neither our interventions, nor our inquiries into what happened the day before, were very helpful. An exhaustive assessment of overt behavioral excesses and deficits usually reveals major disabilities in chronic schizophrenic patients. The systematic application of interventions is undermined by behaviors that wax and wane. At one moment the whole behavioral repertoire is purposefully present, at another moment no stimulation and coaching seems adequate to help the patient express the appropriate skills. Induced behavioral change will not be sustained. After some time of intensive treatment we have to acknowledge that despite the recurrently emerging hope for improvement we are back at our starting point. I often advise patients, nursing staff and family members not to appreciate change over a short period of time but to value only sustained slowly progressing changes over months. This is an important protection against burn-out.

DAILY LIFE SCENE

Special cases can be very illuminating. In his book “Descartes’ Error”, Antonio Damasio (1994) presents an extensive case study of Elliot, a patient who had a frontal lobotomy as a consequence of a benign but large tumor. The surgery was successful but the patient lost his job and consequently his social network. No problems were found, however, with traditional neuropsychologic tests for frontal lobe psychopathology. Elliot responded — unexpectedly considering the lobotomy — within the normal range and sometimes even more adequately than normal subjects on assessments of emotional monitoring and associative thinking. He was aware of social conventions and the appropriate cognitions, emotions and behaviors in specific situations. He was able to reason the appropriate answers in these assessment tasks. Nevertheless, he was unable to hold a job and his professional and social relations were extremely impaired. Outside of the artificial testing situation — in normal daily

life — he was unable to select and perform the appropriate behavioral responses. In the complexity of daily life his emotional and cognitive self-assessment was defective.

The case is interesting and intriguing because it depicts a symptomatology that bears a similarity to the negative symptom type of schizophrenia that is most refractory to treatment and characterizes the largest part of the chronic mental patient population. As in the case of Elliot, in chronic schizophrenia a marginal handicap in a testing situation has a massive impact on daily life with a full adaptational failure as a result. Cognitions, emotions and behaviors are different in artificial compared to normal situations. We once went on a holiday trip with the chronic mental patients of our ward and observed a behavioral change during the first day. Patients reacted almost “normally” when we visited an art gallery or went to a restaurant. The presupposition that the treatment environment might cause the symptoms — even partially — has haunted psychiatry for years. We, ourselves, have often wondered how to design better residential living environments in which all the available skills of our patients could be preserved.

DISCUSSION

What is the reason for this erratic behavior in schizophrenia? Do schizophrenic patients live in non-stimulating environments or are they just unwilling?

Is the variability of schizophrenic symptomatology the consequence of a *generalization* problem? Some critics of psychiatry have argued that institutional environments drain the purposefulness of behaviors. On a ward the patients have no need to cook, wash dishes or do the laundry. Consequently, patients are not stimulated to use their skills resulting in a severe behavioral atrophy over time. We hope that once they are brought into the appropriate environment, behaviors will be restored. I think this is a naive — technocratic — assumption. In our holiday example after one or more days the lethargy and positive symptoms we knew from the ward behaviors crept in again. This shows that the problem is not a poor generalization or an inappropriate treatment environment, but is due to the volatile character of schizophrenic behavior making lasting behavioral change difficult to attain. New or stimulating environments can facilitate this but offer no structural solution. Of course, the environment is important for the expression of behaviors in schizophrenic patients, but a mere change is insufficient to automatically reactivate lost skills for lasting periods of time. Nevertheless, interventions for schizophrenia — and other illnesses with a different symptomatic spectrum in artificial and normal daily life situations — should be planned in the context in which the skills will be applied. The reason, however, is not the poor generalization of learned skills but the deficiency in the cognitive functions they need to assimilate complex environmental cues. These “higher order” psychological functions should be trained. They are different in a complex and purposeful situation mirroring normal daily life compared to the artificial, abstract learning environment with restricted goals in which we often treat

schizophrenic patients. In other words, the location of the learning environment is not important, but its implicit purposefulness and embedded complexity are.

Is this then a *motivational* problem? When we recognize that most schizophrenic patients are aware of the appropriate behaviors, cognitions and emotions in specific environments, training programs that target these areas will result in an alienating experience and be experienced as childish by our patients. Demotivation is understandable in this situation. Can this demotivation explain the erratic behavioral repertoire of schizophrenic patients? A traditional assessment strategy in psychology stipulates that the maximum skills-level of an individual can be defined based on his or her optimal level of functioning over a specific period of time. This level defines what is theoretically possible given the bio-physiologic restraints of the patient. A schizophrenic patient who demonstrates skills sometimes, shows the functioning level that he can reach. Treatment should enable us to counter the momentary a-motivation. However, in line with our current knowledge of schizophrenia, negative symptoms or a-motivation are not considered a reflection of being “unwilling” but of being “unable”. The intriguing aspects of the varying schizophrenic behavioral expression should therefore be taken seriously as intrinsic parts of the illness.

Abstract treatment programs such as the training of neuropsychologic functions or social skills for schizophrenic patients are not very successful. In contrast, the content of the treatment for schizophrenic patients should be as purposeful as possible because the illness itself is primarily an illness of creating this purposefulness in daily life. This can be realized by avoiding hospital admissions and implementing treatment in the normal living environment of the patients. When — due to the severity of the symptomatology — hospitalization is requested and treatment has to be implemented in an artificial context, the environment itself should create its own purposefulness. Such an intervention can even be more powerful than when implemented in the normal living situation. This is the reason why a Fountain-house model, for instance, is more effective than an industrial therapy unit. This is also why social skills training packages for schizophrenia have moved from training abstract skills as “creating eye contact” towards readily applicable but more complex skills such as “symptom self-management”.

Basic deficits in psychologic functioning undoubtedly exist in schizophrenic patients, but it is difficult to understand how these are related to the failure of daily life adaptation. This shows that the most important missing basic functions of schizophrenic patients are not distractibility or concentration problems but in contrast complex environment processing skills that give meaning to the situation. Often schizophrenic patients are able to reason out appropriate solutions but still get confused when it comes to making the appropriate behavioral choices. This is why schizophrenic patients are often able to recognize their own hallucinations and delusions — they can acknowledge the environmental signals that their perceptions are not shared by others — but they are unable to deacthet the related emotions. We have introduced the notion of “*immediate emotional response*” (IER) to point to the emotional level that drives, for instance, the hallucination-related anxiety. The “*subject’s self-narrative*” (SSN), in contrast, is the reflective cognitive mechanism that assimilates and integrates the available information into a coherent self-image. The SSN and its related

mechanisms can create a delusion when — to the patient — the delusion seems to be the only strategy that restores self-integration. Reflective cognitions — being aware of the patient's self and his or her goals — are part of the SSN mechanisms and crucial to select appropriate behaviors. These “higher order” cognitive functions are often disturbed in schizophrenia. Consequently, patients suffering from schizophrenia have problems generating purposeful behaviors. This can result in behavioral atrophy and even in extreme situations to catatonic stupor. In contrast, when the decision for behavior selection is speeded up and we avoid or minimize the activation of the self-reflecting functions of the SSN, a much more appropriate outcome can be obtained. We can, for instance, speed up this decision, making process by maximizing the purposefulness of a situation making a behavioral selection almost automatic. This is the situation in which we should treat schizophrenic patients. The mechanisms that regulate these processes are not the SSNs anymore, but the IERs.

An appropriate term to describe the illness process in schizophrenia is “emotional aphasia”. Schizophrenic subjects are aware of their cognitions and emotions but are unable to integrate them into a purposeful synthesis. Emotions, therefore, have no appropriate meaning. Diagnostic interventions as well as treatment should focus primarily on the arena of daily life adaptation where environmental processing and experience are combined within a purposeful situation.

Towards an Innovation in Psychiatry

ECOLOGICAL PSYCHIATRY

The Experience Sampling Method (ESM) — described in this study — was developed to assess moment-to-moment emotional and cognitive changes (IERs). It is the best possible assessment technique of these momentary subject experiences. Together with traditional assessment techniques in psychiatry — which are appropriate to assess SSNs — ESM can help us to understand the process of illness and well-being.

Some of the parameters that are computed from the ESM data reflect the assessments based on traditional instruments. Schizophrenic patients report “hearing voices” and normal subjects do not. Anxiety patients are more anxious in daily life and depressive patients more depressed from moment to moment. However, the assessment of moment-to-moment experiences is not meant to validate cross-sectional assessments. It would be very unfortunate if both assessment strategies were validated against each other and attempts were made to merge the result of both approaches. There is a qualitative difference between ESM and traditional assessment strategies such as interviews and questionnaires. This difference is important for our understanding of psychic phenomena and invaluable for the future of psychiatric research. Consequently, we should explore how to maximize the differences between both assessment strategies. I am convinced that it is as inappropriate to use ESM to assess the SSNs as it is inappropriate to use interview and questionnaire techniques to assess moment-to-moment experiences and contextual adaptational strategies (IERs). We

should be more conscious of the content of our research because this leads to the appropriate choice of an assessment strategy. Questionnaires and clinical interviews will continue to generate the best data for psychotherapy because they target the subject's self-image and narrative. In contrast, to measure the effects of a neuroleptic or the processes underneath psychopathologic experience, ESM is more appropriate. It generates a kind of information that was previously unavailable in psychiatry.

Ecological psychiatry — the study of person/environment interactions in psychopathology — was until recently restricted to the study the environmental triggers of relapses. Assessment techniques such as ESM can expand the field of ecological psychiatry to assess symptoms of pathology itself: how patients suffering from different psychopathologies process and interpret their environment. Perhaps, in the future, moment-to-moment experiences will be used to describe an additional set of psychiatric "symptoms" that will lead to new syndromal summaries and can fertilize fundamental research into the heterogeneous etiology of schizophrenia and other mental illnesses.

NEGOTIATED MENTAL CARE

ESM is not designed to focus on psychopathology alone. The content of ESM assessment forms is not restricted to questions related to psychic symptomatology but also contains the mundane aspects of daily life adaptation. Using ESM we have demonstrated that the experiential reality of patients suffering from schizophrenia matches the lives of normal or less ill subjects in many ways.

The development of diagnostic tools that customize interventions in the care for chronic mental patients is warranted as a solution for the poor sensitivity of traditional treatment strategies. Due to its ability to detect between-subject differences in adaptation, ESM is an important innovative assessment tool that allows customizing rehabilitative interventions. ESM can be very valuable in bridging the gap of inequality between the clinician "who knows" and the patient "who does not know". In the traditional therapeutic relation the clinician is the specialist while the patient just waits to hear his verdict — the therapeutic plan. Through ESM the daily life reality of the patient is brought into the clinician's office. The dialogue between the professional, patient and care givers is therefore build on the same information. This results in a "collaborative ethos" in which the therapist, patient and families regard themselves as equal partners in the management of schizophrenia. The motivational potentials of a shared responsibility in the treatment plan cannot be overestimated. There is a great demand for communication and assessment strategies that foster this shared responsibility. As we have argued, this purposeful collaboration allows the optimized use of the patient's skills. The cases treated with the help of ESM demonstrate the need to customize interventions to the dynamics observed in specific patients. In ESM the patient is the specialist in his own life and the data that he has collected targets the interventions. Discussion and the involved questioning of the data makes the patient a partner in his own treatment planning. As a consequence ESM offers a base for a true future *negotiated mental health care*.

Appendices

.....

Appendix 1: ESM protocol

SUBJECT SELECTION

In this study we present data collected over the last 7 years from chronic mental patients who are actually in ambulatory or clinical treatment in Maastricht. In addition we sampled normal control subjects, usually nursing staff. The samples were a convenient group. No attempts were made to obtain a random sample. Subjects for this study were recruited from mental health facilities in the South of the Netherlands. The selected group can be considered a convenience sample from the Maastricht population of chronic mental patients. The whole selection trajectory is described in Figure A1.1. The most important aspect of the procedure is that we selected subjects through their treating professionals. We held instruction sessions in the different clinical teams and invited the mental health professionals to refer patients to our research team. We had no control over the professionals who sympathized with our research project (40%), and those who did not. Furthermore, the ESM procedure was considered a diagnostic instrument and as a consequence only those patients with puzzling diagnostic problems or therapeutic challenges, were referred (drop-out guessed to be more than 50%). On the other hand, the therapists assessed the chances that the subjects would bring this ESM procedure to a positive end. This self-selection by the therapist excluded mentally retarded subjects, subjects with reading and writing difficulties and subjects in an acute psychotic phase or who were considered unstable to stand the stress of the ESM assessment. From the remaining 20% of chronic mental patients approximately 20% refused to participate or were not able or willing to subscribe the informed consent. Sometimes, family members forced them to withdraw their participation. We know from different studies that this figure is not much higher in other psychiatric populations, nor in normal settings. Of course, chronic schizophrenics used different arguments when they refuse to participate. One said he could not participate because the beeping device attracts cosmic signals that control his brain, another argued that a voice forced him to withdraw because — as the ruler of the world — he had other things to attend to. Typical reactions of dropping-out chronic depressive subjects are that it is too much trouble. Dropping-out normal control subjects say that the ESM method is an intrusion in their private life. Although clearly different arguments are used, the method related drop-out is usually not higher than 25% in most groups.

Figure A1.2 Subject selection and “step by step” “drop-out” trajectory.

In summary, we sampled 16% of the chronic mental population in Maastricht. Considering the fact that most of the drop out in the trajectory are due to a very conservative therapist selection ($60\% + 50\% \times (100 - 60\%) = 80\%$) and refusal of informed consent (20%) and only 25% dropped out as a consequence of ESM-related problems, we are confident that it is possible to sample up to 75% of the chronic psychiatric population satisfactory. To reach this figure we need support of the treatment staff, invest in a personal therapeutic relation by coaching the subjects intensively on a day-to-day basis and allow postponing the sampling period for some weeks during extremely psychotic episodes. In our sample, a small number of subjects were assessed during treatment in seclusion rooms or when apparently mute — catatonic.

ESM PROCEDURE

The clinicians introduced the method briefly to the patient and checked if he or she was willing to participate in the study. If they agreed the research assistant made an appointment with the patient for the next week. This appointment is the “briefing” session. After that session, the patients collect ESM reports for a whole week. Consequently, they have a last contact to debrief the patient. The whole procedure is described extensively in Chapter 3.

Figure A1.2 Procedure of the ESM data gathering in schizophrenic subjects.

Data-processing could be completed within a week of two. When the ESM data was collected with a clear clinical question in mind, a diagnostic feed-back was given to the treating clinician who stayed in charge of the whole therapeutic process.

RESPONSE DISTRIBUTION

Table A1.1 gives an overview of the response distribution found in different ESM studies.

	Students ⁸¹	Chronics ⁸²	Depression ⁸³	Anxiety ⁸⁴	Stress ⁸⁵
Missed Beeps	33.0%	32.9%	18.7%	33.3%	36.7%
Skipped Beeps	4.0%	9.1%	3.8%	8.7%	7.3%
Skipped Subj.	4.8%	15.8%	4.3%	7.0%	2.2%
	Students	Chronics	Depression	Anxiety	Stress
Students	-	0.90	0.86	0.75	0.58
Chronics	-	-	0.91	0.73	0.82
Depression	-	-	-	0.48	0.86
Anxiety	-	-	-	-	0.32

Table A1.1 Overview of drop out at the subject and beep level (with reference samples).

Irrespective of the population the distributions are comparable ($\chi^2_{(48)}=0.607$, n.s.). Overall, the frequencies of beeps responded upon before the actual beep is below 10% with a very low frequency for stressed individuals (2%) and a higher frequency in the other groups (from 6 to 9%). These are not all invalid beeps as we argued in the discussion on the ESM reliability time window in Chapter 6. All groups have a high response frequency of immediate responses (from 9 to 28%). This is too quick to be real — two minutes are needed to fill in the ESMs — and is probably caused by retrospective guesses of the beep time. The anxiety subjects were the “quickest” responders. Here obsessive time checking might play a role. There is a slight drop for the 1 minute interval. A second peak evolves homogeneously over the 2 — 5 minute range. The next category, although still high in some groups, is a combination code which might fool the eye. In all the groups, the frequency of responses diminishes after the 2-3 minutes peak. This process, however, is smoother for the stressed individuals who are the slowest to respond (Figure A1.3).

⁸¹ Analysis using the students of the RIR/ESM study described in Chapter 6 (Delespaul & Reis, 1994).
⁸² Analysis using chronic mental patients sampled for this study (Chapter 9 - 11).
⁸³ Analysis using depressive subjects sampled in a general practitioners practice by Barge (1994).
⁸⁴ Analysis using subjects with DSM-III-R 300.01 and 300.21 diagnoses (Dijkman-Caes, 1993).
⁸⁵ Analysis using subjects of a white-collar stress protocol by the Van Eck & Nicolson (1994).

Figure A1.3 Response distribution: interval between beep and response (with smoothed insert).

Using the criterion of the reliability window (Chapter 6B) — from -5 and +15 minutes between beep and response — the number of “unreliable beeps” was 4% for the depressive subjects and between 7 and 9% for the other groups. The number of subjects that were skipped as unreliable using a different criterion — at least 1/3 beeps valid — was low (< 7%) for most groups, but double as high in the chronic mental patient group. Considering the formal thought invalidity of this group, this higher percentage is not a surprise.

.....

Appendix 2: ESM sampling device

SAMPLING DEVICE

In ESM, subjects have to know when to fill out the ESM questionnaires. In order to allow the subjects not to bother checking their watches continually, which would lead to a change in their daily life activity patterns and cognitions, the subjects carry a signaling device. In this study, we used a commercially available device, a SEIKO RC-1000⁸⁶ digital watch that is pre-programmed by a personal computer to signal the subjects to fill out the assessment forms. We did select a small device that could easily be carried around and that gives a discrete digital sound. An additional advantage of this watch is that subjects used a standard reference to check on the time. This gives us additional control over the elapsed time between beep and report. As a preventive measure, we changed batteries once a year.

Figure A2.1 SEIKO RC-1000 that is used as the ESM sampling device.

A section in Chapter 4 is devoted to the selection of the appropriate sampling device. We refer to that chapter for an in depth discussion of the different devices. It is important to mention that — although we favor the watches — those are out of production for years and the alternatives that are available today are often not very appropriate for our purpose.

⁸⁶ The SEIKO RC-4000 and RC-4500, but also the CASIO BGP-20 are alternatives.

INSTRUCTION

The watches could have technical problems and the instruction during the briefing session is often forgotten by the patients directly after the meeting. Therefore, a "reminder"- instruction for the watches (and the ESM procedure in general) was included in each ESM booklet. The text was:

IMPORTANT NOTES

- Fill out all the questions. When you are alone, say so. If you are at home, check a 0 (zero) for the number of kilometers you are from home. If you consumed nothing between beeps, check on the "nothing"-entry.
- Follow the order of questions as they appear in the questionnaires. Avoid skipping entries.
- Never go back to previous assessment forms. We are interested in your current emotions, cognitions, activities. You can check on the "complaints" that were written down on the inside front cover.
- Don't wonder to long on each question. We are interested in your spontaneous answers (don't think what you think is unimportant — don't be shy — you're answers will be kept confidentially).
- When open ended answers are requested (What am I thinking? Who am I with?) we ask you to be short and concrete. You don't have to explain thinks to us, just use you're own language as you would normally do: use for instance first names to describe persons. If you feel the need to write down diary kind of memo's, use the note-areas.
- On the one to seven scales, check only one number...
- Fill in the forms as soon as possible after the beep, preferably within one minute...
- When you're in a place, where the sound of the beep would be disturbing (a church, a classical concert,...) you can switch it off. Remember to switch it on again immediately afterwards. You can switch the beeping watch on and off by simultaneously pressing both colored buttons. A "#" will appear in the upper left corner of the display when the beep is "on".
- The watches are not water-proof. Don't keep them on when showering, washing dishes,...
- When the watches become difficult to read, or you see a message on it, always press the "TIME"-button. It will make them operational again.
- If you received no beep in half a day, contact us.

If you encounter any problem, don't postpone to contact us:

John Johnson (name of the briefing person)
working days and times: tel.
other moments tel.

Figure A2.2 Checklist of the briefing instructions that is printed in each ESM booklet.

Appendix 3: ESM sampling scheme

THEORETIC STARTING POINT

The sampling scheme that was used in the ESM data collection procedures was a random interval schedule, stratified within fixed interval periods. This schedule is an optimal choice between fixed and random schedules. It has all the advantages of a random schedule (unpredictability and representative sampling), without the disadvantages (unequal spread of observations). We sample between 7:30 in the morning and 10:30 in the evening. This gives us 10 equal interval time blocks of 90 minutes each. A random moment within each of those moments is generated.

Figure A3.1 Time sampling schedule used in the ESM procedures.

DESCRIPTION OF ACTUAL SCHEDULES

To avoid confounds as a result of the fact that a random scheme was pre-programmed in the watches — therefore only pseudo-random which could result in capitalizing on the specificity of the random generation, for instance, a high number during the 8 o'clock news — we worked with a set of 20 different schedules. Because a schedule lasts for a whole week each subject got the same schedule for the whole ESM period except when the watch unexpectedly broke down during that week (which happened only once). Table A3.1 gives an overview of the most important time series and their characteristics.

Schedule	frequency	$\Delta T_i / T_{i+1}$: $\bar{X}$ (s.d.)	<30	30-60	60-90	90-120	120-150	150-180
AF004	4 (2%)	89.78 (17.33)		2	29	29	3	
AR001	14 (7%)	90.14 (31.64)	1	11	19	18	12	2
RR001	59 (31%)	88.73 (29.90)	2	8	21	21	10	1
RR002	11 (6%)	90.68 (31.56)	1	12	16	20	12	2
RR003	16 (8%)	91.38 (26.81)	1	7	24	22	9	
RR007	53 (28%)	88.25 (31.51)	1	16	13	23	8	2
RR010	13 (7%)	88.21 (36.51)	3	11	22	14	7	6
RR011	4 (2%)	90.81 (33.47)		13	16	21	1	
$\bar{X}$ $\bar{X}$ (s.d.)		89.75 (1.23)	2%	16%	32%	33%	15%	3%
Min. - Max.		88.21 - 91.38	(3%)	(15%)	(32%)	(32%)	(15%)	(3%)
$\bar{X}_{sd}$ (s.d.)		29.84 (5.76)						
Min. - Max.		17.33 - 36.51						

Table A3.1 Reference data of the beep schedules (subset of schedules used by ≥ 4 subjects). Right:: realized frequency distribution of intervals (theoretical proportions between parentheses).

Randomization was successful in generating time schedules with a satisfying between beep interval of 90 minutes for all series. Also, the intervals had a large spreading from one moment to the other. Looking at the distribution of the intervals, we see an unequal spread. This is a consequence of the fact that, while each moment within an interval had an equal change to become the beep moment, the interval will be a result of conditional probabilities of the two consecutive beeps. To find differences between 150 and 180 minutes, the beep of the first period should fall in the first third and the one of the next interval in the last third. All frequencies match the theoretical expectations except for the first interval, which is slightly lower than expected due to the fact that intervals of less than 15 minutes were not allowed.

Due to logistic restrictions (briefing and debriefing session are held on the same day) we sampled for six days, in which a whole week-end was always included. This resulted in a over-representation of the week-day information in the results (2 out of six instead of 2 out of 7). To correct for this fact, time budgets that are different in week-end and week-days should be corrected using a weighted average procedure:

$$A\% = \frac{1}{7} \times (5 \times \frac{5}{4} A_{w\%} + 2 \times \frac{2}{2} A_{we\%})$$

with: $A_{w\%}$: the time budget category for the week-days
 $A_{we\%}$: the time budget category for the week-end-days

RESPONSE DISTRIBUTION

i. Time of day

Figure A3.2 Mean frequency distribution responses at different moments of the day (max. 6).

The subjects of our sample (n=200) responded significantly differently to the beeps ($F_{(9,1713)}=10.758$, $p<.005$). At the average, one third of the beeps were missed for all the moments of the day. However, significant higher numbers of missing data were found for the first beeps of the day — when the subjects were still asleep and the last beep of the day when the subjects were already to bed. Attempts to model the data showed the best fit in the 4th degree polynomial that also described the slightly higher response frequencies in the 10:30-13:30 interval (beep 3 and 4).

ii. Day of the week

There was no significant decay of response frequencies over the seven days of the week ($F_{(6,1017)}=0.235$, n.s.). The decay that was found in the beginning of the day was not related to the week-end days. For this population of chronic mental patients, all days of the week are the same.

iii. Day within the sampling period

There was no significant decay of response frequencies over the six days of the sampling week ($F_{(5,1006)}=0.623$, n.s.). The length of sampling week was not tiring. We also checked on the responses over the six sampling days. If the data would be influenced by tiredness, the responses would be less extreme from the first day to the next ones, while the variability would diminish over time (Figure A3.3).

Figure A3.3 Hypothetical presentation of poor ESM response quality over time.

We illustrate the trends using the pathology data (item 1-8). None of the items showed a drift over time. The daily means were the same from day 1 to day 6. However, we found significant differences in the variability (by subject) over the consecutive days of sampling. The variability was highest the first days and decayed over time. Approximately after two days, the variability stabilized and stayed at the same level for the next 4 days (for 4 from 8 pathology items).

Pathology Item	Drift (mean/day 1-6)		Reduced sensitivity (variability/day 1-6)								$\bar{x}$ s.d.
	Test	$\bar{X}$ $\bar{X}$	Test	1	2	3	4	5	6		
suspicious	$F_{(5,992)}=0.10$ n.s.	1.97	$F_{(5,955)}=1.66$ n.s.								0.63
difficult to express	$F_{(5,856)}=0.17$ n.s.	2.51	$F_{(5,824)}=5.50$ ***	1.22	0.87	0.83	0.77	0.68	0.72		
obsession	$F_{(5,856)}=0.70$ n.s.	2.77	$F_{(5,820)}=5.97$ ***	1.25	1.01	0.81	0.76	0.75	0.89		
paranoia	$F_{(5,855)}=0.11$ n.s.	2.46	$F_{(5,822)}=2.86$ *	1.15	0.94	0.87	0.86	0.74	0.82		
de realization	$F_{(5,948)}=0.65$ n.s.	1.91	$F_{(5,912)}=2.13$ *	0.66	0.50	0.46	0.50	0.37	0.47		
hearing voices	$F_{(5,896)}=0.20$ n.s.	1.79	$F_{(5,860)}=1.23$ n.s.								0.40
seeing things	$F_{(5,477)}=0.23$ n.s.	1.64	$F_{(5,457)}=0.19$ n.s.								0.41
fear losing control	$F_{(5,568)}=0.19$ n.s.	2.10	$F_{(5,546)}=0.97$ n.s.								0.56

Table A3.2 Empirical data ESM response quality over time (day 1 to 6).

Because this reduced variability over time is restricted to some items, our hypothesis is that subjects have difficulties positioning themselves upon the scales. After some time they understand that more extreme positions are possible and therefore restrict themselves to a more restricted range for their personal normal experience. This would explain this reduction of variability that stabilizes already after the second day.

Figure A3.4 Empirical presentation of ESM response quality over time (from day 1 to 6).

.....

Appendix 4: ESM booklets

A. ESM BOOKLETS

The ESM assessment forms are bundled in pocket size booklets that can easily be carried. Due to the amount of assessment sheets, it is impossible to bundle all the sheets for the whole assessment week (60 sheets) in one booklet. Therefore we decided to bundle them by day and give each subject 6 booklets. The bundling by day allows us to use the cover sheets for the day level assessments. In the inside front cover sheet we have printed the assessment of the sleep quality for the past night. The back cover sheet is used to assess the past day. The questions in the ESM Forms can be divided in Likert type scales and open ended questions that are coded afterwards by the research assistant.

When the beep went off, I was thinking about?.....

This thought is... Not Medium High

Pleasant 1 2 3 4 5 6 7

Clear 1 2 3 4 5 6 7

Normal 1 2 3 4 5 6 7

I have trouble concentrating 1 2 3 4 5 6 7

I feel... Not Medium High

Cherful 1 2 3 4 5 6 7

Unsure 1 2 3 4 5 6 7

Lonely 1 2 3 4 5 6 7

Relaxed 1 2 3 4 5 6 7

Anxious 1 2 3 4 5 6 7

Satisfied 1 2 3 4 5 6 7

Irritated 1 2 3 4 5 6 7

Sad 1 2 3 4 5 6 7

Guilty 1 2 3 4 5 6 7

My 1st complaint is present Not Medium High

My 2nd complaint is present 1 2 3 4 5 6 7

My thoughts are... Not Medium High

suspicious 1 2 3 4 5 6 7

difficult to express 1 2 3 4 5 6 7

I can't get rid of them 1 2 3 4 5 6 7

influenced by other 1 2 3 4 5 6 7

I feel unreal Not Medium High

I hear voices 1 2 3 4 5 6 7

I see things 1 2 3 4 5 6 7

I fear to loose control 1 2 3 4 5 6 7

Where am I?

I'm alone? yes / no

If not, with whom?

How many men?

Women?

Children?

I like this company Not Medium High

I prefer being alone 1 2 3 4 5 6 7

We are acting together 1 2 3 4 5 6 7

What am I doing?

Na Medium High

I prefer doing something else 1 2 3 4 5 6 7

I find myself active 1 2 3 4 5 6 7

It consumes my energy 1 2 3 4 5 6 7

This activity is a challenge Na Medium High

I'm skilled at it 1 2 3 4 5 6 7

I'm hungry Na Medium High

I'm tired 1 2 3 4 5 6 7

I don't feel well 1 2 3 4 5 6 7

I LAY DOWN / SIT / STAND / RUN / SPORT (Circle your choice)

Since I woke up I used... (more choices possible)

O NOTHING When was that?..... hours

O ALCOHOL When was that?..... hours

O MEDICATION When was that?..... hours

O COFFEE O TABACCO O

Since I woke up the most important thing that happened was:

This was: very unpleasant -3 -2 -1 0 1 2 3 very pleasant Na Medium High

This beep was disturbing 1 2 3 4 5 6 7

It is now exactly hours..... minutes

Remarks:

Do not write below this line

DayNo Online Where Event/Story

BeepNo Calligraphy What1 Even/What

Goods What2 Even/Who

ThinkTime Who1 Medic1

ThinkRetain Who2 Medic2

ThinkValue Medic3

Figure A4.1a Format of the ESM booklet: design of ESM forms (booklet 063).

[1] Booklet "060"

The first format of the ESM booklet used with chronic mental patients was designed according to the format that has been used in other studies in Maastricht (e.g. Dijkman-Caes, 1993). It has a modular format with standard set of items for cognitions, moods, activity motivation, physical complaints and substance use. In addition, symptom assessment was made possible through a so-called psychopathology module. Separate modules were designed for the most important diagnostic categories.

[2] Booklet "161"

The second booklet was designed for a study on the effect of work rehabilitation. It allowed a better assessment of experiences during activities and could be used to customize the selection of work offers using the subject's best mental state. To allow this, a number of additional items were added to assess psychopathology. The changes were only piloted in 10 subjects chronic mental patients of which only a small number were schizophrenic subjects.

[3] Booklet "063"

Part of the new activity related items that were targeted in booklet 161, were included in booklet 063. The new psychopathology related item formulations of booklet 161 were less successful and we went back to the original item-set (with the addition of visual hallucinations). In addition, the item set on moods was slightly modified to match the actual item set of the other ESM studies (and allow comparisons). Important additional changes are the inclusion of assessments by the subject of the social environment. We also added a new open ended question that assesses the occurrence of salient events between beeps. The subjects describe those events and rate if they influenced the subject positively or negatively. Figure A4.1b illustrates the format of the additional questions asked in the morning when the subjects wake up and in the evening when they go to sleep. The questions in the morning are primarily an assessment of the sleep quality in the previous night. The two entries for the complaints are reminders. The same idiographic complaints should be kept during the whole sampling week. The questions for the evening assessment are related to day structuring activities (Zeitgebers) such as meals and work. To avoid possible confusion with the time coding we use a "yes/no ath.....m"-format in the most recent booklets (063/064). This allows us to make a distinction between when a subject has not eaten or not answered. We also to assess ESM reactivity. We ask subjects to report when they switched off the beeping device and to assess if the ESM procedure altered their mood or changed their activities during the day.

Good morning

Fill in the next question when you wake up:

Date?d.....d.....

How late is it now?hrs. min.

How long did it took before I slept last night?hrs. min.

How often did I wake up? times.

I slept well

Not
1 2

Moderate
3 4 5

Very
6 7

My first complaint is:

My second complaint is:

Do not write behind this line

Subject number

Day number

Time Series Code

Booklets Code

Codebook version

063

3

I switched off the watch

From: h. m. to h. m. because

From: h. m. to h. m. because

Answer the next questions just before you go to sleep:

It was a normal day
The filling in of the booklets
influenced my day
Without the booklets I would
have done other things

Not
1 2

Moderate
3 4 5

Very
6 7

I had breakfast

no / yes

at h. min.

I had lunch

no / yes

at h. min.

I had dinner

no / yes

at h. min.

I went to work

no / yes

from h. min.
to h. min.

I go to sleep now...

It is exactly h. min.

Good night!

Figure A4.1b Morning (left) and evening (right) ESM forms (booklet 063).

[4] Booklet "064"

The last booklet — number 064 — contains a subset of the number 063 version and was developed concurrently. 7-point Likert scales were substituted by “yes/no” scales. The booklet is used for those subjects who were unable during the briefing session, to assess the 7-point Likert scales.

What were you thinking before the beep went off?

This thought is...

Pleasant
Clear

no / yes
no / yes

I feel...

Cheerful
Lonely
Relaxed

Anxious
Irritated
Insecure

no / yes
no / yes
no / yes

no / yes
no / yes
no / yes

My primary complaint bothers me

no / yes

I hear voices
I see things

no / yes
no / yes

I'm aggressive
I'm suspicious

no / yes
no / yes

With whom am I?

How many men? Women?

I prefer being alone

no / yes

Where am I?

What am I doing?

I would prefer doing something else

no / yes

I'm hungry
I'm tired

no / yes
no / yes

Since the last beep I had ...
0 NOTHING 0 ALCOHOL
0 MEDICATION 0 COFFEE
0

This beep disturbed me

no / yes

The current time is

..... hrs.. min.

Notes

Do not write behind this line

[coding information]

Figure A4.1c Special booklet (booklet 064) for subjects who have problems with the Likert scales.

B. DROP OUT BY BOOKLET TYPE

The first booklet in use in the schizophrenia program was “060”. This booklet had a low overall effectiveness of 65% which was due to the lack of routine in the application of the protocol in the beginning of the study. Booklets “161” and the actual booklet “063”, reached a much higher effectiveness. Only 15% of the subjects who are recently been sampled (booklet “063”), did not reach the very conservative criteria applied in this study. As can be seen the “success” of booklet “064” is much lower but this is a special booklet, designed for those mental patients who showed an inability to perform adequately with the 7-point Likert scales during the briefing session. It is much simplified and uses a yes/no-substitute for the ratings. Using this simplified booklet we could sample 5 out of 8 subjects adequately. Those subjects would have been lost otherwise. On the other hand, it is interesting to note that most of the schizophrenic subjects (83%) were able to understand the ESM-forms and the abstract 7-point Likert scales. A previous attempt in the pilot phase using semantic differential scales, however, was unsuccessful. A large number of schizophrenic subjects became confused by the cues at each pole of the scales. They gave two answers to the same scale, hereby mislabeling the dimensional nature of it and trying to answer the both cueing items independently as if they were actually two different scales.

Count (% Valid)	Schizophrenia	Depression	Mental Health C.	Normal Control	Total
Booklet "060"	30 (65%)	6 (60%)	5 (38%)	17 (85%)	58 (65%)
"161"	3 (60%)	5 (100%)	0 (0%)	5 (83%)	13 (81%)
"063"	20 (83%)	12 (86%)	6 (75%)	8 (100%)	46 (85%)
"064"	4 (67%)	0 (0%)	1 (50%)	0 (0%)	5 (63%)
Column Total	57 (70%)	23 (79%)	12 (52%)	30 (88%)	122 (73%)

Table A4.1 Subject drop-out by “booklet”.

C. DAY LEVEL QUESTIONS

i. Assessments in the morning

		1	2	3	4 ⁸⁷
Date?		*	*	*	*
Sleep	Wake up time?	*	*	*	*
	Sleep delay? (open)	*	*	*	
	Number of times up? (open)	*	*	*	
	Sleep Quality? (Likert)	*	*	*	*

ii. Assessments in the evening

		1	2	3	4
	Normal day? (Likert)	*	*	*	
	ESM was disturbing today? (Likert)		*		
	ESM influenced activities? (Likert; o=open)	*	o	*	
	ESM influenced moods? (Likert; o=open)		o	*	
	Breakfast time?	*	*	*	*
	Lunch time?	*	*	*	*
	Dinner time?	*	*	*	*
	To work time?	*	*	*	
	From work time?	*	*	*	
Disconnect	From/to time?		*	*	
	Reason?		*	*	
	To bed time?	*	*	*	*

D. BEEP LEVEL QUESTIONS

i. Item comparisons of the Likert type scales

The items used in the different versions of the “schizophrenia”-booklet are presented. The stars indicate the availability of an item in the version of the booklet. The “+”, “-” and “.”-symbols indicate the weight values that are attributed at the different scales in the standard a priori summations e.g. in the “Quality of Life”-assessments of the personalized rehabilitation protocol. The numbers in the front indicate the item reference code in the so called “Basic Properties” item pools⁸⁸.

Thought Scales (BP-[Data] item set)

		1	2	3	4
01	Pleasant Thought	+	*	*	*
02	Clear Thought	+	*	*	*
03	Racing Thoughts	-	*		*
04	Normal Thought	+	*	*	
05	Confusing Thought	-		*	
08	I'm concentrated	+		*	

⁸⁷ 1= booklet “060”; 2= “161”; 3= booklet “063”; 4= booklet “064”.

⁸⁸ Booklets are composed by selecting items from sets: general ‘data’-items (thoughts, moods, physical feelings and activity motivation), the ‘social interaction’ (self assessment of the social interaction); ‘events’(assessment of between beep events); and ‘use’ (set of psychotropic substances). For different disease entities we have an additional set of items that target the specific psychopathology.

Mood Scales (BP-[Data] item set)

		1	2	3	4
11	Cheerful	+	*	*	*
12	Unsure	-	*	*	*
13	Lonely	-	*	*	*
14	Relaxed	+	*	*	*
15	Anxious	-	*	*	*
16	Angry/Irritated	-	*	*	*
19	Sad	-	*		

		1	2	3	4
20	Satisfied	+	*	*	
21	Listless/Apathetic	-	*		
24	Gloomy	-		*	
26	Guilty	-		*	
49	Self-Assured	+	*		

Activity Scales (BP-[Data] item set)

		1	2	3	4
27	Prefer something else	-	*	*	*
28	I'm active	+	*	*	*
29	I'm in control	+	*		
30	I've trouble concentrate	-	*		
31	I'm focused	+		*	
32	I'm skilled	+		*	*

		1	2	3	4
33	I like to do this	+	*		
34	This costs energy	-	*	*	
35	Stimulating environment	+		*	
36	I'm forced to do this	-	*		
68	This is a challenge	+		*	

“Physical” Scales (BP-[Data] item set)

		1	2	3	4
38	I'm hungry	.	*	*	*
39	I'm tired	.	*	*	*
41	I don't feel well	.	*	*	
42	I feel miserable	.	*		

Social Interaction Scales (BP-[Social Interaction] item set)

		1	2	3	4
01	Prefer being alone	-		*	*
02	Like this company	+	*	*	*
03	Interacting	+	*	*	*

Pathology Scales (BP-[Pathology / Schizophrenia subset] item set)

		1	2	3	4
01	Suspiciousness	-	*	*	*
02	Thoughts are difficult to express	-	*	*	
03	I can't get rid of my thoughts (obsession)	-	*	*	
04	Thoughts are influenced by other people	-	*	*	
05	I feel unreal (de realiz.)	-	*	*	*

		1	2	3	4
06	I hear voices (auditory hallucination)	-	*		*
07	I see “things” (visual hallucination)	-	*	*	*
08	I fear to loose control	-	*	*	*
09	I feel helpless	-	*		
10	Thoughts coming from outside	-	*		

Use Scales (BP-[Use] item set)

		1	2	3	4
01	Nothing	*		*	*
02	Alcohol	*	*	*	*
03	Medication	*	*	*	*

		1	2	3	4
04	Coffee	*	*	*	*
05	Tobacco		*	*	
06	Other	*	*	*	

Disturbance

		1	2	3	4
01	Disturbing beep	*	*	*	*

ii. Open ended questions

“Thought” assessment

	1	2	3	4
What am I thinking?	*	*	*	*

“Context” assessment

	1	2	3	4
What am I doing?	*	*	*	*
Who am I with?	*	*	*	*
Am I alone (Y/N)?	*	*	*	
Number of persons?		*		
Number of men	*		*	*
Number of women	*		*	*
Number of children	*		*	
Are you acting together (by person) ⁸⁹		o		
Where am I?	*	*	*	*
How many km’s from home	*			
How late is it?	*	*	*	*

“Event” assessment (Between now and the previous beep...)

	1	2	3	4
Did something important happen?		*	*	
Did I had problems with my complaints ?		*		
Which complaint? (o=open)		o		
Other substances used?	*	*	*	*

E. OTHER BOOKLETS: PATHOLOGY MODULES

In different other ESM studies we have used a set of item-formulations to assess the core psychopathology. These sets are called “pathology”-modules. They were defined using expert panels or a focus group technique involving the patients themselves.

<i>Anxiety</i> (Dijkman)	shortness of breath	L7	<i>Pain</i> (Portegijs)	I'm a helpless case	L7
	palpitations	L7		I should slow down	L7
	dizziness	L7		I feel handicapped	L7
	feeling unreal	L7		I'm not self-confident/weak	L7
	fear of dying	L7		I should have more energy	L7
				This day is lost	L7
<i>Depression</i> (Meertens/ Kraan)	[extended mood list]		<i>Stress</i> (Van Eck/ Nicolson)	stressful event	L7
	[event assessment — see stress]			description	O
				start/end time	O
<i>Drugs (craving)</i> (Kaplan)	thinking about using	L7		annoying	L7
	being in control	L7		expected	L7
	feeling restless	L7		important	L7
	need dope urgently	L7		influence result	L7
		L7		replication	L7+0

Table A4.1 Pathology modules in different studies (L7: Likert type scale/ O: open ended question).

⁸⁹ This question was asked for all the individual persons present at the moment of the beep. It appeared to be a needlessly complex item.

Appendix 5: Codebook information (including re-coding)

A. SUBJECT LEVEL CODES

Demographic codes

• Sex

i. Original Basic Properties Codes:

1: Man; 2: Woman; 9: NA

ii. No recoding

• Age

i. Original Basic Properties Codes:

Age in years.

ii. Recoding

AGECAT: "Age by category":

20= [20-29]; 30= [30-39]; 40= [40-49]; 50= [50+]; 9=NA (9)

• Life circumstances

i. Original Basic Properties Codes:

1	Alone	5	Special Psychiatric Environment
2	With partner/family	6	Special Non Psychiatric Environment
3	With parents/relatives	7	Other
4	Lodging	9	NA

ii. Recoding

LIFECIRC: "Living circumstances":

1= Alone (1); 2= Family (2, 3); 3= Health Care Environment (5); 4= Other (4, 6, 7); 9=NA (9)

• Work

i. Original Basic Properties Codes:

1	Household	7	Reintegrating
2	School/Education	8	Unemployed
3	Regular Job/Employed	9	Sheltered Work
4	>3 Months Sick	10	Unpaid volunteer
5	Pensioner	11	Other
6	Unfitted for work	99	Not Available

ii. Recoding

WORK1: "Location of daily activities":

1= Home (1, 4, 5, 6, 8); 2= Out (2, 3, 9, 10); 9=NA (99)

WORK2: "Salary through daily activities":

1= No (1, 2, 4, 5, 6, 8, 10); 2= Yes (3, 9); 9=NA (99)

WORK3: "Day structure by daily activities":

1= No (4, 5, 6, 8); 2= Yes (1, 2, 3, 9, 10); 9=NA (99)

Psychopathology (cross-sectional)

• BPRS “Brief Psychiatric Rating Scale”

i. Original Basic Properties Codes:

In this study we used two different versions of the BPRS over time (1 Overall & Gorman, 1962; 2 Lukoff et al., 1986; Ventura et al., 1993a; 1993b). We first describe the item definition and integration of the both versions. In all situations we used a 7-point Likert scale with the code 1 as the “not applicable”-code, codes 2 - 7 growing levels of intensity (using the Lukoff et al. (1986) codebook) and 9 as “not scored”. In addition I add the a-priori item selection that I used to define the positive (“P”) and negative (“N”) itemsets. Factor analytic studies of the BPRS were not available to me at that moment. In comparison with the work of Dingemans et al. (1983; 1994) who studied the factor structure of both BPRS versions the thought disorder items loaded on the positive item clusters. The items that load on the positive factor (D_p) for both versions were “suspiciousness”, “unusual thought content”, “hallucinations” and “conceptual disorganization”. Only two items loaded on the negative factor (D_n) for both versions: “blunted affect” and “emotional withdrawal”. In addition they also found a “depression”, “mania” and “disorientation”-factor.

	1	2	P	N	D _p	D _n		1	2	P	N	D _p	D _n
01 Somatic concern	*	*					13 Motor retardation	*	*		*		*
02 Anxiety	*	*					14 Blunted affect	*	*		*		*
03 Depression	*	*					15 Tension	*	*				
04 Guilt	*	*					16 Mannerism & posturing	*	*				*
05 Hostility	*	*	*		*		17 Uncooperativeness	*	*	*		*	*
06 Suspiciousness	*	*	*		*		18 Emotional withdrawal	*	*	*		*	*
07 Unusual thought content	*	*		*	*		19 Suicidality	*					
08 Grandiosity	*	*	*		*		20 Self-neglect	*					(*)
09 Hallucinations	*	*	*		*		21 Bizarre behavior	*			*		
10 Disorientation	*				*		22 Elevated mood	*					
11 Conceptual disorganization	*	*			*		23 Motor hyperactivity	*					
12 Excitement	*						24 Distractibility	*					

Factor analysis codes: *: only in the 16-item version; *: only in the 24 item version (parenthesis for items that also load on another factors); #: in both versions.

ii. Recoding

BPRSMSEV: “Severity using BPRS items (min. 1 / max. 7)”: $\bar{X}$ of non-blank BPRS items.
BPRSSEVC: “BPRSSEV-code by category”:
1= Low (1.00 - 1.99); 2= Medium (2.00 - 2.99); 3= High (3.00+); 9=NA (9)
POSSYMP (PS): “BPRS Positive symptom score”:
BPRS[05] + BPRS[06] + BPRS[08] + BPRS[09] (only if all items are non-blank).
NEGSYMP (NS): “BPRS Negative symptom score”:
BPRS[07] + (BPRS[13] + BPRS[18])/2 + BPRS[14] + BPRS[17] (only if all items non-blank).
DIFPONE: “Positive/Negative Symptom score (Δ of BPRS item subsets)”: $X = \text{Round}((PS - NS)/4) - 1$ = Negative Symptoms ($X < 0$); 0= Unspecified ($X = 0$); 1= Positive Symptoms ($X > 0$); 9=NA (9)
POSNEGR: “Positive/Negative category-score”
 $X = 9$ (high score criterion: four items with mean item score > 2.00)
0= Low Symptoms ($PS < X$ and $NS < X$); 1= Positive Symptoms ($PS \Rightarrow X$ and $NS < X$);
2= Negative Symptoms ($PS < X$ and $NS \Rightarrow X$); 3= High Symptoms Mixed ($PS \Rightarrow X$ and $NS \Rightarrow X$); 9=NA (9)

• SFCL “Social Functioning Checklist”

The Social Functioning Checklist is a small, five-item list, rated using a 7-point Likert scale with 1 = extreme low level functioning and 7 = optimal adaptation and functioning.

i. Original Basic Properties Codes:

- 01 General Level of (Social) Functioning
- 02 Functioning in Work and Productivity
- 03 Functioning in Independent Living and Self Care
- 04 Functioning in the Direct Social Network
- 05 Functioning in the Remote Social Network

ii. Recoding

SFCLMSEV: "Severity using mean SFCL item scores (min. 1 / max. 7)":
Mean of non-blank SFCL-items.

SFCLSEVC: "SFCLSEV-code by category":

1= Low (1.00 - 1.99); 2= Medium (2.00 - 2.99); 3= High (3.00+); 9=NA (9)

- **SCIZTYPE**

The Type code is based on the fourth digit of the DSM-III-R code for schizophrenia.

ii. Recoding

SCIZTYPE: "Code of phase using the 4th digit of the DSM-III-R Axis I code": X=295.xN

1= Disorganized (X=1); 2= Catatonic (X=2); 3= Paranoid (X=3); 4= Schizophreniform (X=4); 6= Residual (X=6); 7= Schizo-Affective (X=7); 9= Undifferentiated (else)

- **PARANOIA**

The Paranoia code is based on the fourth digit of the DSM-III-R code for schizophrenia.

ii. Recoding

PARANOIA: "Code of paranoia using the 4th digit of the DSM-III-R Axis I code": X=295.xN

0= No Paranoia (X<3); 1= Paranoia (X=3)

- **SCIZPHAS (n / category is to low)**

The Phase code is based on the fifth digit of the DSM-III-R code for schizophrenia.

ii. Recoding

SCIZPHAS: "Code of phase using the 5th digit of the DSM-III-R Axis I code": X=295.Nx

0= Unspecified (X=0); 1= Subchronic (X=1); 2= Chronic (X=2); 3= Subchronic + Acute (X=3); 4= Chronic + Acute (X=4); 9= NA (else)

- **GROUP**

The Group code is based on the DSM-III-R Axis I codes.

ii. Recoding

GROUP: "Code of the DSM-III-R Axis I diagnosis":

1= "Chronic Schizophrenia" (295.xx; 297.10; 298.80; 298.90); 2= "Chronic Depression" (296.xx; 300.40; 309.00); 3= "Chronic Control Subjects" (xxx.xx) 4= "Normal Controls" (no DSM-III-R code on Axis I and 2)

B. DAY LEVEL CODES

We haven't developed a specific codebook for the Day-level data. All data are analyzed in their raw data-form as recorded by the subjects.

C. BEEP LEVEL CODES

Thought codes

- *Online*

This item assesses the relation between thought content and activity. It is not intended to be a measurement of psychopathology, but reflects the (normal) cognitive detachment of current activities.

i. extended codebook

0	Off line (thoughts divergent from activities)	8	Can't code
1	On line (thoughts related to current activities)	9	Missing

ii. combined categories

The recode statement used in the BP analyses was:

ONLINE(0=0)(1=1)(Else=9)(9=missing)

labels0 "Off Line"1 "On Line"9 Can't Code

iii "Basic Properties" reference data

Item	n	%
0 Off Line	1474	39.2%
1 On Line	2291	60.8%
	3765	—

• Thought Content

This item assesses the content of the thoughts. It is not intended to be a measurement of psychopathology, but reflects the (normal) content topics of the thoughts.

i. extended codebook

0 Nothing	22 Pain	41 Objects, other anima	54 Social interaction
9 Other self	23 Health, disease	42 Location, place	55 Meals
10 Mood	29 Other physical	43 Situation	56 Organizational act.
11 Sex	30 Partner	44 World view, life in	59 Other activities
12 Aggression	31 Children	45 Knowledge, language,	60 ESM
13 Death, suicide	32 Other relatives	49 Other contextual	98 Cant code
14 Self in relation to	33 Friends	50 Inactivity	99 Missing
19 Other psychic	34 Colleagues, acquaintance	51 Work, study	
20 Appearance, own body	39 Other persons	52 Maintenance act.	
21 Tired, hunger, thirst	40 Pets	53 Hobby, leisure	

ii. combined categories

The recode statement used in the BP analyses was:

THOUCONT(0,9=0)(10-19=1)(20-29=2)(30-32=3)(34-39=4)(40-49=5)(50-59=6)(60=7)

(Else=9)(99=missing)

labels0 Nothing1 Psychic2 Physical3 Family members

4 Other persons5 Context6 Activities7 ESM

9 Can't Code

iii "Basic Properties" reference data

Item	n	%	Item	n	%
00 Nothing	196	3.7%	39 Other persons	185	3.5%
09 Other self	225	4.3%	40 Pets	23	0.4%
10 Mood	77	1.5%	41 Objects, other anima	746	14.2%
11 Sex	21	0.4%	42 Location, place	196	3.7%
12 Aggression	14	0.3%	43 Situation	825	15.7%
13 Death, suicide	32	0.6%	44 World view, life in	154	2.9%
14 Self related to other	178	3.4%	45 Knowledge, language,	31	0.6%
19 Other psychic	179	3.4%	49 Other contextual	411	7.8%
20 Appearance, own body	88	1.7%	50 Inactivity	69	1.3%
21 Tired, hunger, thirst	99	1.9%	51 Work, study	146	2.8%
22 Pain	29	0.6%	52 Maintenance act.	268	5.1%
23 Health, disease	37	0.7%	53 Hobby, leisure	129	2.5%
29 Other physical	24	0.5%	54 Social interaction	51	1.0%
30 Partner	58	1.1%	55 Meals	101	1.9%
31 Children	72	1.4%	56 Organizational act.	3	0.1%
32 Other relatives	134	2.6%	59 Other activities	38	0.7%
33 Friends	78	1.5%	60 ESM	160	3.1%
34 Colleagues, acquaint.	162	3.1%		5239	—

• *Goofs*

The “goofs” are assessments of psychopathology by the research staff, as they are made based on the thought descriptions of the subjects. The first scores fall within the normal range (up to worrying), but the highest scores are clearly pathological (psychotic).

i. *extended codebook*

1	Clear thoughts	4	Worrying	8	Can't code
2	Daydreaming, fantasy	5	Obsession, preoccupation	9	Missing
3	Sighs, laments, complaints	6	Delusion, psychotic thought		

ii. *combined categories*

The recode statement used in the BP analyses was:

GOOFS (1=1) (2=2) (3,4,5=3) (6=4) (Else=9) (9=missing)
labels 1 Clear Thought 2 Daydreaming 3. Worrying 4 Psychotic 9 NA

iii “Basic Properties” reference data

Item		n	%	Item		n	%
1	Clear thoughts	4668	89.7%	4	Worrying	39	0.7%
2	Daydreaming	27	0.5%	5	Obsession	287	5.5%
3	Sighs	49	0.9%	6	Delusion	135	2.6%
						5205	—

• *Thought “type”*

The “Thought Type” is an assessment of the type of thought, done by the research staff. The code tends to reflect the way the subjects evaluates and assimilates his environment mentally. The instruction is to give only definite codes to avoid confusion in the data.

i. *extended codebook*

0	Neutral	3	Weight, Choices	8	Can't code
1	Evaluation	4	Planning	9	Missing
2	Introspection	5	Anticipation		

ii. *combined categories*

The recode statement used in the BP analyses was:

THOUTYPE (0=0) (1,2,3=1) (4,5=2) (Else=9) (9=missing)
labels 0 Neutral Thought 1 Evaluation 2. Planning 9 Can't Code

iii “Basic Properties” reference data

Item		n	%
0	Neutral	2727	52.5%
1	Evaluation	891	17.2%
2	Introspection	352	6.8%
3	Weight, Choices	649	12.5%
4	Planning	500	9.6%
5	Anticipation	74	1.4%
		5193	—

• *Thought “time”*

The “Thought Time” is an assessment of the timing of thought. This assessment is done by the research staff. The code reflects the time focus of thoughts. The instruction is to give only definite codes to avoid confusion in the data.

i. *extended codebook*

1	Retrospection (past)	3	Current Time (present)	9	Missing
2	Prospection (future)	8	Can't code		

ii. combined categories

no changes

iii "Basic Properties" reference data

Item		n	%
1	Retrospection	426	21.6%
2	Prospection	1158	58.8%
3	Current Time	386	19.6%
		1970	—

• Thought "valence"

The "Thought valence" is an assessment of the positive/negative dichotomy of thoughts. This assessment is done by the research staff. This is a coded variable which should match the self-assessed value attributed by the subject to the thought. A large number of thoughts were difficult to code. The instruction is to give only definite codes to avoid confusion in the data.

i. extended codebook

0	Negative	8	Can't code
1	Positive	9	Missing

ii. combined categories

no changes

iii "Basic Properties" reference data

Item		n	%
0	Negative	388	40.7%
1	Positive	566	59.3%
		954	—

Context codes

• Where

The codes on where reflect the physical place persons are in. As far as possible we have restricted ourselves to the place on itself, however, due to the fact that places are defined as "meaningful places" we were sometimes forced to compromise on this general rule. An example can illustrate this:

A women works in a shop, one floor down her home. When she works and she is downstairs the place will be coded as her working place, when she shops in another shop it will be coded as "public place indoor" and when she is upstairs waiting for clients coming in, her apartment will be her working place. Finally, when she is doing household shores, waiting for clients to come in, the place will be coded as "home".

Our primary option was to reconstruct the meaningfulness of places by combining codes. Being in a shop would then be coded "public places - indoors" irrespectively of the activity performed by the subject in that place, being "working" or "grocery chores" as a "household activity". Because those combined definitions of meaningfulness of the environment would create confusion in interpreting situations, we did not pursue this strategy further on.

i. extended codebook

10	At home	40	Health care setting	90	Other places
19	Home, outdoor	50	Public places outdoor	98	Can't code
20	Network	60	Public places indoor	99	Empty
30	Work	70	Transport		

ii. combined categories

The recode statement used in the BP analyses was:

WHERE (10=1) (19=1) (20=2) (30=3) (40,50,60=4) (70=5) (Else=9) (9=missing)
labels 1 Home 2 Network 3 Workplace 4 Public Place
5 Transport 9 Can't Code

iii "Basic Properties" reference data

Item	n	%
10 At home	3421	70.2%
20 Network	411	8.4%
30 Work	209	4.3%
40 Health Care Setting	192	3.9%

Item	n	%
50 Public Places Out	244	5.0%
60 Public Places In	214	4.4%
70 Transport	162	3.3%
90 Other	19	0.4%
	4872	—

• What(1/2)

The "what" code reflects the activity of the subject at the moment of the beep. The subjects are stimulated to describe those activities in their own words, using for instance, slang and personal vocabulary. The coding is done by the research assistant who is in charge of the briefing and debriefing and takes time during the debriefing session to check through the booklets in anticipation of possible coding problems. The coding list is based on the frequency—lists of the International Time Budget studies (Szalai et al., 1972). There are two "what"-codes to account for mixed activities such as watching television and smoking a cigarette. The codebook stipulates that the main activity (most active, with the highest level of involvement) should be coded in "what1" while "what2" is reserved for secondary activities.

i. extended codebook

00 Nothing	24 Shopping & services	46 Read book, periodical	68 Related transport
01 Sleep, nap, rest	25 Personal care service	48 Related transport	69 Other meals
02 Pasting	26 Personal medical care	49 Other leisure act.	71 Political/civic act
11 Work	27 Personal hygiene	51 Conversation	72 Organized religious
12 Volunteer work	28 Related transport	52 Physical interaction	78 Related transport
13 School, study	29 Other maintenance	53 Visits	79 Other organizational
14 Study, irregular	31 Personal mental activity	54 Café, pub visits	88 Other transport
15 Work & study breaks	39 Other personal activity	55 Parties	89 Other activities
18 Related transport	41 Active sports	58 Related transport	98 Can't code
21 Household chores	42 Taking a walk	59 Other social	99 Empty
22 Child care	43 Hobbies, art...	61 Regular meals	
19 Other work, study	44 Movie, theater...	62 Special meals	
23 Care to adults	45 TV, radio...	63 Snacks	

ii. combined categories

The recode statement used in the BP analyses was:

WHAT (00,01,02,31=0) (11,12,13,14,15,19=1) (21 thru 27,29=2) (41 thru 46,49=3) (51 thru 55, 59,62=4) (61,63,69=5) (71,72,79,89=6) (18,28,48,58,68,78,88=7) (Else=9) (9=missing)

labels 0 Not Much 1 Work/Study 2 Maintenance 3 Leisure 4 Social
5 Meals 6 Other 7 Transport 9 Missing

iii "Basic Properties" reference data

Item: WHAT1	n	%
0 Nothing	230	5.1%
2 Sleep, nap, rest	350	7.8%
11 Work	421	9.4%
12 Volunteer work	142	3.2%
13 School, study	71	1.6%
14 Study, irregular	40	0.9%
15 Work & study break	107	2.4%
18 Related transport	72	1.6%
19 Other work, study	21	0.5%
21 Household chores	511	11.4%
22 Child care	56	1.2%
24 Shopping & services	90	2.0%
25 Personal care service	3	0.1%
26 Personal medical care	58	1.3%
27 Personal hygiene	210	4.7%
28 Related transport	111	2.5%
29 Other maintenance	21	0.5%
31 Personal mental activity	343	7.6%
41 Active sports	4	0.1%
42 Taking a walk	66	1.5%

Item: WHAT1/b	n	%
43 Hobbies, art	115	2.6%
44 Movie, theater...	9	0.2%
46 Read book, newspaper	206	4.6%
48 Related transport	21	0.5%
49 Other leisure act	17	0.4%
51 Conversation	387	8.6%
52 Physical	8	0.2%
53 Visits	62	1.4%
54 Café, pub visits	40	0.9%
55 Parties	8	0.2%
58 Related transport	47	1.0%
59 Other social	13	0.3%
61 Regular meals	233	5.2%
62 Special meals	5	0.1%
63 Snacks	335	7.5%
71 Political/civic act	3	0.1%
72 Organized religious	2	0.0%
79 Other organizational	2	0.0%
88 Other transport	28	0.6%
89 Other activities	25	0.6%
	4493	—

Item: WHAT2	n	%	Item: WHAT2/cont'd	n	%
2 Sleep, nap, rest	23	2.7%	41 Active sports	1	0.1%
11 Work	8	0.9%	42 Taking a walk	4	0.5%
12 Volunteer work	5	0.6%	43 Hobbies, art	21	2.4%
14 Study, irregular	2	0.2%	45 TV, radio	63	7.3%
15 Work & study break	5	0.6%	46 Read book, newspaper	32	3.7%
18 Related transport	3	0.3%	48 Related transport	1	0.1%
19 Other work, study	3	0.3%	49 Other leisure act	5	0.6%
21 Household chores	37	4.3%	51 Conversation	191	22.2%
22 Childcare	2	0.2%	52 Physical	1	0.1%
24 Shopping & services	7	0.8%	53 Visits	153	17.7%
25 Personal care service	1	0.1%	54 Café, pub visits	20	2.3%
26 Pers. medical care	10	1.2%	55 Parties	1	0.1%
27 Personal hygiene	10	1.2%	58 Related transport	1	0.1%
28 Related transport	2	0.2%	61 Regular meals	34	3.9%
29 Other maintenance	1	0.1%	63 Snacks	169	19.6%
31 Personal mental activity	39	4.5%	89 Other activities	7	0.8%
				862	—

• Who(1/2)

The “who” code reflects the persons in the environment of the subject at the moment of the beep. The subjects are stimulated to describe those persons in their own words, using for instance, first names. The coding is done by the research assistant who is in charge of the briefing and debriefing and takes time during the debriefing session to check through the booklets in anticipation of possible coding problems. There are two “who”-codes to account for a mixed company. The codebook stipulates that the lower codes (more proximate relations) have a priority over other codes.

i. extended codebook

00 Alone	19 Other relatives	40 Colleagues
01 Alone with pets	20 Partner (non-household)	41 Neighbors
10 Partner (household)	21 Children: pre-school	42 Health care profess
11 Children: pre-school	22 Children: adolescent	49 Other acquaintances
12 Children: adolescent	23 Children: grown up	50 Strangers, others
13 Children: grown up	24 Parents	98 Can't code
14 Parents	29 Other relatives	99 Empty
15 Brothers, sisters	30 Friends	

ii. combined categories

The recode statement used in the BP analyses was:

WHO	(00,01=0) (10 thru 19=1) (20 thru 29=2) (30 thru 49=3) (50,59=4) (98=9) (99=missing) (else=9)			
labels	0 'Alone	1 Family-In	2 Family-Out	3 Friends/Acquaintances
	4 Strangers	9 Can't Code		

iii “Basic Properties” reference data

Item: WHO 1/a	n	%	Item: WHO 1/b	n	%
0 Alone	2141	39.5%	22 Children: adolescent	7	0.1%
1 Alone with pets	94	1.7%	23 Children: grown up	9	0.2%
10 Partner (household)	416	7.7%	24 Parents	174	3.2%
11 Children: pre-school	53	1.0%	29 Other relatives	96	1.8%
12 Children: adolescent	8	0.1%	30 Friends	272	5.0%
13 Children: grown up	21	0.4%	40 Colleagues	535	9.9%
14 Parents	208	3.8%	41 Neighbors	25	0.5%
15 Brothers, sisters	16	0.3%	43 Health care profess	186	3.4%
19 Other relatives	657	12.1%	49 Other acquaintances	208	3.8%
20 Partner (non-household)	39	0.7%	50 Strangers, others	248	4.6%
21 Children: pre-school	2	0.0%		5415	—

Item: WHO2/a			Item: WHO2/b		
	n	%		n	%
10 Partner (household)	8	1.1%	24 Parents	14	1.9%
11 Children: pre-school	79	10.5%	29 Other relatives	87	11.5%
12 Children: adolescent	14	1.9%	30 Friends	25	3.3%
13 Children: grown up	24	3.2%	40 Colleagues	8	1.1%
14 Parents	7	0.9%	41 Neighbors	7	0.9%
15 Brothers, sisters	13	1.7%	43 Health care profess	251	33.3%
19 Other relatives	38	5.0%	49 Other acquaintances	73	9.7%
20 Partner (non-household)	14	1.9%	50 Strangers, others	88	11.7%
23 Children: grown up	4	0.5%		754	—

Event codes

The between beep events are coded using the codebook that is described above for “thought content”, “activities” (What) and “persons involved” (Who). These content based eventcodes are not used in this study. We have, however, developed a coding scheme for events that happen over time. This coding scheme is been used for the study of hallucinations (Chapter 10).

Procedures for dynamic analysis of events

The definition of an “event” is defined by specific cut-off scores described at the end of this section. For the dynamic analysis of events, we used the following general procedure:

Procedure:

- Step 1:
- select those beeps that:

 - come from subjects with variability in the target events;
 - were answered within the “reliability time-window”; and
- Step 2:
- compute the eventcode (0=no; 1=yes) based on the cut-off scores (see end of section).
- Step 3:
- compute the transitions from time “t-1” to time “t” and from “t” to “t+1” and define the “Sequence Moment Codes” as:

 - 0= no event;
 - 1= last beep with no event;
 - 2= start event;
 - 3= prolonged event;
 - 4= last beep of an event;
 - 5= first after an event.

Figure A5.1 Overview of the time t-1/ time t transition codes.

When a transition can be coded in two ways the next priority list is used [0T] priority over [00]; [1T] priority over [11] and [10] has priority over [01]. Each blank beep restarts counting and generates a new period.

Program:

```
GET FILE.
sort cases by beepcode.
missing values item (99).

compute okla1=0.
compute okla2=0.
if(lag(beepno,1)=(beepno-1)) okla1=1.
if(lag(beepno,2)=(beepno-2)) okla2=1.
if(lag(beepno,1)=(beepno-2)) okla2=2.

if(item=1) Event=0.
if(item>1) Event=1.
if(item=9) Event=9.
missing values Event (99).

compute Event1=lag(Event,1).
if(okla1=0) Event1=9.
compute Event2=lag(Event,2).
if(okla2=0) Event2=9.
if(okla2=2) Event2=lag(Event,1).

compute Eventpat=(Event2*100)+(Event1*10)+Event.

if((Eventpat=999)or(Eventpat=991)or(Eventpat=990)
   or(Eventpat=199)or(Eventpat=191)or(Eventpat=190))
   Eventt=9.
if((Eventpat=909)or(Eventpat=900)or(Eventpat=9)or(Eventpat=0))
   Eventt=0.
if((Eventpat=901)or(Eventpat=101)or(Eventpat=1))Eventt=1.
if((Eventpat=919)or(Eventpat=911)or(Eventpat=910)
   or(Eventpat=119)or(Eventpat=10)or(Eventpat=11))Eventt=2.
if((Eventpat=119)or(Eventpat=111))Eventt=3.
if((Eventpat=110))Eventt=4.
if((Eventpat=109)or(Eventpat=100))Eventt=5.

save/outfile "Temp".
get file "Temp".
sort cases by beepcode(D).
if(lag(beepcode,1)=(Beepcode+1))Eventtn=lag(Eventt,1).
if(Event=9)Eventtn=9.
save/outfile "Temp2".

get file "Temp2".
sort cases by beepcode.
if (missing(Eventtn)and(Event=0)and(lag(Event,1)=0))Eventtn=0.
if (missing(Eventtn)and(Event=0)and(lag(Event,1)=1))Eventtn=5.
if (missing(Eventtn)and(Event=1)and(lag(Event,1)=0))Eventtn=2.
if (missing(Eventtn)and(Event=1)and(lag(Event,1)=1))Eventtn=3.
save/outfile "Event.SYS".
```

Example:

Raw data:	1135721111162411...
Event detection: (see event cut-off score)	0011110000011100...
Sequence moment code:	9123345000123450...

Cut-off scores of psychopathological symptoms are defined arbitrarily at $\pm 25\%$ of the beeps in all groups.

Psychopathology-items	1	2	3	4	5	6	7	Cut off
• suspicion	4821 71.6%	356 5.3%	358 5.3%	535 7.9%	202 3.0%	78 1.2%	387 5.7%	3+
• thoughts difficult to express	3276 56.4%	436 7.5%	294 5.1%	765 13.2%	368 6.3%	207 3.6%	466 8.0%	4+
• obsession	2938 50.7%	392 6.8%	373 6.4%	845 14.6%	418 7.2%	247 4.3%	583 10.1%	4+
• thoughts controlled	3481 60.2%	368 6.4%	303 5.2%	574 9.9%	336 5.8%	225 3.9%	493 8.5%	4+
• de realization	4787 73.7%	361 5.6%	283 4.4%	496 7.6%	254 3.9%	93 1.4%	218 3.4%	2+
• auditory hallucinations	4875 81.1%	235 3.9%	118 2.0%	135 2.2%	116 1.9%	67 1.1%	465 7.7%	2+
• visual hallucinations	2840 84.9%	126 3.8%	56 1.7%	65 1.9%	65 1.9%	27 0.8%	167 5.0%	2+
• fear to loose control	2892 72.1%	254 6.3%	178 2.6%	206 5.1%	129 3.2%	51 1.3%	303 4.4%	2+

Table A5.1 Overview of the response distributions in the pathology-module to define the cut-off scores for “events”.

Appendix 6: ESM file structure⁹⁰

The fundamental data of the ESM are a set of hundreds (or thousands) of self-reports obtained from participants at random times. For each self-report one will have a small number of variables on context, moods, thoughts—whatever questions have been asked on the ESM self-report form (the ESF). Along with these data one typically has information on the 10, 50, or 200 participants who were in the study. This information will include at least their ages and sexes, but probably also some questionnaire or interview data obtained before or after the period of paging; in some cases one may also have data from school or medical records, from teachers, nurses, parents or other sources. These two types of data, the beeper data and the data on one's subjects, correspond to the two types of computer files one will want to create to store and organize one's data. The first is commonly called a "Beeper File"; the second is commonly called a "Subject File".

A. BEEPER FILE

In the Beeper File the individual self-report or ESF is the unit of organization, what in SPSS-X or SAS is referred to as the "case". Typically all of the information on the self-report form comprises only 30 to 50 variables and fits onto one 80-character computerline. But since there are hundreds or thousands of self-reports, some dozens for each subject, this file typically contains many, many lines.

In this file the person is not the significant unit of organization, but one needs to include the subject ID number as a variable in order to have it available. Note that the number of cases per person will vary, since some people will have responded to the beeper more often than others. We also recommend that you create a variable that identifies the order in which each person filled out his or her beeper sheets — the "page number" — should you decide to do any sequential analysis. In the Maastricht database the page-number is substituted for a day-counter and a within day beep-counter because the night is considered as an important break in the continuity of the data sequence. You may also wish to repeat two or three of the most fundamental subject variables (e.g. sex, age, diagnostic group) on each case⁹¹.

Typically we create a "z-score" version of the Beeper File. In this file all the continuous variables have been standardized to remove differences between individuals in how they responded to each item. These z-scores are created by subtracting the subject's overall mean for the item and then dividing by the subject's standard deviation. In SAS you use PROC STANDARD; in SPSS-X you use CONDESCRIPTIVES (now called DESCRIPTIVES) preceded by SORT CASES BY ID and temporary SPLIT FILE BY ID commands⁹². The Z-score File provides a file in which variance due to subject main effects have been removed, thus it allows one to make comparisons within individuals on a scale where values are relative to each person's distribution for that particular item. Z-scores, however, are no panacea and may be misleading by introducing spurious variability in the data. E.g. when the subject's s.d. are small deviation from the mean will be blown out and may cause unwanted significance.

Rarely, if ever, would we recommend that final analysis of mood data be carried out on the raw Beeper File alone. The values on this file are affected both by person level and moment-level sources of variance, thus any simple analysis leads one into the Perspective Fallacy discussed in Example 1. A common example of this mistake that can be avoided using z-scored data, is the factor analysis of the Beeper File with the objective of simplifying a set of mood variables to a smaller number of factors. The problem

⁹⁰ This appendix was originally part of Larson & Delespaul (1992) and is adapted for this study.

⁹¹ This is important for the pragmatics of analysis.

⁹² Spot-checking these computations is important. The first author has found instances when SPSS-X mishandles cases when a person has no variance.

when this is done on raw data is that one does not know whether the factors obtained represent qualities of people or qualities of moments or situations.

Record j field 1 field 2 field i field n

with: Subject Identification Part

field 1	Subject ID	
field 2	Subject File data	optional
...		

Beep Identification Part

field ...	Beep Sequence Number (day/beep)	
field	Date	
field	Beep time	
field	Answering time	optional

Beep Experiential data

field	e.g. happiness rating	customized to study...
...		

Beep Situational data

field	e.g. activity, companionship,... customized to study...	
...		

Figure A6.1 Typical field structure of a Beep File record.

B. SUBJECT FILE

In the Subject File the individual subject is the unit of organization or the "case". Since you probably have relatively few subjects, but much information on each one, this file — in contrast to the Beeper File — is likely to be short and fat. In fact we are going to suggest making it a lot fatter in a moment and you may even decide that you need to have several Subject Files, indexed by subject ID. The initial data on the Subject File will include all of your demographic data on the participants, questionnaire and interview data, and information from archives or from other sources. In addition to these you probably want to add various ESM summary variables, computed from the Beeper File, that describe the typical experience of each subject.

ESM summary variables can be computed on an ad hoc basis, as needed, but you will find it useful (and cheaper) to generate a standard skeleton of summary variables that you anticipate using again and again. These might include each persons' mean score for each ESM scaled item, the standard deviation for these items, the percentage of time the person is alone, with family members, and with friends, and the percentage of time he or she is engaged in various activities, thinks various categories of thought, or experiences various extreme emotional states. These might also include the person's mean states in certain contexts (e.g. alone, with family, with friends); typically the first author stores the raw values of these contextual means, but frequently converts them, when necessary, to mean z-scores by using the overall means and overall standard deviations available on this file.

Basically, you need to decide what the key variables are likely to be for your project—and be prepared to create more as your understanding of the data set matures. We recommend extreme care and ample discussion with colleagues before selecting aggregated data for analysis. Seemingly harmless conversions may have unexpected consequences for the analysis. In this case also we advise not to trust one summary score but design different ones that should give the same results in the analysis. When divergence is found it should be interpreted in relation to the summary procedure used. E.g. summary z-scores of experience in situations will reach 0.00 when the frequency of the situation is high.

These summary scores are computed from your Beeper File. In SPSS-X one can use the procedure AGGREGATE and in SAS the procedure PROC SUMMARY. Subject ID is used as the BREAK or BY variable. The output file created by these procedures is then merged with your Subject File using ID as the matching variable and the enlarged Subject File is saved for future use. As with all file manipulations, investigators are warned that they must check and double check all these computations by listing out data

before and after the manipulations. Numerous horror stories could be told at this point about people who created files that seemed right, but months later were discovered to be wrong, requiring computer runs to be redone and findings to be rethought.

Figure A6.2 Typical field structure of a Subject File record.

C. ADDITIONAL VARIATIONS

In addition to a Beeper File and a Subject File a particular investigator may require a file organized by some other unit. Mayers (1978) was interested in investigating high school students' experience in different school classes, therefore organized a file in which the classes for each person were the fundamental case — with the ESM data aggregated using ID and class as the break variables. Increasingly investigators are having ESM participants fill out diary reports at the end of each day. This leads to the creation of files in which the day is the case and the unit of analysis. It is also conceivable that investigators may find it useful to create files in which a specific episode (a drinking binge or a stressful event) defines a sequence of ESM sheets and becomes the unit of organization, though often these analyses can be carried out on the Beeper-File.

The organization of data is the key to analysis of ESM data since the type of analysis follows directly from the file organization. Given a choice between hiring an Research Assistant with talents in data management or statistical analysis we highly recommend the former.

**Appendix 7: ESM output for clinical practice:
Personalized Rehabilitation**

The present appendix contains the unabbreviated computer-output for the third patient described in chapter 13. This overview is generated by the software that was developed to facilitate clinical decision making based on ESM data. It is rather complex but when we assess it globally it allows us to scan the ESM data quickly. It starts with subject identification material and general ESM characteristics. The next part assesses the time budgets of which the first three (What? - Who? - Where?) and the fifth (Thought Content?) are the most important. Part 2 is the assessment of mental state, at the item level. Data are broken down on “days of the week”, “week-end/week” and “parts of the day”, which allows us to see very quickly when significant moments that diverge from the normal patterns occurred. The next three tables contain the mental state items broken down by context for Who? - What? and Where? Screening significance, although problematic from a statistical point of view considering the number of tests that are computed, yield interesting information for clinical decision making. The last four tables contain experiential parameters related to Quality of Life and are assessments of the general success of adaptation.

A. REPORT IDENTIFICATION

Subject: 739; Period 1
ESM Start: Friday, August 6 1993
Selection restricted to valid beeps (-5/+15 minutes interval)
Beeps = 52/ # Days = 12/ Response% =43%: Valid Subject ⁹³
Report: dd. 13-05-1994

Man ; 29yrs; never married; unfitted for work; living in a special psychiatric environment;
DSM-III-R: 299.80 Asperger Syndrome

B. TIME BUDGETS

Context

What

Figure A7.1a Time budget of context: activities.

⁹³ The subject was sampled in two periods. The subject did not reach the validity criterion during the first period (33% of the beeps should be valid, while only 23% were). The second period was clearly better: 63% of valid beeps.

Who

Figure A7.1b Time budget of context: persons.

Where

Figure A7.1c Time budget of context: places.

Thought Assessment

Online

#	Item Description	n	%	F	S	S	M	T	W	T	W	T	F	S	S
0	Off line	32	62.74	3	1	0	3	0	2	1	6	4	3	3	6
1	On line	9	17.64	1	0	0	0	0	0	0	2	1	0	2	3
8	Can't code	10	19.60	0	1	1	1	0	0	2	0	3	0	1	1

Categories: 3 %/Category x (s.d.): 33.33 (25.49) % Change over time: 49.01%

Figure A7.2a Time budget of thoughts: Online.

Thought Relation

#	Item Description	n	%	F	S	S	M	T	W	T	W	T	F	S	S
0	Neutral	13	25.49	1	1	1	2	0	0	1	2	3	0	0	2
1	Evaluation	3	5.88	0	0	0	0	0	0	0	0	0	0	2	1
2	Introspection	21	41.17	1	1	0	0	0	2	2	5	3	2	2	3
3	Weigh, Choices	12	23.52	1	0	0	1	0	0	0	1	2	1	2	4
4	Planning	2	3.92	1	0	0	1	0	0	0	0	0	0	0	0

Categories: 5 %/Category x (s.d.): 20.00 (15.40) % Change over time: 72.54%

Figure A7.2b Time budget of thoughts: thought relation.

Thought Content

Figure A7.2c Time budget of thoughts: thought content.

Goofs

Figure A7.2d Time budget of thoughts: thought pathology.

Thought Time

Figure A7.2e Time budget of thoughts: thought time perspective.

Thought Pos./Neg. Value

Figure A7.2f Time budget of thoughts: thought value.

C. MENTAL STATE

Overview: Days of the Week

Item	n	x (s.d.)	range	Sunday (n=11)	Monday (n=8)	Tuesday (n=4)	Wednesday (n=10)	Thursday (n=8)	Friday (n=7)	Saturday (n=8)	
Pleasant thought	48	2.04 (1.44)	1-5	2.2 (1.6)	2.0 (2.0)	3.0 (1.0)	1.7 (1.3)	2.5 (1.6)	1.9 (1.5)	1.5 (1.2)	F(6,41)=0.60 n.s.
Clear thought	48	6.65 (0.86)	4-7	6.3 (1.1)	7.0 (0.0)	7.0 (0.0)	6.4 (1.3)	6.8 (0.7)	6.7 (0.8)	7.0 (0.0)	F(6,41)=0.77 n.s.
Normal thought	48	5.33 (2.44)	1-7	6.1 (2.0)	4.0 (3.5)	6.3 (1.2)	5.1 (2.2)	5.5 (2.8)	4.1 (2.9)	6.0 (2.4)	F(6,41)=0.80 n.s.
Difficult to concentrate	48	1.88 (1.18)	1-6	1.6 (0.5)	1.0 (0.0)	2.3 (2.5)	2.3 (1.7)	2.3 (0.7)	1.7 (0.8)	1.7 (0.8)	F(6,41)=0.93 n.s.
I feel ...											
Cheerful	51	1.49 (1.14)	1-6	1.4 (0.9)	1.0 (0.0)	3.3 (2.2)	1.0 (0.0)	1.1 (0.4)	2.1 (1.5)	1.4 (1.1)	F(6,44)=3.28 p<0.010 **
Unsure	51	1.04 (0.28)	1-3	1.0 (0.0)	1.5 (1.0)	1.0 (0.0)	1.0 (0.0)	1.0 (0.0)	1.0 (0.0)	1.0 (0.0)	F(6,44)=2.25 p<0.055
Lonely	51	1.00 (0.00)	1-1								
Relaxed	51	1.10 (0.50)	1-4	1.0 (0.0)	1.0 (0.0)	1.5 (1.0)	1.0 (0.0)	1.0 (0.0)	1.4 (1.1)	1.0 (0.0)	F(6,44)=1.23 n.s.
Anxious	51	1.00 (0.00)	1-1								
Satisfied	51	1.00 (0.00)	1-1								
Angry/Irritated	51	4.57 (1.65)	1-7	4.2 (1.3)	5.0 (1.4)	2.5 (1.3)	5.8 (1.4)	4.8 (1.0)	4.0 (2.3)	4.7 (1.6)	F(6,44)=2.76 p<0.023 *
Sad	51	4.39 (1.73)	1-7	3.8 (1.0)	4.5 (1.3)	3.3 (1.5)	6.2 (1.1)	4.6 (1.6)	3.4 (2.2)	4.0 (1.8)	F(6,44)=3.62 p<0.006 **
Guilty.	51	1.00 (0.00)	1-1								
Not motivated	51	2.29 (1.60)	1-6	3.1 (1.8)	1.0 (0.0)	2.0 (2.0)	1.6 (1.3)	3.1 (1.4)	1.4 (1.1)	2.9 (1.6)	F(6,44)=2.41 p<0.042 *
Active	51	1.06 (0.24)	1-2	1.2 (0.4)	1.0 (0.0)	1.3 (0.5)	1.0 (0.0)	1.0 (0.0)	1.0 (0.0)	1.0 (0.0)	F(6,44)=1.34 n.s.
Difficult	51	1.25 (0.96)	1-7	1.1 (0.3)	1.0 (0.0)	1.0 (0.0)	1.9 (2.0)	1.3 (0.5)	1.0 (0.0)	1.1 (0.4)	F(6,44)=1.01 n.s.
Challenging	49	1.00 (0.00)	1-1								
Skilled	49	5.71 (1.46)	2-7	5.5 (1.5)	6.3 (1.5)	4.8 (1.5)	5.9 (1.3)	5.5 (1.3)	6.0 (1.5)	6.0 (1.9)	F(6,42)=0.51 n.s.
Hunger	51	1.25 (0.98)	1-7	1.1 (0.3)	1.0 (0.0)	1.0 (0.0)	1.3 (0.9)	1.1 (0.4)	2.1 (2.3)	1.0 (0.0)	F(6,44)=1.25 n.s.
Tired	51	2.24 (1.19)	1-5	1.5 (0.5)	2.8 (1.5)	1.5 (0.6)	3.9 (1.1)	2.0 (0.5)	1.9 (0.9)	1.9 (0.7)	F(6,44)=9.45 p<0.000 ***
Not feeling well	51	2.86 (1.60)	1-7	1.6 (0.5)	3.0 (1.2)	2.0 (1.4)	4.8 (1.4)	2.6 (1.1)	3.3 (2.0)	2.3 (1.0)	F(6,44)=6.71 p<0.000 ***
Pathology	n	x (s.d.)	range	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	
I feel suspicious	52	1.00 (0.00)	1-1								
Thoughts ...											
...difficult to express	52	3.98 (2.68)	1-7	2.5 (1.9)	7.0 (0.0)	1.0 (0.0)	4.6 (2.8)	3.4 (2.4)	4.9 (2.9)	5.1 (2.7)	F(6,45)=3.69 p<0.005 **
...can't get rid	51	2.39 (2.12)	1-7	1.0 (0.0)	3.0 (1.8)	1.0 (0.0)	2.9 (2.2)	2.1 (2.0)	3.5 (2.9)	3.5 (2.6)	F(6,45)=2.15 p<0.065
...influenced by others	52	1.19 (0.84)	1-6	1.0 (0.0)	1.0 (0.0)	1.0 (0.0)	1.8 (1.8)	1.0 (0.0)	1.3 (0.8)	1.0 (0.0)	F(6,45)=1.22 n.s.
Feeling unreal	52	1.04 (0.19)	1-2	1.0 (0.0)	1.0 (0.0)	1.0 (0.0)	1.0 (0.0)	1.0 (0.0)	1.3 (0.5)	1.0 (0.0)	F(6,45)=2.60 p<0.030 *
Hearing voices	52	1.00 (0.00)	1-1								
Seeing things	52	1.00 (0.00)	1-1								
Fear losing control	52	1.00 (0.00)	1-1								
Substances	n	x (s.d.)	range	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	
Nothing	52	0.48 (0.50)	0-1	0.7 (0.5)	0.3 (0.5)	0.3 (0.5)	0.5 (0.5)	0.5 (0.5)	0.1 (0.4)	0.6 (0.5)	F(6,45)=1.42 n.s.
Alcohol	52	0.10 (0.30)	0-1	0.0 (0.0)	0.3 (0.5)	0.0 (0.0)	0.1 (0.3)	0.1 (0.4)	0.3 (0.5)	0.0 (0.0)	F(6,45)=1.07 n.s.
Medication	52	0.50 (0.50)	0-1	0.4 (0.5)	0.8 (0.5)	0.8 (0.5)	0.4 (0.5)	0.4 (0.5)	0.9 (0.4)	0.4 (0.5)	F(6,45)=1.32 n.s.
Coffee	52	0.08 (0.27)	0-1	0.1 (0.3)	0.3 (0.5)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.1 (0.4)	0.1 (0.4)	F(6,45)=0.67 n.s.
Tobacco	52	0.02 (0.14)	0-1	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.1 (0.4)	0.0 (0.0)	F(6,45)=1.08 n.s.
Other	52	0.00 (0.00)	0-0								
Between beep events	n	x (s.d.)	range	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	
Disurbing event	29	1.28 (1.16)	1-7	1.0 (0.0)	1.0 (0.0)	1.0 (0.0)	1.0 (0.0)	1.4 (0.9)	2.2 (2.7)	1.0 (0.0)	F(6,22)=0.66 n.s.
Pleasant event	29	1.76 (1.46)	1-7	1.7 (1.2)	1.0 (0.0)	1.7 (1.2)	1.0 (0.0)	1.8 (1.8)	2.6 (2.6)	2.2 (1.1)	F(6,22)=0.67 n.s.
Social interaction	n	x (s.d.)	range	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	
Pleasant company	14	3.43 (1.55)	1-5	4.4 (0.5)	-- (--)	2.5 (2.1)	1.0 (0.0)	4.7 (0.6)	-- (--)	2.5 (0.7)	F(4, 9)=8.05 p<0.005 **
Prefer being alone	51	4.75 (1.53)	1-7	5.0 (1.8)	4.0 (0.0)	5.3 (1.5)	5.2 (1.5)	3.5 (1.9)	4.6 (0.5)	5.4 (1.3)	F(6,44)=1.66 n.s.
Active interaction	14	1.36 (1.08)	1-5	1.2 (0.4)	-- (--)	3.0 (2.8)	1.0 (0.0)	1.0 (0.0)	-- (--)	1.0 (0.0)	F(4, 9)=1.64 n.s.
Idiographic complaint	n	x (s.d.)	range	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	
Tension	52	4.92 (1.43)	1-7	4.2 (0.9)	5.3 (1.5)	4.5 (0.6)	6.1 (1.0)	5.1 (1.2)	4.3 (2.3)	4.9 (1.4)	F(6,45)=2.29 p<0.051
Negativism/Sadness	52	4.50 (1.55)	1-7	4.0 (0.8)	4.3 (1.3)	3.0 (1.4)	6.3 (0.9)	4.6 (1.3)	3.7 (2.0)	4.4 (1.4)	F(6,45)=4.96 p<0.001 ***

Table A7.1a Experience in time: day of the weeks.

Overview: Week / Week-End

Items	n	x (s.d.)	range	Week (n=34)	Weekend (n=18)	
Pleasant thought	48	2.04 (1.44)	1-5	2.0 (1.4)	2.1 (1.5)	F(1,46)=0.00 n.s.
Clear thought	48	6.65 (0.86)	4-7	6.7 (0.8)	6.5 (0.9)	F(1,46)=1.09 n.s.
Normal thought	48	5.33 (2.44)	1-7	5.4 (2.4)	5.3 (2.5)	F(1,46)=0.01 n.s.
Difficult to concentrate	48	1.88 (1.18)	1-6	2.0 (1.4)	1.6 (0.6)	F(1,46)=1.08 n.s.
I feel...						
Cheerful	51	1.49 (1.14)	1-6	1.4 (1.1)	1.7 (1.2)	F(1,49)=0.66 n.s.
Unsure	51	1.04 (0.28)	1-3	1.1 (0.3)	1.0 (0.0)	F(1,49)=0.54 n.s.
Lonely	51	1.00 (0.00)	1-1			
Relaxed	51	1.10 (0.50)	1-4	1.1 (0.3)	1.2 (0.7)	F(1,49)=0.52 n.s.
Anxious	51	1.00 (0.00)	1-1			
Satisfied	51	1.00 (0.00)	1-1			
Angry/Irritated	51	4.57 (1.65)	1-7	4.8 (1.6)	4.1 (1.7)	F(1,49)=2.18 n.s.
Sad	51	4.39 (1.73)	1-7	4.8 (1.7)	3.7 (1.5)	F(1,49)=5.29 p<0.024 *
Guilty	51	1.00 (0.00)	1-1			
Not motivated	51	2.29 (1.60)	1-6	2.2 (1.5)	2.4 (1.8)	F(1,49)=0.24 n.s.
Active	51	1.06 (0.24)	1-2	1.0 (0.2)	1.1 (0.3)	F(1,49)=1.36 n.s.
Difficult	51	1.25 (0.96)	1-7	1.4 (1.2)	1.1 (0.2)	F(1,49)=1.22 n.s.
Challenging	49	1.00 (0.00)	1-1			
Skilled	49	5.71 (1.46)	2-7	5.7 (1.5)	5.7 (1.5)	F(1,47)=0.00 n.s.
Hunger	51	1.25 (0.98)	1-7	1.1 (0.5)	1.5 (1.5)	F(1,49)=1.78 n.s.
Tired	51	2.24 (1.19)	1-5	2.6 (1.3)	1.6 (0.7)	F(1,49)=8.80 p<0.005 **
Not feeling well	51	2.86 (1.60)	1-7	3.2 (1.6)	2.3 (1.5)	F(1,49)=3.94 p<0.050
Pathology	n	x (s.d.)	range	Week	Weekend	
I feel suspicious	52	1.00 (0.00)	1-1			
Thoughts...						
...difficult to express	52	3.98 (2.68)	1-7	4.3 (2.7)	3.4 (2.5)	F(1,50)=1.36 n.s.
...can't get rid	51	2.39 (2.12)	1-7	2.6 (2.1)	1.9 (2.1)	F(1,49)=1.49 n.s.
...influenced by others	52	1.19 (0.84)	1-6	1.2 (1.0)	1.1 (0.5)	F(1,50)=0.25 n.s.
Feeling unreal	52	1.04 (0.19)	1-2	1.0 (0.0)	1.1 (0.3)	F(1,50)=4.09 p<0.046 *
Hearing voices	52	1.00 (0.00)	1-1			
Seeing things	52	1.00 (0.00)	1-1			
Fear losing control	52	1.00 (0.00)	1-1			
Use	n	x (s.d.)	range	Week	Weekend	
Nothing	52	0.48 (0.50)	0-1	0.5 (0.5)	0.5 (0.5)	F(1,50)=0.04 n.s.
Alcohol	52	0.10 (0.30)	0-1	0.1 (0.3)	0.1 (0.3)	F(1,50)=0.07 n.s.
Medication	52	0.50 (0.50)	0-1	0.5 (0.5)	0.6 (0.5)	F(1,50)=0.33 n.s.
Coffee	52	0.08 (0.27)	0-1	0.1 (0.2)	0.1 (0.3)	F(1,50)=0.44 n.s.
Tobacco	52	0.02 (0.14)	0-1	0.0 (0.0)	0.1 (0.2)	F(1,50)=1.92 n.s.
Other	52	0.00 (0.00)	0-0			
Between beep events	n	x (s.d.)	range	Week	Weekend	
Disturbing event	29	1.28 (1.16)	1-7	1.1 (0.4)	1.8 (2.1)	F(1,27)=1.90 n.s.
Pleasant event	29	1.76 (1.46)	1-7	1.6 (1.1)	2.3 (2.1)	F(1,27)=1.27 n.s.
Social interaction	n	x (s.d.)	range	Week	Weekend	
Pleasant company	14	3.43 (1.55)	1-5	2.9 (1.7)	4.4 (0.5)	F(1,12)=3.66 p<0.077
Prefer being alone	51	4.75 (1.53)	1-7	4.7 (1.6)	4.8 (1.4)	F(1,49)=0.07 n.s.
Active interaction	14	1.36 (1.08)	1-5	1.4 (1.3)	1.2 (0.4)	F(1,12)=0.15 n.s.
Idiog. complaint	n	x (s.d.)	range	Week	Weekend	
Tension	52	4.92 (1.43)	1-7	5.3 (1.2)	4.2 (1.5)	F(1,50)=7.50 p<0.008 **
Negativism/Sadness	52	4.50 (1.55)	1-7	4.8 (1.6)	3.9 (1.3)	F(1,50)=4.56 p<0.035 *

Table A7.1b Experience in time: week/weekend days.

Overview: Parts of the Day

Items	n	x (s.d.)	range	Morning (n=15)	Aftern (n=21)	Evening (n=16)	
Pleasant thought	48	2.04 (1.44)	1-5	1.8 (1.4)	1.9 (1.3)	2.5 (1.6)	F(2,45)=0.99 n.s.
Clear thought	48	6.65 (0.86)	4-7	6.8 (0.6)	6.6 (1.0)	6.5 (0.9)	F(2,45)=0.41 n.s.
Normal thought	48	5.33 (2.44)	1-7	6.1 (1.9)	5.5 (2.4)	4.5 (2.9)	F(2,45)=1.43 n.s.
Difficult to concentrate	48	1.88 (1.18)	1-6	1.7 (1.4)	1.6 (0.7)	2.5 (1.3)	F(2,45)=3.07 p<0.055
I feel ...							
Cheerful	51	1.49 (1.14)	1-6	1.6 (1.1)	1.7 (1.4)	1.2 (0.8)	F(2,48)=0.85 n.s.
Unsure	51	1.04 (0.28)	1-3	1.0 (0.0)	1.1 (0.4)	1.0 (0.0)	F(2,48)=0.71 n.s.
Lonely	51	1.00 (0.00)	1-1				
Relaxed	51	1.10 (0.50)	1-4	1.0 (0.0)	1.2 (0.8)	1.0 (0.0)	F(2,48)=1.42 n.s.
Anxious	51	1.00 (0.00)	1-1				
Satisfied	51	1.00 (0.00)	1-1				
Angry/Irritated	51	4.57 (1.65)	1-7	4.9 (1.5)	4.4 (1.6)	4.6 (1.9)	F(2,48)=0.34 n.s.
Sad	51	4.39 (1.73)	1-7	4.4 (1.9)	4.4 (1.6)	4.4 (1.9)	F(2,48)=0.01 n.s.
Guilt.	51	1.00 (0.00)	1-1				
Not motivated	51	2.29 (1.60)	1-6	2.4 (1.8)	2.4 (1.6)	2.1 (1.5)	F(2,48)=0.13 n.s.
Active	51	1.06 (0.24)	1-2	1.1 (0.4)	1.0 (0.2)	1.0 (0.0)	F(2,48)=1.41 n.s.
Difficult	51	1.25 (0.96)	1-7	1.1 (0.3)	1.4 (1.3)	1.2 (0.8)	F(2,48)=0.63 n.s.
Challenging	49	1.00 (0.00)	1-1				
Skilled	49	5.71 (1.46)	2-7	5.1 (1.4)	5.7 (1.3)	6.3 (1.6)	F(2,46)=2.27 n.s.
Hunger	51	1.25 (0.98)	1-7	1.2 (0.8)	1.4 (1.3)	1.1 (0.5)	F(2,48)=0.32 n.s.
Tired	51	2.24 (1.19)	1-5	2.0 (1.0)	2.3 (1.3)	2.4 (1.1)	F(2,48)=0.39 n.s.
Not feeling well	51	2.86 (1.60)	1-7	2.7 (1.4)	2.9 (1.6)	3.0 (1.9)	F(2,48)=0.12 n.s.
Pathology	n	x (s.d.)	range	Morning	Aftern	Evening	
I feel suspicious	52	1.00 (0.00)	1-1				
Thoughts ...							
...difficult to express	52	3.98 (2.68)	1-7	3.2 (2.6)	4.8 (2.7)	3.6 (2.6)	F(2,49)=1.85 n.s.
...can't get rid	51	2.39 (2.12)	1-7	2.5 (2.2)	2.8 (2.3)	1.8 (1.8)	F(2,48)=0.91 n.s.
...influenced by others	52	1.19 (0.84)	1-6	1.0 (0.0)	1.1 (0.7)	1.4 (1.3)	F(2,49)=1.11 n.s.
Feeling unreal	52	1.04 (0.19)	1-2	1.0 (0.0)	1.1 (0.3)	1.0 (0.0)	F(2,49)=1.54 n.s.
Hearing voices	52	1.00 (0.00)	1-1				
Seeing things	52	1.00 (0.00)	1-1				
Fear losing control	52	1.00 (0.00)	1-1				
Use	n	x (s.d.)	range	Morning	Aftern	Evening	
Nothing	52	0.48 (0.50)	0-1	0.4 (0.5)	0.5 (0.5)	0.5 (0.5)	F(2,49)=0.27 n.s.
Alcohol	52	0.10 (0.30)	0-1	0.1 (0.3)	0.0 (0.2)	0.2 (0.4)	F(2,49)=1.11 n.s.
Medication	52	0.50 (0.50)	0-1	0.6 (0.5)	0.5 (0.5)	0.4 (0.5)	F(2,49)=0.43 n.s.
Coffee	52	0.08 (0.27)	0-1	0.2 (0.4)	0.0 (0.2)	0.0 (0.0)	F(2,49)=2.48 p<0.092
Tobacco	52	0.02 (0.14)	0-1	0.0 (0.0)	0.0 (0.2)	0.0 (0.0)	F(2,49)=0.73 n.s.
Other	52	0.00 (0.00)	0-0				
Between beep events	n	x (s.d.)	range	Morning	Aftern	Evening	
Disturbing event	29	1.28 (1.16)	1-7	1.0 (0.0)	1.5 (1.6)	1.0 (0.0)	F(2,26)=0.75 n.s.
Pleasant event	29	1.76 (1.46)	1-7	1.6 (1.0)	1.3 (0.7)	4.0 (2.6)	F(2,26)=8.82 p<0.001 **
Social interaction	n	x (s.d.)	range	Morning	Aftern	Evening	
Pleasant company	14	3.43 (1.55)	1-5	2.0 (1.7)	3.6 (1.3)	5.0 (0.0)	F(2,11)=3.05 p<0.087
Prefer being alone	51	4.75 (1.53)	1-7	5.4 (1.5)	4.4 (1.6)	4.6 (1.5)	F(2,48)=1.65 n.s.
Active interaction	14	1.36 (1.08)	1-5	2.3 (2.3)	1.1 (0.3)	1.0 (0.0)	F(2,11)=1.74 n.s.
Idiographic complaint	n	x (s.d.)	range	Morning	Aftern	Evening	
Tension	52	4.92 (1.43)	1-7	5.1 (1.1)	4.7 (1.3)	5.1 (1.8)	F(2,49)=0.57 n.s.
Negativism/Sadness	52	4.50 (1.55)	1-7	4.5 (1.6)	4.4 (1.6)	4.6 (1.6)	F(2,49)=0.11 n.s.

Table A7.1c Experience in time: week/weekend days.

Overview: Social Context⁹⁴

Items	n	x (s.d.)	range	Alone (n=36)	Family (n=13)	Friends (n=2)	Other (n=1)	
Pleasant thought	48	2.04 (1.44)	1-5	1.8 (1.3)	2.7 (1.5)	1.0 (0.0)	4.0 (-)	F(3,44)=2.38 p<0.081
Clear thought	48	6.65 (0.86)	4-7	6.8 (0.8)	6.3 (1.0)	7.0 (0.0)	7.0 (-)	F(3,44)=0.99 n.s.
Normal thought	48	5.33 (2.44)	1-7	4.9 (2.7)	6.3 (1.8)	5.5 (2.1)	5.0 (-)	F(3,44)=0.98 n.s.
Difficult to concentrate	48	1.88 (1.18)	1-6	1.9 (1.1)	1.6 (0.7)	1.0 (0.0)	6.0 (-)	F(3,44)=6.23 p<0.002 **
I feel...								
Cheerful	51	1.49 (1.14)	1-6	1.5 (1.1)	1.5 (1.4)	1.0 (0.0)	2.0 (-)	F(3,47)=0.19 n.s.
Unsure	51	1.04 (0.28)	1-3	1.1 (0.3)	1.0 (0.0)	1.0 (0.0)	1.0 (-)	F(3,47)=0.14 n.s.
Lonely	51	1.00 (0.00)	1-1					
Relaxed	51	1.10 (0.50)	1-4	1.1 (0.5)	1.2 (0.6)	1.0 (0.0)	1.0 (-)	F(3,47)=0.09 n.s.
Anxious	51	1.00 (0.00)	1-1					
Satisfied	51	1.00 (0.00)	1-1					
Angry/Irritated	51	4.57 (1.65)	1-7	4.8 (1.7)	3.9 (1.2)	6.5 (0.7)	2.0 (-)	F(3,47)=2.83 p<0.047 *
Sad	51	4.39 (1.73)	1-7	4.5 (1.9)	4.0 (1.4)	5.0 (2.8)	4.0 (-)	F(3,47)=0.36 n.s.
Guilty	51	1.00 (0.00)	1-1					
Not motivated	51	2.29 (1.60)	1-6	2.1 (1.6)	2.7 (1.5)	1.0 (0.0)	5.0 (-)	F(3,47)=1.84 n.s.
Active	51	1.06 (0.24)	1-2	1.0 (0.2)	1.1 (0.3)	1.0 (0.0)	2.0 (-)	F(3,47)=7.68 p<0.000 ***
Difficult	51	1.25 (0.96)	1-7	1.3 (1.1)	1.2 (0.4)	1.0 (0.0)	1.0 (-)	F(3,47)=0.16 n.s.
Challenging	49	1.00 (0.00)	1-1					
Skilled	49	5.71 (1.46)	2-7	5.8 (1.5)	5.8 (1.4)	4.0 (0.0)	4.0 (-)	F(3,45)=1.53 n.s.
Hunger	51	1.25 (0.98)	1-7	1.2 (1.1)	1.2 (0.4)	2.5 (2.1)	1.0 (-)	F(3,47)=1.17 n.s.
Tired	51	2.24 (1.19)	1-5	2.4 (1.3)	1.8 (0.8)	2.5 (2.1)	2.0 (-)	F(3,47)=0.64 n.s.
Not feeling well	51	2.86 (1.60)	1-7	3.1 (1.6)	2.1 (1.0)	3.5 (3.5)	4.0 (-)	F(3,47)=1.60 n.s.
Pathology	n	x (s.d.)	range	Alone	Family	Friends	Other	
I feel suspicious	52	1.00 (0.00)	1-1					
Thoughts ...	52	3.98 (2.68)	1-7	4.6 (2.6)	3.0 (2.4)	1.0 (0.0)	1.0 (-)	F(3,48)=2.67 p<0.057
...difficult to express	51	2.39 (2.12)	1-7	2.6 (2.2)	1.8 (1.9)	2.5 (2.1)	1.0 (-)	F(3,47)=0.56 n.s.
...can't get rid	52	1.19 (0.84)	1-6	1.3 (1.0)	1.0 (0.0)	1.0 (0.0)	1.0 (-)	F(3,48)=0.64 n.s.
...influenced by others	52	1.04 (0.19)	1-2	1.1 (0.2)	1.0 (0.0)	1.0 (0.0)	1.0 (-)	F(3,48)=0.29 n.s.
Feeling unreal	52	1.00 (0.00)	1-1					
Hearing voices	52	1.00 (0.00)	1-1					
Seeing things	52	1.00 (0.00)	1-1					
Fear losing control								
Use	n	x (s.d.)	range	Alone	Family	Friends	Other	
Nothing	52	0.48 (0.50)	0-1	0.4 (0.5)	0.6 (0.5)	1.0 (0.0)	0.0 (-)	F(3,48)=1.56 n.s.
Alcohol	52	0.10 (0.30)	0-1	0.1 (0.4)	0.0 (0.0)	0.0 (0.0)	0.0 (-)	F(3,48)=0.79 n.s.
Medication	52	0.50 (0.50)	0-1	0.6 (0.5)	0.4 (0.5)	0.0 (0.0)	1.0 (-)	F(3,48)=1.38 n.s.
Coffee	52	0.08 (0.27)	0-1	0.1 (0.3)	0.0 (0.0)	0.0 (0.0)	0.0 (-)	F(3,48)=0.62 n.s.
Tobacco	52	0.02 (0.14)	0-1	0.0 (0.2)	0.0 (0.0)	0.0 (0.0)	0.0 (-)	F(3,48)=0.14 n.s.
Other	52	0.00 (0.00)	0-0					
Between beep events	n	x (s.d.)	range	Alone	Family	Friends	Other	
Disturbing event	29	1.28 (1.16)	1-7	1.3 (1.4)	1.3 (0.7)	1.0 (-)	1.0 (-)	F(3,25)=0.04 n.s.
Pleasant event	29	1.76 (1.46)	1-7	1.8 (1.7)	1.5 (0.9)	1.0 (-)	3.0 (-)	F(3,25)=0.41 n.s.
Social interaction	n	x (s.d.)	range	Alone	Family	Friends	Other	
Pleasant company	14	3.43 (1.55)	1-5	- - (-)	3.6 (1.5)	1.0 (-)	4.0 (-)	F(2,11)=1.44 n.s.
Prefer being alone	51	4.75 (1.53)	1-7	5.0 (1.3)	3.8 (1.7)	7.0 (0.0)	4.0 (-)	F(3,47)=4.31 p<0.009 **
Active interaction	14	1.36 (1.08)	1-5	- - (-)	1.1 (0.3)	1.0 (-)	5.0 (-)	F(2,11)=85.8 p<0.000 ***
Idiographic complaint	n	x (s.d.)	range	Alone	Family	Friends	Other	
Tension	52	4.92 (1.43)	1-7	5.1 (1.6)	4.3 (0.9)	5.5 (0.7)	4.0 (-)	F(3,48)=1.36 n.s.
Negativism/Sadness	52	4.50 (1.55)	1-7	4.7 (1.6)	4.0 (1.4)	5.0 (2.8)	3.0 (-)	F(3,48)=1.02 n.s.

Table A7.2a Experience in context: social context.

⁹⁴ The 'Friends' summary code contains the two observations with a 'health care professional'; the 'Other' code was a missing value code.

Overview: Activities

Items	n	x (s.d.)	range	Nothing (n=16)	Maintenance (n=7)	Leisure (n=28)	Other (n=1)	
Pleasant thought	48	2.04 (1.44)	1-5	1.6 (1.1)	2.5 (1.6)	2.1 (1.5)	4.0 (-)	F(3,44)=1.36 n.s.
Clear thought	48	6.65 (0.86)	4-7	6.8 (0.8)	6.7 (0.8)	6.6 (0.9)	7.0 (-)	F(3,44)=0.27 n.s.
Normal thought	48	5.33 (2.44)	1-7	5.5 (2.6)	4.7 (2.9)	5.3 (2.4)	7.0 (-)	F(3,44)=0.31 n.s.
Difficult to concentrate	48	1.88 (1.18)	1-6	1.5 (0.7)	2.8 (1.6)	1.8 (1.2)	2.0 (-)	F(3,44)=1.85 n.s.
I feel...								
Cheerful	51	1.49 (1.14)	1-6	1.8 (1.5)	1.2 (0.4)	1.4 (1.0)	1.0 (-)	F(3,47)=0.51 n.s.
Unsure	51	1.04 (0.28)	1-3	1.1 (0.5)	1.0 (0.0)	1.0 (0.0)	1.0 (-)	F(3,47)=0.72 n.s.
Lonely	51	1.00 (0.00)	1-1					
Relaxed	51	1.10 (0.50)	1-4	1.1 (0.5)	1.0 (0.0)	1.1 (0.6)	1.0 (-)	F(3,47)=0.10 n.s.
Anxious	51	1.00 (0.00)	1-1					
Satisfied	51	1.00 (0.00)	1-1					
Angry/Irritated	51	4.57 (1.65)	1-7	4.7 (1.6)	4.3 (1.6)	4.6 (1.8)	4.0 (-)	F(3,47)=0.10 n.s.
Sad	51	4.39 (1.73)	1-7	3.9 (1.9)	5.0 (1.5)	4.5 (1.7)	4.0 (-)	F(3,47)=0.68 n.s.
Guilty	51	1.00 (0.00)	1-1					
Not motivated	51	2.29 (1.60)	1-6	1.9 (1.3)	4.5 (1.0)	2.1 (1.5)	1.0 (-)	F(3,47)=5.70 p<0.002 **
Active	51	1.06 (0.24)	1-2	1.0 (0.0)	1.3 (0.5)	1.0 (0.2)	1.0 (-)	F(3,47)=3.59 p<0.020 *
Difficult	51	1.25 (0.96)	1-7	1.1 (0.3)	1.0 (0.0)	1.4 (1.3)	1.0 (-)	F(3,47)=0.44 n.s.
Challenging	49	1.00 (0.00)	1-1					
Skilled	49	5.71 (1.46)	2-7	4.9 (1.3)	5.5 (1.6)	6.3 (1.3)	6.0 (-)	F(3,45)=3.58 p<0.021 *
Hunger	51	1.25 (0.98)	1-7	1.3 (0.8)	1.0 (0.0)	1.3 (1.2)	1.0 (-)	F(3,47)=0.19 n.s.
Tired	51	2.24 (1.19)	1-5	2.2 (1.4)	2.0 (0.0)	2.3 (1.2)	2.0 (-)	F(3,47)=0.14 n.s.
Not feeling well	51	2.86 (1.60)	1-7	2.9 (1.5)	2.7 (1.2)	2.9 (1.8)	2.0 (-)	F(3,47)=0.14 n.s.
Pathology	n	x (s.d.)	range	Nothing	Maintenance	Leisure	Other	
I feel suspicious	52	1.00 (0.00)	1-1					
Thoughts ...								
...difficult to express	52	3.98 (2.68)	1-7	3.5 (2.6)	4.3 (2.8)	4.1 (2.8)	7.0 (-)	F(3,48)=0.62 n.s.
...can't get rid	51	2.39 (2.12)	1-7	2.1 (1.4)	3.3 (2.8)	2.4 (2.3)	1.0 (-)	F(3,47)=0.68 n.s.
...influenced by others	52	1.19 (0.84)	1-6	1.0 (0.0)	1.0 (0.0)	1.4 (1.1)	1.0 (-)	F(3,48)=0.77 n.s.
Feeling unreal	52	1.04 (0.19)	1-2	1.0 (0.0)	1.0 (0.0)	1.1 (0.3)	1.0 (-)	F(3,48)=0.57 n.s.
Hearing voices	52	1.00 (0.00)	1-1					
Seeing things	52	1.00 (0.00)	1-1					
Fear losing control	52	1.00 (0.00)	1-1					
Use	n	x (s.d.)	range	Nothing	Maintenance	Leisure	Other	
Nothing	52	0.48 (0.50)	0-1	0.4 (0.5)	0.4 (0.5)	0.5 (0.5)	1.0 (-)	F(3,48)=0.42 n.s.
Alcohol	52	0.10 (0.30)	0-1	0.0 (0.0)	0.1 (0.4)	0.1 (0.4)	0.0 (-)	F(3,48)=0.87 n.s.
Medication	52	0.50 (0.50)	0-1	0.6 (0.5)	0.6 (0.5)	0.5 (0.5)	0.0 (-)	F(3,48)=0.49 n.s.
Coffee	52	0.08 (0.27)	0-1	0.2 (0.4)	0.0 (0.0)	0.0 (0.2)	0.0 (-)	F(3,48)=1.37 n.s.
Tobacco	52	0.02 (0.14)	0-1	0.0 (0.0)	0.0 (0.0)	0.0 (0.2)	0.0 (-)	F(3,48)=0.27 n.s.
Other	52	0.00 (0.00)	0-0					
Between beep events	n	x (s.d.)	range	Nothing	Maintenance	Leisure	Other	
Disturbing event	29	1.28 (1.16)	1-7	1.0 (0.0)	1.0 (0.0)	1.4 (1.6)	3.0 (-)	F(3,25)=1.09 n.s.
Pleasant event	29	1.76 (1.46)	1-7	1.3 (0.7)	2.3 (1.6)	1.9 (1.7)	1.0 (-)	F(3,25)=0.73 n.s.
Social interaction	n	x (s.d.)	range	Nothing	Maintenance	Leisure	Other	
Pleasant company	14	3.43 (1.55)	1-5	2.5 (1.7)	4.0 (-)	3.6 (1.5)	5.0 (-)	F(3,10)=0.88 n.s.
Prefer being alone	51	4.75 (1.53)	1-7	4.8 (1.7)	5.0 (1.4)	4.8 (1.4)	1.0 (-)	F(3,47)=2.21 p<0.098
Active interaction	14	1.36 (1.08)	1-5	1.0 (0.0)	5.0 (-)	1.1 (0.4)	1.0 (-)	F(3,10)=54.6 p<0.000 ***
Idiogr. complaint	n	x (s.d.)	range	Nothing	Maintenance	Leisure	Other	
Tension	52	4.92 (1.43)	1-7	5.2 (1.0)	4.9 (1.5)	4.8 (1.6)	4.0 (-)	F(3,48)=0.36 n.s.
Negativism/Sadness	52	4.50 (1.55)	1-7	4.1 (1.7)	4.7 (1.4)	4.7 (1.6)	4.0 (-)	F(3,48)=0.50 n.s.

Table A7.2b Experience in context: activities.

Overview: Places

Items	n	x (s.d.)	range	Home (n=42)	Network (n=8)	Other (n=2)	
Pleasant thought	48	2.04 (1.44)	1-5	1.7 (1.3)	3.1 (1.4)	4.5 (0.7)	F(2,45)=8.28 p<0.001 **
Clear thought	48	6.65 (0.86)	4-7	6.8 (0.7)	6.1 (1.1)	6.0 (1.4)	F(2,45)=2.73 p<0.074
Normal thought	48	5.33 (2.44)	1-7	5.2 (2.5)	5.9 (2.2)	6.0 (1.4)	F(2,45)=0.33 n.s.
Difficult to concentrate	48	1.88 (1.18)	1-6	1.8 (1.1)	1.9 (0.6)	4.0 (2.8)	F(2,45)=3.84 p<0.028 *
I feel...							
Cheerful	51	1.49 (1.14)	1-6	1.6 (1.2)	1.1 (0.4)	1.5 (0.7)	F(2,48)=0.48 n.s.
Unsure	51	1.04 (0.28)	1-3	1.0 (0.3)	1.0 (0.0)	1.0 (0.0)	F(2,48)=0.12 n.s.
Lonely	51	1.00 (0.00)	1-1				
Relaxed	51	1.10 (0.50)	1-4	1.1 (0.6)	1.0 (0.0)	1.0 (0.0)	F(2,48)=0.23 n.s.
Anxious	51	1.00 (0.00)	1-1				
Satisfied	51	1.00 (0.00)	1-1				
Angry/Irritated	51	4.57 (1.65)	1-7	4.8 (1.7)	3.9 (0.8)	3.0 (1.4)	F(2,48)=2.02 n.s.
Sad	51	4.39 (1.73)	1-7	4.5 (1.9)	3.9 (0.6)	4.0 (0.0)	F(2,48)=0.50 n.s.
Guilty	51	1.00 (0.00)	1-1				
Not motivated	51	2.29 (1.60)	1-6	2.1 (1.6)	2.5 (1.4)	4.5 (0.7)	F(2,48)=2.24 n.s.
Active	51	1.06 (0.24)	1-2	1.0 (0.2)	1.0 (0.0)	2.0 (0.0)	F(2,48)=45.46 p<0.000 ***
Difficult	51	1.25 (0.96)	1-7	1.3 (1.1)	1.1 (0.4)	1.0 (0.0)	F(2,48)=0.17 n.s.
Challenging	49	1.00 (0.00)	1-1				
Skilled	49	5.71 (1.46)	2-7	5.7 (1.5)	6.0 (1.3)	4.0 (0.0)	F(2,46)=1.58 n.s.
Hunger	51	1.25 (0.98)	1-7	1.3 (1.1)	1.3 (0.5)	1.0 (0.0)	F(2,48)=0.07 n.s.
Tired	51	2.24 (1.19)	1-5	2.4 (1.3)	1.6 (0.5)	2.0 (0.0)	F(2,48)=1.35 n.s.
Not feeling well	51	2.86 (1.60)	1-7	3.1 (1.7)	1.9 (0.6)	2.5 (2.1)	F(2,48)=2.01 n.s.
Pathology	n	x (s.d.)	range	Home	Network	Other	
I feel suspicious	52	1.00 (0.00)	1-1				
Thoughts ...							
...difficult to express	52	3.98 (2.68)	1-7	4.4 (2.7)	2.8 (2.0)	1.0 (0.0)	F(2,49)=2.67 p<0.078
...can't get rid	51	2.39 (2.12)	1-7	2.7 (2.2)	1.1 (0.4)	1.0 (0.0)	F(2,48)=2.45 p<0.095
...influenced by others	52	1.19 (0.84)	1-6	1.2 (0.9)	1.0 (0.0)	1.0 (0.0)	F(2,49)=0.31 n.s.
Feeling unreal	52	1.04 (0.19)	1-2	1.0 (0.2)	1.0 (0.0)	1.0 (0.0)	F(2,49)=0.24 n.s.
Hearing voices	52	1.00 (0.00)	1-1				
Seeing things	52	1.00 (0.00)	1-1				
Fear losing control	52	1.00 (0.00)	1-1				
Use	n	x (s.d.)	range	Home	Network	Other	
Nothing	52	0.48 (0.50)	0-1	0.5 (0.5)	0.8 (0.5)	0.0 (0.0)	F(2,49)=2.21 n.s.
Alcohol	52	0.10 (0.30)	0-1	0.1 (0.3)	0.0 (0.0)	0.0 (0.0)	F(2,49)=0.64 n.s.
Medication	52	0.50 (0.50)	0-1	0.5 (0.5)	0.3 (0.5)	1.0 (0.0)	F(2,49)=2.09 n.s.
Coffee	52	0.08 (0.27)	0-1	0.1 (0.3)	0.0 (0.0)	0.0 (0.0)	F(2,49)=0.50 n.s.
Tobacco	52	0.02 (0.14)	0-1	0.0 (0.2)	0.0 (0.0)	0.0 (0.0)	F(2,49)=0.11 n.s.
Other	52	0.00 (0.00)	0-0				
Between beep events	n	x (s.d.)	range	Home	Network	Other	
Disturbing event	29	1.28 (1.16)	1-7	1.3 (1.2)	1.7 (1.2)	1.0 (0.0)	F(2,26)=0.22 n.s.
Pleasant event	29	1.76 (1.46)	1-7	1.8 (1.5)	1.7 (1.2)	2.0 (1.4)	F(2,26)=0.03 n.s.
Social interaction	n	x (s.d.)	range	Home	Network	Other	
Pleasant company	14	3.43 (1.55)	1-5	1.6 (0.9)	4.4 (0.5)	4.5 (0.7)	F(2,11)=26.43 p<0.000 ***
Prefer being alone	51	4.75 (1.53)	1-7	5.1 (1.3)	2.8 (1.0)	4.5 (0.7)	F(2,48)=11.70 p<0.000 ***
Active interaction	14	1.36 (1.08)	1-5	1.0 (0.0)	1.0 (0.0)	3.5 (2.1)	F(2,11)=13.10 p<0.002 **
Idiographic complaint	n	x (s.d.)	range	Home	Network	Other	
Tension	52	4.92 (1.43)	1-7	5.1 (1.5)	4.0 (0.8)	4.0 (0.0)	F(2,49)=2.77 p<0.070
Negativism/Sadness	52	4.50 (1.55)	1-7	4.7 (1.7)	3.9 (0.6)	3.5 (0.7)	F(2,49)=1.32 n.s.

Table A7.2c Experience in context: places.

D. QUALITY OF LIFE

QoL-Index computed using mean differences of z-scored Likert scales.
[M] cognition, mood and activity assessments; [P] pathology items; (S.D.≥0.75).

“Quality of Life”-Index: split-half technique

Difference of contrasted situations 50% (low) - 50% (high) frequency activity codes.

High frequency activities used (max. 4):
45 TV, radio... 38%
01 Sleep, nap, rest 15%

[M] Items	n	x (s.d.)	range	Low (n=24)	High (n=28)	Δ=	
Pleasant thought	48	2.04 (1.44)	1-5	0.0 (1.0)	0.0 (1.0)	-0.07	F(1,46)=0.05 n.s.
Clear thought	48	6.65 (0.86)	4-7	0.0 (1.0)	0.0 (1.0)	-0.04	F(1,46)=0.02 n.s.
Normal thought	48	5.33 (2.44)	1-7	-0.1 (1.1)	0.1 (1.0)	0.14	F(1,46)=0.22 n.s.
Difficult to concentrate	48	1.88 (1.18)	1-6	-0.1 (1.2)	0.1 (0.9)	0.12	F(1,46)=0.16 n.s.
I feel ..							
Cheerful	51	1.49 (1.14)	1-6	0.0 (1.2)	0.0 (0.9)	-0.05	F(1,49)=0.03 n.s.
Angry/Irritated	51	4.57 (1.65)	1-7	-0.1 (1.0)	0.1 (1.0)	0.14	F(1,49)=0.24 n.s.
Sad	51	4.39 (1.73)	1-7	0.0 (1.1)	0.0 (1.0)	0.00	F(1,49)=0.00 n.s.
Not motivated	51	2.29 (1.60)	1-6	-0.2 (1.1)	0.2 (0.9)	0.36	F(1,49)=1.63 n.s.
Difficult	51	1.25 (0.96)	1-7	-0.1 (1.3)	0.0 (0.7)	0.09	F(1,49)=0.11 n.s.
Skilled	49	5.71 (1.46)	2-7	-0.4 (1.1)	0.3 (0.8)	0.70	F(1,47)=6.66 p<0.013 *
Not feeling well	51	2.86 (1.60)	1-7	0.1 (0.8)	-0.1 (1.1)	-0.14	F(1,49)=0.25 n.s.

ESM-QoLn[M] = 0.11 ESM-QoLn[M]n = 54

[P] Items	n	x (s.d.)	range	Low (n=24)	High (n=28)	Δ=	
Thoughts ...							
...difficult to express	52	3.98 (2.68)	1-7	-0.1 (1.0)	0.1 (1.0)	0.16	F(1,50)=0.32 n.s.
...can't get rid	51	2.39 (2.12)	1-7	-0.1 (1.1)	0.1 (0.9)	0.21	F(1,49)=0.54 n.s.
...influenced by others	52	1.19 (0.84)	1-6	0.2 (0.0)	-0.2 (1.3)	-0.42	F(1,50)=2.39 n.s.
Tension	52	4.92 (1.43)	1-7	-0.1 (0.9)	0.1 (1.1)	0.21	F(1,50)=0.56 n.s.
Negativism/Sadness	52	4.50 (1.55)	1-7	0.0 (1.0)	0.0 (1.0)	-0.05	F(1,50)=0.03 n.s.

ESM-QoLn[P] = 0.02 ESM-QoLn[P]n = 50

Table A7.3a Quality of Life: High against low frequency behaviors.

“Quality of Life”-Index: contrasted situations

Difference of contrasted situations: 33% (low) - 33% (high) frequency activity codes.

High frequency activities used (max. 4):

45 TV, radio... 38%

Low frequency activities used :

46 Read book, newspaper 12%
27 Personal hygiene 8%
31 Personal mental act. 4%
26 Personal medical care 2%
21 Household chores 2%
43 Hobbies, art... 2%
51 Conversation 2%
29 Other maintenance 2%
42 Taking a walk 2%

[M] Items	n	x (s.d.)	range	Low (n=17)	Medium (n=15)	High (n=20)	Δ =
Pleasant thought	48	2.04 (1.44)	1-5	0.1 (1.1)	-0.2 (0.9)	0.1 (1.0)	-0.03 F(2,45)=0.52 n.s.
Clear thought	48	6.65 (0.86)	4-7	0.0 (1.0)	0.0 (1.1)	0.0 (0.9)	-0.02 F(2,45)=0.01 n.s.
Normal thought	48	5.33 (2.44)	1-7	-0.1 (1.1)	0.2 (0.9)	0.0 (1.0)	0.10 F(2,45)=0.43 n.s.
Difficult to concentrate	48	1.88 (1.18)	1-6	-0.3 (1.3)	0.3 (0.6)	0.0 (1.0)	0.26 F(2,45)=1.30 n.s.
I feel...							
Cheerful	51	1.49 (1.14)	1-6	-0.2 (0.7)	0.1 (1.3)	0.1 (1.0)	0.31 F(2,48)=0.51 n.s.
Angry/Irritated	51	4.57 (1.65)	1-7	-0.2 (0.9)	0.0 (1.0)	0.1 (1.1)	0.32 F(2,48)=0.44 n.s.
Sad	51	4.39 (1.73)	1-7	-0.2 (1.0)	0.1 (1.0)	0.1 (1.0)	0.22 F(2,48)=0.34 n.s.
Not motivated	51	2.29 (1.60)	1-6	-0.4 (1.2)	0.2 (0.8)	0.2 (0.9)	0.58 F(2,48)=1.95 n.s.
Difficult	51	1.25 (0.96)	1-7	-0.2 (1.6)	0.1 (0.4)	0.1 (0.7)	0.25 F(2,48)=0.44 n.s.
Skilled	49	5.71 (1.46)	2-7	0.0 (1.1)	-0.7 (0.8)	0.6 (0.7)	0.60 F(2,46)=8.36 p<0.001 **
Not feeling well	51	2.86 (1.60)	1-7	0.0 (0.6)	0.1 (1.0)	-0.1 (1.2)	-0.05 F(2,48)=0.08 n.s.

ESM-QoLc[M] = 0.23 ESM-QoLc[M]n = 59

[P] Items	n	x (s.d.)	range	Low (n=17)	Medium (n=15)	High (n=20)	Δ =
...troublesome ...	52	3.98 (2.68)	1-7	-0.4 (0.9)	0.3 (1.0)	0.1 (1.1)	0.43 F(2,49)=1.97 n.s.
...can't get rid	51	2.39 (2.12)	1-7	-0.3 (1.2)	0.2 (0.7)	0.1 (1.0)	0.37 F(2,48)=0.86 n.s.
...influenced by others	52	1.19 (0.84)	1-6	0.2 (0.0)	0.2 (0.0)	-0.4 (1.6)	-0.59 F(2,49)=2.28 n.s.
Tension	52	4.92 (1.43)	1-7	-0.1 (1.0)	-0.1 (0.7)	0.2 (1.2)	0.33 F(2,49)=0.59 n.s.
Negativism/Sadness	52	4.50 (1.55)	1-7	-0.1 (1.0)	0.1 (1.0)	0.0 (1.1)	0.16 F(2,49)=0.24 n.s.

ESM-QoLc[P] = 0.14 ESM-QoLc[P]n = 55

Table A7.3b Quality of Life: Contrasted situations: high against low 33% frequency behaviors.

“Quality of Life”-Index: “Flow”⁹⁵

Flow or pursue of positive: Difference of contrasted situations using z-scored c scores Lo/Lo - Hi/Hi. based on the balance of z-scored challenge & skills: criterion is positive/negative value

[M]Items:	n	x (s.d.)	range	Lo/Lo (n=19)	Lo/Hi (n=33)	Δ=	
Pleasant thought	48	2.04 (1.44)	1-5	0.1 (1.0)	0.0 (1.0)	-0.09	F(1,46)= 0.08 n.s.
Clear thought	48	6.65 (0.86)	4-7	-0.1 (1.1)	0.0 (1.0)	0.14	F(1,46)= 0.22 n.s.
Normal thought	48	5.33 (2.44)	1-7	0.1 (1.0)	0.0 (1.0)	-0.14	F(1,46)= 0.21 n.s.
Difficult to concentrate	48	1.88 (1.18)	1-6	-0.2 (1.1)	0.1 (1.0)	0.24	F(1,46)= 0.64 n.s.
I feel...							
Cheerful	51	1.49 (1.14)	1-6	0.0 (1.1)	0.0 (0.9)	0.02	F(1,49)= 0.01 n.s.
Angry/Irritated	51	4.57 (1.65)	1-7	0.2 (1.2)	-0.1 (0.9)	-0.24	F(1,49)= 0.70 n.s.
Sad	51	4.39 (1.73)	1-7	0.1 (1.1)	0.0 (0.9)	-0.12	F(1,49)= 0.16 n.s.
Not motivated	51	2.29 (1.60)	1-6	-0.2 (1.1)	0.1 (0.9)	0.39	F(1,49)= 1.82 n.s.
Difficult	51	1.25 (0.96)	1-7	0.2 (0.3)	-0.1 (1.2)	-0.25	F(1,49)= 0.74 n.s.
Skilled	49	5.71 (1.46)	2-7	-1.2 (0.4)	0.7 (0.3)	N.A.	F(1,47)=<.<.< p<0.000 ***
Not feeling well	51	2.86 (1.60)	1-7	-0.2 (1.2)	0.1 (0.9)	0.29	F(1,49)= 1.03 n.s.

ESM -Flow[M] = 0.03 ESM -Flow[M]n = 51

[P]Items:	n	x (s.d.)	range	Lo/Lo (n=19)	Lo/Hi (n=33)	Δ=	
Thoughts ...	52	3.98 (2.68)	1-7	0.4 (0.9)	-0.2 (1.0)	-0.58	F(1,50)= 4.28 p<0.041 *
... difficult to express	51	2.39 (2.12)	1-7	-0.1 (1.0)	0.1 (1.0)	0.14	F(1,49)= 0.23 n.s.
... can't get rid	52	1.19 (0.84)	1-6	0.2 (0.0)	-0.1 (1.2)	-0.36	F(1,50)= 1.58 n.s.
... influenced by others							
Tension	52	4.92 (1.43)	1-7	0.1 (1.0)	-0.1 (1.0)	-0.15	F(1,50)= 0.26 n.s.
Negativism/Sadness	52	4.50 (1.55)	1-7	0.1 (1.2)	-0.1 (0.9)	-0.19	F(1,50)= 0.42 n.s.

ESM -Flow[P] = -0.23 ESM -Flow[P]n = 41

Table A7.3c Quality of Life: “Flow” theory.

⁹⁵ The item ‘challenged’ was always scored 1 (therefore, with no variability) and as a consequence, only the ‘skills’ will be used to define the moments.

“Quality of Life”-Index: “Escape”

“Escape of boredom”: Difference of contrasted situations using z-scored challenge and skill items Lo/Lo until +.40 /+ .40 compared to the Rest.

[M]Items:	n	x (s.d.)	range	Lo&Lo (n=25)	Hi&Hi (n=27)	Δ =	
Pleasant thought	48	2.04 (1.44)	1-5	0.1 (1.0)	-0.1 (1.0)	-0.18	F(1,46)= 0.38 n.s.
Clear thought	48	6.65 (0.86)	4-7	-0.1 (1.2)	0.1 (0.8)	0.21	F(1,46)= 0.54 n.s.
Normal thought	48	5.33 (2.44)	1-7	0.0 (1.0)	0.0 (1.0)	-0.06	F(1,46)= 0.04 n.s.
Difficult to concentrate	48	1.88 (1.18)	1-6	0.0 (1.0)	0.0 (1.0)	-0.01	F(1,46)= 0.00 n.s.
I feel...							
Cheerful	51	1.49 (1.14)	1-6	0.0 (1.0)	0.0 (1.0)	0.09	F(1,49)= 0.09 n.s.
Angry/Irritated	51	4.57 (1.65)	1-7	0.2 (1.1)	-0.1 (0.9)	-0.30	F(1,49)= 1.11 n.s.
Sad	51	4.39 (1.73)	1-7	0.1 (1.0)	-0.1 (1.0)	-0.13	F(1,49)= 0.20 n.s.
Not motivated	51	2.29 (1.60)	1-6	-0.1 (1.0)	0.1 (1.0)	0.28	F(1,49)= 0.97 n.s.
Difficult	51	1.25 (0.96)	1-7	-0.2 (1.3)	0.1 (0.6)	0.30	F(1,49)= 1.13 n.s.
Skilled	49	5.71 (1.46)	2-7	-0.8 (0.7)	0.9 (0.0)	N.A.	F(1,47)=<.<< p<0.000 ***
Not feeling well	51	2.86 (1.60)	1-7	-0.1 (1.1)	0.1 (0.9)	0.22	F(1,49)= 0.60 n.s.

ESM-Escape[M] = 0.04 ESM-Escape[M]n = 52

[P]Items:	n	x (s.d.)	range	Lo&Lo (n=25)	Hi&Hi (n=27)	Δ =	
Thoughts ...							
...difficult to express	52	3.98 (2.68)	1-7	0.2 (0.9)	-0.2 (1.0)	-0.39	F(1,50)= 2.01 n.s.
...can't get rid	51	2.39 (2.12)	1-7	0.1 (0.9)	-0.1 (1.1)	-0.14	F(1,49)= 0.25 n.s.
...influenced by others	52	1.19 (0.84)	1-6	0.1 (0.7)	-0.1 (1.2)	-0.17	F(1,50)= 0.35 n.s.
Tension	52	4.92 (1.43)	1-7	0.2 (0.9)	-0.2 (1.1)	-0.33	F(1,50)= 1.41 n.s.
Negativism/Sadness	52	4.50 (1.55)	1-7	0.1 (1.0)	-0.1 (1.0)	-0.17	F(1,50)= 0.39 n.s.

ESM-Escape[P] = -0.24 ESM-Escape[P]n = 41

Table A7.3d Quality of Life: “Escape” theory.

Bibliography

.....

- Addington, D., Addington, J. & Schissel, B. (1990). A depression rating scale for schizophrenics. *Schizophrenia Research*, 3(4), 247-251.
- Adler, D.A., Berlant, J., Dochery, J.P., Dorwart, R.A., Drake, R.E., Ellison, J.M., Goldman, H.H., Lehman, A.F., Pajer, K.A., Siris, S.G. & Lehrer, D. (1992). *Beyond Symptom Suppression: Improving Long-Term Outcomes of Schizophrenia: GAP Report 134*. Washington: American Psychiatric Press, Inc.
- Alford, B.D. & Correia, C.J. (1994). Cognitive therapy of schizophrenia: Theory and empirical status. *Behavior Therapy*, 25, 17-33.
- Alliger, G.M. & Williams, K.J. (1993). Using signal-contingent experience sampling methodology to study work in the field: A discussion and illustration examining task perceptions and mood. *Personnel Psychology*, 46, 525-549.
- Allison, P.D. (1984). *Event History Analysis*. Beverly Hills: Sage.
- Altergott, K. (1988). *Daily life in Later Life. Comparative Perspectives*. Beverly Hills, CA: Sage Publications.
- Altergott, K. & Duncan, S. (1988). National Context and Daily life in Later Life. In K. Altergott (Ed.), *Daily life in Later Life. Comparative Perspectives*, (pp. 229-249). Beverly Hills, CA: Sage Publications.
- American Psychiatric Association (APA). (1982). *DSM-III. Diagnostic and Statistical Manual of Mental Disorders*. Washington D.C.: American Psychiatric Association (APA).
- American Psychiatric Association (APA). (1987). *DSM-III-R. Diagnostic and Statistical Manual of Mental Disorders*. Washington D.C.: American Psychiatric Association (APA).
- American Psychiatric Association (APA). (1987). *Quick Reference to the Diagnostic Criteria from DSM-III-R*. Washington: American Psychiatric Association.
- American Psychiatric Association (APA). (1991). *The DSM-IV Options Book: Work in Progress*. Washington DC: APA.
- American Psychiatric Association (APA). (1993). *The DSM-IV Draft Criteria*. Washington DC: APA.
- American Psychiatric Association (APA). (1994). *DSM-IV. Diagnostic and Statistical Manual of Mental Disorders*. (4 ed.). Washington DC: American Psychiatric Press.
- Anastasi, A. (1976). *Psychological Testing*. (4 ed.). New York: MacMillan Publishing Co.
- Anderson, C.M., Reiss, D.J. & Hogarty, G.E. (1986). *Schizophrenia and the Family: a Practitioner's Guide to Psychoeducation and Management*. New York: The Guilford Press.
- Andreasen, N.C. (1990). Positive and negative symptoms: Historical and conceptual aspects. In N.C. Andreasen (Ed.), *Schizophrenia: Positive and Negative Symptoms and Syndromes*, (Vol. 24, pp. 1-42). New York: Karger.
- Andreasen, N.C. (Ed.). (1990). *Schizophrenia: Positive and Negative Symptoms and Syndromes*. (Vol. 24). New York: Karger.
- Andreasen, N.C. (Ed.). (1994). *Schizophrenia: From Mind to Molecule*. Washington: American Psychiatric Press, Inc.
- Andreasen, N.C. & Carpenter, W.T. (1993). Diagnosis and classification of schizophrenia. *Schizophrenia Bulletin*, 19(2), 199-214.
- Andreasen, N.C. & Flaum, M. (1991). Schizophrenia: The characteristic symptoms. *Schizophrenia Bulletin*, 17(1), 27-49.
- Andreasen, N.C., Flaum, M., Swayze, V.W., Tyrrell, G. & Arndt, S. (1990). Positive and negative symptoms in schizophrenia. A critical reappraisal. *Archives of General Psychiatry*, 47(July), 615-621.
- Andreasen, N.C. & Olson, S.A. (1982). Negative versus positive schizophrenia: Definition and validation. *Archives of General Psychiatry*, 39, 789-794.
- Andreasen, N.C., Olson, S.A., Dennert, J.W. & Smith, M.R. (1982). Ventricular enlargement in schizophrenia: Relation to positive and negative symptoms. *American Journal of Psychiatry*, 139, 297-302.
- Andreasen, N.C., Swayze III, V.W., Flaum, M., O'Leary, D.S. & Alliger, R. (1994). The neural mechanisms of mental phenomena. In N.C. Andreasen (Ed.), *Schizophrenia: From Mind to Molecule*, (pp. 49-92). Washington: American Psychiatric Press, Inc.
- Andresen, B., Stark, F.M. & Gross, J. (Eds.). (1992). *Mensch, Psychiatrie, Umwelt: Ökologische Perspektiven für die soziale Praxis*. Bonn: Psychiatrie-Verlag.
- Annals of Behavioral Medicine. (1993). Mini series on daily recording and ambulatory monitoring. *Annals of Behavioral Medicine*, 15(1), 3-50.
- Anthony, W.A., Cohen, M. & Farkas, M.D. (Eds.). (1990). *Psychiatric Rehabilitation*. Boston, MA: Boston University, Center for Psychiatric Rehabilitation.

- Argyle, M. (Ed.). (1992). *The Social Psychology of Everyday Life*. Londen: Routledge.
- Arieti, S. (1974). *Interpretation of Schizophrenia*. (2 ed.). New York: Basic Books.
- Ås, D. (1978). Studies of time use: Problems and prospects. *Acta Sociologica*, 21, 125-141.
- Asaad, G. (1990). *Hallucinations in Clinical Psychiatry. A Guide for Mental Health Professionals*. (Vol. 2). New York: Brunner/Mazel.
- Asselbergs, L. (1986). *Ambulante behandeling van schizofrenie. Een literatuurstudie*. (Vol. 91). Utrecht: NcGv.
- Asselbergs, L. (1989). *Sociaal Beperkt. Het Sociaal Functioneren van Schizofreniepatiënten*. (Vol. 89-7). Utrecht: NcGv.
- Awad, A.G. (1993). Methodological and design issues in clinical trials of new neuroleptics: An overview. *British Journal of Psychiatry*, 163(Supplement 22), 51-57.
- Baars, H.M.J. (1994). *Sociale Netwerken van Ambulante Chronische Psychiatrische Patiënten*. Unpublished doctoraatsthesis, Rijksuniversiteit Limburg te Maastricht.
- Bachrach, L. (1984). The homeless mentally ill and mental health services: an analytic review of the literature. In H.R. Lamb (Ed.), *The Homeless Mentally Ill*, (pp. 11-54). Washington DC: APA Press.
- Bachrach, L.L. (1987). *Leona Bachrach speaks: Selected Speeches and Lectures*. (Vol. 35). San Francisco: Jossey-Bass Inc.
- Bakeman, R. & Gottman, J.M. (1986). *Observing Interaction: an Introduction to Sequential Analysis*. (1 ed.). Cambridge: Cambridge University Press.
- Baldessarini, R. (1985). *Chemotherapy in Psychiatry*. Cambridge MA: Harvard University Press.
- Banaji, M.R. & Crowder, R.G. (1989). The bankruptcy of everyday memory. *American Psychologist*, 44, 1185-1193.
- Bannister, D. (1977). The logical requirements of research into schizophrenia. *Schizophrenia Bulletin*, 4(2), 72-77.
- Barge-Schaapveld, D.Q.C.M., Nicolson, N.A. & deVries, M.W. (1994,). *Effects of Antidepressant Treatment on Subjective Daily Life Experience*. Paper presented at "Quality of Life and Disabilities in Mental Disorders - 7th European Symposium of the Association of European Psychiatrists - Section of Psychiatric Epidemiology and Social Psychiatry", Vienna, Austria.
- Barker, R.G. (1968). *Ecological Psychology: Concepts and methods for studying the environment of human behavior*. Stanford CA: Stanford University Press.
- Barker, R.G. (1978). *Habitats, Environments, and Human Behavior: Studies in Ecological Psychology and Eco-Behavioral Science from the Midwest Psychological Field Station, 1947-1972*. San Francisco: Jossey-Bass.
- Barrett, B.H., Johnston, J.M. & Pennypacker, H.S. (1986). Behavior: its units, dimensions, and measurement. In R.O. Nelson & S.C. Hayes (Eds.), *Conceptual Foundations of Behavioral Assessment*, (pp. 156-200). New York: The Guilford Press.
- Barrowclough, C., Johnston, M. & Tarrier, N. (1994). Attributions, expressed emotion, and patient relapse: An attributional model of relatives' response to schizophrenic illness. *Behavior Therapy*, 25, 67-88.
- Barrowclough, C. & Tarrier, N. (1992). *Families of Schizophrenic Patients. Cognitive Behavioral Interventions*. London: Chapman & Hall.
- Barrowclough, C. & Tarrier, N. (1994). Interventions with families. In M. Birchwood & N. Tarrier (Eds.), *Psychological Management of Schizophrenia*, (pp. 53-76). Chichester: John Wiley & Sons.
- Barton, R. (1976 (1959)). *Institutional Neurosis*. (3 ed.). Bristol: John Wright & Sons Ltd.
- Bateson, G. (1972). *Steps to an Ecology of the Mind: A Revolutionary Approach to Man's Understanding of Himself*. New York: Ballantine Books.
- Beckers, B.M.C. (1986). *Persoonlijke Vaardigheidstherapieën voor Kansarmen*. Rijksuniversiteit Limburg (Maastricht).
- Beckers, R.J.M. (1982). *Interpersoonlijke Vaardigheidstherapieën voor Kansarmen*. Rijksuniversiteit Limburg (Maastricht).
- Bellack, A.E. (1993,). (...). Paper presented at "Auf dem Weg zu einer integrativen Therapie der Schizophrenie/ Towards a Comprehensive Therapy of Schizophrenia", Bern, Schweiz.
- Bellack, A.E. & Mueser, K.T. (1993). Psychosocial treatment for schizophrenia. *Schizophrenia Bulletin*, 19(2), 317-336.
- Bellack, A.S. (1986). Schizophrenia: Behavior therapy's forgotten child. *Behavior Therapy*, 17, 199-214.
- Bellack, L. (Ed.). (1979). *Disorders of the Schizophrenic Syndrome*. New York: Basic Books, Inc, Publishers.
- Bentall, R.P. (1990). *Reconstructing Schizophrenia*. London: Routledge.
- Bentall, R.P. (1992). The Classification of schizophrenia. In D.J. Kavanagh (Ed.), *Schizophrenia. An Overview and Practical Handbook*, (pp. 23-44). London: Chapman & Hall.

- Bentall, R.P., Haddock, G. & Slade, P.D. (1994). Cognitive behavior therapy for persistent auditory hallucinations: From theory to therapy. *Behavior Therapy*, 25, 51-66.
- Bentall, R.P., Kinderman, P. & Kaney, S. (1994). The self, attributional processes and abnormal beliefs: Towards a model of persecutory hallucinations. *Behavior Research and Therapy*, 32(3), 331-341.
- Bergson, H.-L. (1907). *L'Evolution Créatrice*.
- Bhugra, D. (1991). *Psychiatric Care of the Homeless Mentally Ill. Report of a Working Party*. London: Royal College of Psychiatry.
- Bigelow, D., Brodsky, G., Steward, L. & Olson, M. (1982). The concept and measurement of quality of life as a dependent variable in evaluation of mental health services. In G.J. Stahler & W.R. Tash (Eds.), *Innovative Approaches to Mental Health Evaluation*, (pp. 345-366). New York: Academic Press.
- Binkley, S.A. (1990). *The Clockwork Sparrow: Time, Clocks, and Calendars in Biological Organisms*. Englewood Cliffs: Prentice Hall.
- Birchwood, M., Smith, J., Drury, V., Healy, J., Macmillan, F. & Slade, M.A. (1994). A self-report Insight Scale (IS) for psychosis: Reliability, validity and sensitivity to change. *Acta Psychiatrica Scandinavica*, 89, 62-67.
- Birchwood, M.J., Hallett, S.E. & Preston, M.C. (1988). *Schizophrenia: An Integrated Approach to Research and Treatment*. London: Longman.
- Blanchard, J.J., Mueser, K.T. & Bellack, A.S. (1992). Self- and interview-rated negative mood states in schizophrenia: Their convergence and prediction of thought disturbance. *Journal of Psychopathology and Behavioral Assessment*, 14(3), 277-291.
- Bleuler, E. (1911 (1950)). *Dementia Praecox, or the Group of Schizophrenias*. New York: International Universities Press.
- Bleuler, E. (1920). *Lehrbuch der Psychiatrie*. (3 ed.). Berlin: Verlag van Jules Springer.
- Bleuler, M. (1972). *Die schizophrenen Geistesstörungen im Lichte langjähriger Kranken- und Familiengeschichten*. Stuttgart: Thieme.
- Bleuler, M. (1978). *The Schizophrenic Disorders: Long-Term Patient and Family Studies*. New Haven: Yale University Press.
- Bornstein, P.H., Hamilton, S.B. & Tepper-Bornstein, M. (1986). Self-monitoring procedures. In A.R. Ciminero, K.S. Calhoen & H.E. Adams (Eds.), *Handbook of Behavioral Assessment*, (2 ed., pp. 176-222). New York: John Wiley & Sons.
- Bornstein, P.H., Mungas, D.M., Quevillon, R.P., Kniivila, C.M., Miller, R.K. & Holombo, L.K. (1978). Self-monitoring training: Effects on reactivity and accuracy of self-observation. *Behavioral Research and Therapy*, 16, 217-220.
- Braff, D.L. (1993). Information processing and attention dysfunctions in schizophrenia. *Schizophrenia Bulletin*, 19(2), 233-260.
- Brandstätter, H. (1983). Emotional responses to other persons in everyday life situations. *Journal of Personality and Social Psychology*, 45, 871-883.
- Brandstätter, H. (1991). Emotions in everyday life situations. Time sampling of subjective experience. In M. Argyle (Ed.), *Subjective Well-Being: an Interdisciplinary Perspective*, (pp. 173-192). Oxford: Pergamon Press.
- Brenner, C. (1975). Affects and psychic conflict. *Psychoanalytic Quarterly*, 44, 5-28.
- Breslin, N.A. & Weinberger, D.R. (1990). Schizophrenia and the normal functional development of the prefrontal cortex. *Development and Psychopathology*, 2, 409-424.
- Bronfenbrenner, U. (1979). *The Ecology of Human Development: Experiments by nature and design*. Cambridge: Harvard University Press.
- Brown, G.W. & Harris, T. (1978). *Social Origins of Depression*. New York: Free Press.
- Brown, G.W. & Rutter, M. (1966). The measurement of family activities and relationships: A methodological study. *Human Relations*, 19, 241-263.
- Bruner, J. (1990). *Acts of Meaning*. Cambridge: Harvard University Press.
- Brunswik, E.C. (1956). *Perception and the Representative Design of Psychological Experiments*. Berkeley, Calif.: University of California Press.
- Bruton, C.J., Crow, T.J., Frith, C.D., Johnstone, E.C., Owens, D.G.C. & Roberts, G.W. (1990). Schizophrenia and the brain: A prospective clinico-neuropathological study. *Psychological Medicine*, 20, 285-304.
- Buchsbaum, M.S. & Haler, R.J. (1987). Functional and anatomical brain imaging: Impact on schizophrenia research. In D. Shore (Ed.), *Special Report: Schizophrenia 1987*, (pp. 129-146). Rockville, M: National Institute of Mental Health.
- Cacioppo, J.T., Klein, D.J., Berntson, G.G. & Hatfield, E. (1993). The psychophysiology of emotion. In M. Lewis & J.M. Haviland (Eds.), *Handbook of Emotions*, (pp. 119-142). New York: The Guilford Press.

- Campbell, D.T. & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multi-method matrix. *Psychological Bulletin*, 16, 81-105.
- Campbell, D.T. & Stanley, J.C. (1963 (1966)). *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally.
- Camras, L.A., Holland, E.A. & Patterson, M.J. (1993). Facial expression. In M. Lewis & J.M. Haviland (Eds.), *Handbook of Emotions*, (pp. 199-208). New York: The Guilford Press.
- Cannon, M., Cotter, D., Sham, P.C., Larkin, C., Murray, R.M., Coffey, V.P. & O'Callaghan, E. (1994). Schizophrenia in an Irish sample following prenatal exposure to the 1957 influenza epidemic: A case-controlled, prospective follow-up study. *Schizophrenia Research*, 11(2), 95.
- Carlsson, A. (1994). The search for the ideal medications: Developing a rational neuropharmacology. In N.C. Andreasen (Ed.), *Schizophrenia: From Mind to Molecule*, (pp. 161-172). Washington: American Psychiatric Press, Inc.
- Carpenter, W. & Heinrichs, D.W. (1987). On the heterogeneity of schizophrenia. In M. Alpert (Ed.), *Controversies in Schizophrenia: Changes and Constanties. Proceedings of the 74th Annual Meeting of the American Psychopathological Association*. New York City, March 1-3, 1984, (pp. 25-37). New York: The Guilford Press.
- Carpenter, W., Strauss, J. & Bartko, J. (1974). Uses of signs and symptoms for the identification of schizophrenic patients. *Schizophrenia Bulletin*, 11, 37-49.
- Carpenter, W.T. & Strauss, J.S. (1979). Diagnostic issues in schizophrenia. In L. Bellak (Ed.), *Disorders of the Schizophrenic Syndrome*, (pp. 291-319). New York: Basic Books, Inc. Publishers.
- Carpenter, W.T.j. (1986). The medical approach to diagnosis and treatment of the schizophrenias. In D.B. Feinsilver (Ed.), *Towards a Comprehensive Model for Schizophrenic Disorders: Psychoanalytic Essays in Memory of Ping-Nie Pao, M.D.*, (Vol. 5, pp. 15-28). Hillsdale: The Analytic Press.
- Castel, R. (1973 (1981)). *Le Psychanalisme. L'Ordre Psychanalytique et le Pouvoir*. (3 ed.). Paris: Flammarion.
- Casti, J. & Karlqvist, A. (Eds.). (1989). *Newton to Aristotle: Toward a Theory of Models for Living Systems*. (Vol. 4). Boston, MA: Birkhäuser.
- Castoriadis-Aulagnier, P. (1975 (1981)). *La Violence de l'Interprétation: due Pictogramme à l'énoncé*. (2 ed.). Paris: Presses Universitaires de France.
- Cattell, R.B. (1946). *Description and Measurement of Personality*. Yonkers, N.Y.: World.
- Cattell, R.B. (1950). *Factor Analysis*. New York: Harper.
- Chadwick, P.D.J. & Birchwood, M. (1994). The omnipotence of voices: A cognitive approach to auditory hallucinations. *British Journal of Psychiatry*, 164, 190-201.
- Chadwick, P.D.J. & Lowe, C.F. (1994). A cognitive approach to measuring and modifying delusions. *Behavioral Research and Therapy*, 32(3), 355-367.
- Chadwick, P.D.J., Lowe, C.F., Horne, P.J. & Higson, P.J. (1994). Modifying delusions: The role of empirical testing. *Behavior Therapy*, 25, 35-49.
- Chapman, L.J. & Chapman, J.P. (1980). Scales for rating psychotic and psychotic-like experiences as continua. *Schizophrenia Bulletin*, 6, 476-489.
- Chapman, L.J., Chapman, J.P., Numbers, J.A., Edell, W.S., Carpenter, B.N. & Beckfield, D. (1984). Impulsive non-conformity as a trait contributing to the prediction of psychoticlike and schizotypal symptoms. *Journal of Nervous and Mental Disease*, 172, 681-691.
- Chapman, L.J., Chapman, J.P. & Raulin, M.L. (1976). Scales for physical and social Anhedonia. *Journal of Abnormal Psychology*, 85, 374-382.
- Chapman, L.J., Chapman, J.P. & Raulin, M.L. (1978). Body-image aberration in schizophrenia. *Journal of Abnormal Psychology*, 87, 399-407.
- Chapple, E.D. (1970). *Culture and the Biological Man: Explanations in Behavioral Anthropology*. New York: Holt, Rinehart & Wilson.
- Chow, S.L. (1987). Science, Ecological Validity and Experimentation. *Journal for the Theory of Social Behaviour*, 17, 181-194.
- Chow, S.L. (1992). Positivism and cognitive psychology: A second look. In C.W. Tolman (Ed.), *Positivism in Psychology: Historical and Contemporary Problems*, (pp. 119-144). New York: Springer-Verlag.
- Christison, G.W., Kirch, D.G. & Wyatt, R.J. (1991). When symptoms persist: Choosing among alternative somatic treatments of schizophrenia. *Schizophrenia Bulletin*, 17(2), 217-246.
- Ciminero, A.R., Calhoun, K.S. & Adams, H.E. (Eds.). (1986). *Handbook of Behavioral Assessment*. New York: John Wiley-Interscience.
- Ciampi, L. (1982). *Affektlogik: über die Struktur der Psyche und ihre Entwicklung. Ein Beitrag zur Schizophrenieforschung*. (1 ed.). Stuttgart: Klett-Cotta.
- Ciampi, L. (1989). The dynamics of complex biological-psychosocial systems. Four fundamental psycho-biological mediators in the long-term evolution of schizophrenia. *British Journal of Psychiatry*, 155(suppl. 5), 15-21.

- Ciampi, L. & Müller, C. (1976). *Lebensweg, und Alter der Schizophrenen. Eine katamnestiche Langzeitstudie bis ins Senium*. Berlin: Springer-Verlag.
- Clardy, J.A., Hyde, T.M. & Kleinman, J.E. (1994). Postmortem neurochemical and neuropathological studies in schizophrenia. In N.C. Andreasen (Ed.), *Schizophrenia: From Mind to Molecule*, (pp. 123-146). Washington: American Psychiatric Press, Inc.
- Cloninger, C.R., Martin, R.L., Guze, S.B. & Clayton, P.J. (1985). Current status of schizophrenia as a disease concept. In M. Alpert (Ed.), *Controversies in Schizophrenia: Changes and Constancies*, (Vol. XLI, pp. 12-24). New York: The Guilford Press.
- Cloninger, C.R., Martin, R.L., Guze, S.B. & Clayton, P.J. (1985). Diagnosis and prognosis in schizophrenia. *Archives of General Psychiatry*, 42, 15-25.
- Cloninger, C.R., Martin, R.L., Guze, S.B. & Clayton, P.J. (1990). The empirical structure of psychiatric comorbidity and its theoretical significance. In J.D. Maser & C.R. Cloninger (Eds.), *Comorbidity of Mood and Anxiety Disorders*, (pp. 439-462). Washington: American Psychiatric Press.
- Cohen, S. & Hoberman, H.M. (1983). Positive events and social supports as buffers of life change stress. *Journal of Applied Social Psychology*, 13, 99-125.
- Corrigan, P.W., Wallace, C.J. & Schade, M.L.G., M.F. (1993). Learning medication self-management skills in schizophrenia: Relationships with cognitive deficits and psychiatric symptoms. *Behavior Therapy*.
- Costello, C.G. (1993). The advantages of the symptom approach to depression. In C.G. Costello (Ed.), *Symptoms of Depression*, (pp. 1-22). New York: John Wiley & Sons, Inc.
- Costello, C.G. (1993). Advantages of the symptom approach to schizophrenia. In C.G. Costello (Ed.), *Symptoms of Schizophrenia*, (pp. 1-26). New York: John Wiley & Sons, Inc.
- Costello, C.G. (Ed.). (1993). *Basic Issues in Psychopathology*. New York: The Guilford Press.
- Costello, C.G. (1993). Cognitive causes of psychopathology. In C.G. Costello (Ed.), *Basic Issues in Psychopathology*, (pp. 320-358). New York: The Guilford Press.
- Costello, C.G. (1993). From symptoms of depression to syndromes of depression. In C.G. Costello (Ed.), *Symptoms of Depression*, (pp. 291-302). New York: John Wiley & Sons, Inc.
- Costello, C.G. (1993). From symptoms of schizophrenia to syndromes of schizophrenia. In C.G. Costello (Ed.), *Symptoms of Schizophrenia*, (pp. 291-304). New York: John Wiley & Sons, Inc.
- Costello, C.G. (1993). Introduction: Conceptual and methodological issues in the sciences of psychopathology. In C.G. Costello (Ed.), *Basic Issues in Psychopathology*, (pp. 1-18). New York: The Guilford Press.
- Costello, C.G. (1993). Psychopathology in the clinic and in the community. In C.G. Costello (Ed.), *Basic Issues in Psychopathology*, (pp. 67-102). New York: The Guilford Press.
- Costello, C.G. (Ed.). (1993). *Symptoms of Depression*. New York: John Wiley & Sons, Inc.
- Costello, C.G. (Ed.). (1993). *Symptoms of Schizophrenia*. New York: John Wiley & Sons, Inc.
- Coveney, P. & Highfield, R. (1990 (1991)). *The Arrow of Time: the Quest to Solve Science's Greatest Mystery*. London: Flamingo, Harper Collins Publishers.
- Cromwell, R.L. (1984 (1983)). Preemptive thinking and schizophrenia research. In W.D. Spaulding & J.K. Cole (Eds.), *Theories of Schizophrenia and Psychosis. Nebraska Symposium on Motivation*, 1983, (pp. 1-46). Lincoln: University of Nebraska Press.
- Cronbach, L.J. (1990). *Essentials of Psychological Testing*. (5 ed.). New York: Harper & Row.
- Crow, T.J. & Stevens, M. (1978). Age disorientation in chronic schizophrenia. *British Journal of Psychiatry*, 133, 137-143.
- Crow, T.J. (1980). Molecular pathology of schizophrenia. More than one disease process? *British Medical Journal*, 280, 66-68.
- Crow, T.J. (1980). Positive and negative symptoms of schizophrenia and the role of dopamine. *British Journal of Psychiatry*, 137, 383-386.
- Crow, T.J., Cross, A.J., Johnson, J.A., Johnstone, E.C., Joseph, M.H., Owen, F., Owens, D.G.C. & Poulter, M. (1984). Catecholamines and schizophrenia: An assessment of the evidence. In E. Usden, A. Carlsson, A. Dahlstrom & J. Engel (Eds.), *Catecholamines. Part C: Neuropharmacology and Central Nervous System - Therapeutic Aspects*, (pp. 11-20). New York: Liss.
- Crow, T.J., Done, D.J. & Sacker, A. (1994). Neurological abnormality in children who develop psychosis in adulthood. *Schizophrenia Research*, 11(2), 96.
- Crowe, R.R. (1994). Molecular genetic research in schizophrenia. In N.C. Andreasen (Ed.), *Schizophrenia: From Mind to Molecule*, (pp. 245-260). Washington: American Psychiatric Press.
- Csikszentmihalyi, M. (1990). *Flow: the Psychology of Optimal Experience*. New York: Harper Perennial.
- Csikszentmihalyi, M. (1991). *Flow. The Psychology of Optimal Experience. Steps Toward Enhancing the Quality of Life*. New York: Harper Perennial.
- Csikszentmihalyi, M. (1993). *The Evolving Self: a Psychology for the Third Millennium*. New York: Harper Collins.

- Csikszentmihalyi, M. & Larson, R. (1984). *Being Adolescent: Conflict and Growth in the Teenage Years*. New York: Basic Books.
- Csikszentmihalyi, M. & Larson, R. (1987). Validity and reliability of the Experience Sampling Method. *Journal of Nervous and Mental Disease*, 175(9), 526-536.
- Csikszentmihalyi, M., Larson, R. & Prescott, S. (1977). The Ecology of Adolescent Activity and Experience. *Journal of Youth and Adolescence*, 6, 281-294.
- Csikszentmihalyi, M., Rathune, K. & Whalen, S. (1993). *Talented Teenagers. The Roots of Success & Failure*. Cambridge: Cambridge University Press.
- Csikszentmihalyi, M. & Selega-Csikszentmihalyi, I. (Eds.). (1988). *Optimal Experience: Psychological Studies of Flow in Consciousness*. Cambridge: Cambridge University Press.
- Daan, S. & Gwinner, E. (Eds.). (1989). *Biological Clocks and Environmental Time. Proceedings of a Symposium in Honor of Prof. Dr. Jürgen Aschoff*. New York: Guilford Press.
- Damasio, A.R. (1994). *Descartes' Error. Emotion, Reason, and the Human Brain*. New York: G.P.Putman's Sons.
- David, A.S. (1994). The neuropsychological origin of auditory hallucinations. In A.S. David & J.C. Cutting (Eds.), *The Neuropsychology of Schizophrenia*, (pp. 269-316). Hillsdale: Lawrence Erlbaum Associates.
- David, A.S. & Cutting, J.C. (Eds.). (1994). *The Neuropsychology of Schizophrenia*. Hillsdale: Lawrence Erlbaum Associates.
- Davidhizar, R.E. (1985). Can clients with schizophrenia describe feelings and beliefs about taking medication? *Journal of Advanced Nursing*, 10, 469-473.
- De Boeck, P. (1988). *Psychology van het Individu*. Leuven: Universitaire Pers Leuven.
- Debus, S. & Machleidt, W. (1990). EEG-Dynamik bei Aggression, Trauer und regressiver Trauer. *Zentralblatt Neurologie Psychiatrie*, 255(3-4), 212.
- Delepaul, P.A.E.G. (1981). *Indicatiestelling voor Psychotherapie*. Unpublished Departement voor Psychologie en Pedagogische Wetenschappen, Katholieke Universiteit Leuven.
- Delepaul, P.A.E.G. (1982). *Literatuuronderzoek Extramuralisatie* : P.M.S. Vijverdal, Maastricht.
- Delepaul, P.A.E.G. (1989). Een token economy systeem op een klinische afdeling voor residentiële verblijvende schizofrenen. *Nieuwsbrief Intramuraal VGT*(2), 10.
- Delepaul, P.A.E.G. (1992,). *The Experience Sampling Method: Questions and Methods (Workshop)*. Paper presented at "Invited Conference on Field Studies", Baltimore (US) March 3 and 4, 1992.
- Delepaul, P.A.E.G. (1992). Technical note: Devices and time-sampling procedures. In M.W. deVries (Ed.), *The Experience of Psychopathology: Investigating Mental Disorders in their Natural Settings*, (pp. 363-376). Cambridge: Cambridge University Press.
- Delepaul, P.A.E.G. (1995). *Assessing Schizophrenia in Daily Life. The Experience Sampling Method*. Rijksuniversiteit van Limburg (Maastricht, The Netherlands).
- Delepaul, P.A.E.G. & deVries, M.W. (1987). The daily life of ambulatory chronic mental patients. *Journal of Nervous & Mental Disease*, 175(9), 537-544.
- Delepaul, P.A.E.G. & deVries, M.W. (1988,). *Experience Sampling Method: A Useful Tool in Behavioral Assessment*. Paper presented at "International Behavior Therapy Congress, September 1988", Edinburgh.
- Delepaul, P.A.E.G. & deVries, M.W. (1991). Gedragsdiagnostiek met behulp van de Experience Sampling Methode. *T. Klin. Psych.*, 21(1), 45-59.
- Delepaul, P.A.E.G. & deVries, M.W. (1991,). *Stemmen Horen (Lecture)*. Paper presented at "Stemmen Horen", Maastricht: 1 februari 1991.
- Delepaul, P.A.E.G. & deVries, M.W. (1991,). *Verbal Hallucinations in Daily Life (Lecture)*. Paper presented at "Schizophrenia and Affective Disorders", Genève, Switzerland, 12-14 september 1991.
- Delepaul, P.A.E.G. & deVries, M.W. (1994,). *The Experience and Social Context of Auditory Hallucinations (Poster)*. Paper presented at "VIIth Biennial European Winter Workshop on Schizophrenia", Les Diablerets, Switzerland, January 23-28, 1994.
- Delepaul, P.A.E.G. & deVries, M.W. (1994). The experience and social context of auditive hallucinations. *Schizophrenia Research*, 11(2), 187.
- Delepaul, P.A.E.G., deVries, M.W. & Radstake, D.W.S. (1995). Persoonlijke rehabilitatie: Een 'Experience Sampling' casus. In G. Pieters & J. Peuskens (Eds.), *Rehabilitatie van de Chronische Psychiatrische Patiënt. Op weg naar een Gemeenschapspsychiatrie*, (pp. 215-234). Leuven: Grant.
- Delepaul, P.A.E.G. & Kerpentier, J.M.F.M. (1990). *Het dagelijkse leven van chronische psychiatrische patiënten. Wat weten hulpverleners ervan?* . Maastricht: IPSE.
- Delepaul, P.A.E.G. & Reis, H. (1995 (Chapter 6)). Comparing Time and Event Sampling.
- Delepaul, P.A.E.G. & Stone, A.A. (1993). *Psychometric Considerations of Sampling Moments* (Unpublished Manuscript) : Rijksuniversiteit Limburg.
- Deleuze, G. (1985 (1989)). *Cinéma 2: L'Image Temps (The Time Image)*. London: The Athlone Press.

- DeLisi, L. & Hoff, A. (Eds.). (1990). *Genetic Marker Studies in Schizophrenia and Affective Disorder: Is there a Biological Evidence for a Continuum of the Clinical Syndromes?* Washington DC: American Psychiatric Press.
- DeLisi, L.E. (Ed.). (1990). *Depression in Schizophrenia*. Washington: American Psychiatric Press, Inc.
- Delle Fave, A. & Massimini, F. (1988). Modernization and the changing contexts of Flow in work and leisure. In M. Csikszentmihalyi & I. Selega-Csikszentmihalyi (Eds.), *Optimal Experience: Psychological Studies of Flow in Consciousness*, (pp. 183-213). Cambridge: Cambridge University Press.
- Delle Fave, A. & Massimini, F. (1992). The ESM and the measurement of clinical change: a case of anxiety disorder. In M.W. deVries (Ed.), *The Experience of Psychopathology: Investigating Mental Disorders in their Natural Settings*, (pp. 280-289). Cambridge: Cambridge University Press.
- Delle Fave, A., Massimini, F. & Borri Gaspardin, M. (1992). Televisione e qualità dell'esperienza soggettiva: l'integrazione tra dati quantitativi e qualitativi. *Ikon*, 24.
- Depue, R.A. & Zald, D.H. (1993). Biological and environmental processes in nonpsychotic psychopathology: A neurobehavioral perspective. In C.G. Costello (Ed.), *Basic Issues in Psychopathology*, (pp. 127-237). New York: The Guilford Press.
- deVries, M.W. (1983). *Temporal Patterning of Psychiatric Symptoms (Abstract)*. Paper presented at "World Psychiatric Association Congress", Vienna.
- deVries, M.W. (1987). Mental disorders in their natural settings. The application of time allocation and Experience Sampling techniques in psychiatry. *Journal of Nervous and Mental Disease*, 175(9), 509-578.
- deVries, M.W. (1992). The experience of psychopathology in natural settings: Introduction and illustration of variables. In M.W. deVries (Ed.), *The Experience of Psychopathology: Investigating Mental Disorders in their Natural Settings*, (pp. 3-26). Cambridge: Cambridge University Press.
- deVries, M.W. (Ed.). (1992). *The Experience of Psychopathology: Investigating Mental Disorders in their Natural Settings*. Cambridge: Cambridge University Press.
- deVries, M.W. (1992). Practical issues in psychiatric applications of ESM. In M.W. deVries (Ed.), *The Experience of Psychopathology: Investigating Mental Disorders in their Natural Settings*, (pp. 317-323). Cambridge: Cambridge University Press.
- deVries, M.W. (1992). The uses of the ESM in psychotherapy. In M.W. deVries (Ed.), *The Experience of Psychopathology: Investigating Mental Disorders in their Natural Settings*, (pp. 255-259). Cambridge: Cambridge University Press.
- deVries, M.W. (1994). Naturalistic assessments improve clinical reasoning. In L.F. Saugstad, H.L. Freeman & M.K. Palomäki (Eds.), *Psychiatry and Mental Health. The Legacy of Ørnulf Ødegård*, (pp. 137-166). Oslo: Letten F. Saugstad Fund.
- deVries, M.W. & Delespaul, P.A.E.G. (1988). *Persons in the Schizophrenics' Environment: Influence on Mental State*. Maastricht: Department of Social Psychiatry. R.U. Limburg, Maastricht.
- deVries, M.W. & Delespaul, P.A.E.G. (1989,). *The Experience of Psychosis in Daily Life*. Paper presented at the 30th Dutch Federation Meeting, San Francisco, California.
- deVries, M.W. & Delespaul, P.A.E.G. (1989). Time, context and subjective experience in Schizophrenia. *Schizophrenia Bulletin*, 15(2), 233-244.
- deVries, M.W. & Delespaul, P.A.E.G. (1992). Variability of schizophrenia symptoms. In M.W. deVries (Ed.), *The Experience of Psychopathology: Investigating Mental Disorders in their Natural Settings*, (pp. 97-109). Cambridge: Cambridge University Press.
- deVries, M.W. & Delespaul, P.A.E.G. (1993). Zelfobservatie en intensieve tijdsampling in de psychiatrie. In C.A.L. Hoogduin, P. Schnabel, W. Vandereycken, K. van der Velden & F.C. Verhulst (Eds.), *Jaarboek Psychiatrie en Psychotherapie*, (Vol. 4, pp. 94-108). Meppel: Van Loghum Slaterus.
- deVries, M.W., Delespaul, P.A.E.G., Dijkman, C.I.M. & Theunissen, J. (1986). I disordini mentali gravi: aspetti temporali ed ambientali. In F. Massimini & P. Inghilleri (Eds.), *L'esperienza quotidiana. Teoria e metodo d'analisi*. (pp. 477-491). Milano: Franco Angeli.
- deVries, M.W., Delespaul, P.A.E.G. & Dijkman-Caes, C.I.M. (1987). Affect and anxiety in daily life. In G. Racagni & E. Smeraldi (Eds.), *Anxious Depression: Assessment and Treatment*, (pp. 21-32). New York: Raven Press.
- deVries, M.W., Delespaul, P.A.E.G., Nicolson, N. & Bayer, H. (1993). *Basic Properties of Illness and Well-Being. Phase I: Variable Description for the study of Quality of Life and Clinical Applications. Interim Report to Solvay-Duphar* : IPSE Foundation.
- deVries, M.W., Delespaul, P.A.E.G. & Theunissen, J. (1984,). *Diurnal variation in the conscious experience of schizophrenics (Abstract)*. Paper presented at "Symposium on Development of Schizophrenia. WPA Congress", Helsinki.
- deVries, M.W., Dijkman-Caes, C.I.M. & Delespaul, P.A.E.G. (1988). De ontbrekende schakel: diagnostiek in de natuurlijke omgeving. *Tijdschrift voor Psychiatrie*, 30(2), 94-113.

- deVries, M.W., Dijkman-Caes, C.I.M. & Delespaul, P.A.E.G. (1990). The sampling of experience: A method of co-measuring the co-occurrence of anxiety and depression in daily life. In J.D. Maser & C.R. Cloninger (Eds.), *Comorbidity of Mood and Anxiety Disorders*, (pp. 869). Washington DC: American Psychiatric Press.
- Dickinson, A. (1980). *Contemporary Animal Learning Theory*. (Vol. 1). Cambridge: Cambridge University Press.
- Diener, E. & Emmons, R.A. (1984). The independence of positive and negative affect. *Journal of Personality and Social Psychology*, 47, 1105-1117.
- Diener, E., Emmons, R.A., Larsen, R.J. & Griffin, S. (1985). The "Satisfaction with Life" scale. *Journal of Personality Assessment*, 49, 71-75.
- Dijkman-Caes, C. & deVries, M.W. (1991). Daily life situations and anxiety in panic disorder and agoraphobia. *Journal of Anxiety Disorders*, 5, 343-357.
- Dijkman-Caes, C.I.M. (1987). *Resultaten Vergelijking ESM en Dagboeken (n=10)* (Internal Publication): Rijksuniversiteit Limburg (Maastricht, NL).
- Dijkman-Caes, C.I.M. (1993). *Panic Disorder and Agoraphobia in Daily Life*. Unpublished Proefschrift, Maastricht.
- Dijkman-Caes, C.I.M. & deVries, M.W. (1987). The social ecology of anxiety. Theoretical and quantitative perspectives. *Journal of Nervous and Mental Disease*, 175(9), 550-557.
- Dijkman-Caes, C.I.M., deVries, M.W., Kraan, H.F. & Volovics, A. (1993). Agoraphobic behavior in daily life: Effects of social roles and demographic characteristics. *Psychological Reports*, 72, 1283-1293.
- Dingemans, P.M., De Winter, M.L., Bleeker, A.C. & Rathod, P. (1983). A cross-cultural study of the reliability and factorial dimensions of the Brief Psychiatric Rating Scale (BPRS). *Psychopharmacology*, 80(2), 190-191.
- Dingemans, P.M., Linszen, D.H., Lenior, M.E. & Smeets, R.M.W. (1994). Factor structure of the extended BPRS (BPRS-E). [in preparation].
- Docherty, J.P., van Kammen, D.P., Siris, S.G. & Marder, S.R. (1978). Stages of onset of schizophrenic psychosis. *American Journal of Psychiatry*, 135, 420-426.
- Dollard, J. & Miller, N.E. (1950). *Personality and Psychotherapy. An Analysis in Terms of Learning, Thinking and Culture*. New York: McGraw Hill.
- Donner, E. (1992). Expanding the experiential parameters of cognitive therapy. In M.W. deVries (Ed.), *The Experience of Psychopathology*, (pp. 260-269). Cambridge: Cambridge University Press.
- Drury, V. (1992). Monitoring recovery from acute psychosis. In M. Birchwood & N. Tarrier (Eds.), *Innovations in the Psychological Management of Schizophrenia: Assessment, Treatment and Services*, (pp. 49-76). Chichester: John Wiley & Sons.
- Drury, V. (1994). Recovery from acute psychosis. In M. Birchwood & N. Tarrier (Eds.), *Psychological Management of Schizophrenia*, (pp. 23-51). Chichester: John Wiley & Sons.
- Eccles, J.C. (1994). *How the Self Controls its Brain*. Berlin: Springer Verlag.
- Eckman, T.A., Liberman, R.P., Phipps, C.C. & Blair, K. (1990). Teaching medication self-management skills to schizophrenics. *Journal of Clinical Psychopharmacology*, 10(1), 33-38.
- Edelman, G.M. (1987). *Neural Darwinism*. New York: Basic Books.
- Edelman, G.M. (1988). *Topobiology*. New York: Basic Books.
- Edelman, G.M. (1989). *The remembered present*. New York: Basic Books.
- Edelman, G.M. (1992). *Bright Air, Brilliant Fire*. New York: Basic Books.
- Edlund, M. (1987). *Psychological Time and Mental Illness*. New York: Gardner Press.
- Ehlers, C.L., Frank, E. & Kupfer, D.J. (1988). Social Zeitgebers and biological rhythms. A unified approach to understanding the etiology of depression. *Arch Gen Psychiatry*, 45(Oct), 948-952.
- Ekman, P., Friesen, W.V. & Ellsworth, P.C. (1972). *Emotions in the Human Face*. New York: Pergamon Press.
- Elias, N. (1987). *Über die Zeit*. : Suhrkamp Verlag.
- Elkins, I.J. & Cromwell, R.L. (1991). A focalized deficit within an elegant system. *Behavioral and Brain Sciences*, 14, 27-28.
- Emmons, R.A. & King, L.A. (1992). Thematic analysis, Experience Sampling and personal goals. In C.P. Smith, J.W. Atkinson, D.C. McClelland & J. Veroff (Eds.), *Motivation and Personality: Handbook of Thematic Content Analysis*, (pp. 73-86). New York: Cambridge University Press.
- Epstein, S. (1979). The Stability of Behavior. I. On predicting most of the people some of the time. *Journal of Personality and Social Psychology*, 37, 1097-1126.
- Epstein, S. (1980). The stability of behavior. II. Implications for psychological research. *American Psychologist*, 35, 790-806.
- Ericsson, A.K. & Simon, H.A. (1984). *Protocol Analysis: Verbal Reports as Data*. Cambridge, MA: MIT Press.

- Erikson, R. & Binder, L. (1986). Cognitive deficits among functionally psychotic patients: a rehabilitation perspective. *Journal of Clinical and Experimental Neuropsychology*, 8, 257-274.
- Evans, I.M. (1985). Building systems models as a strategy for target behavioral selection in clinical assessment. *Beh. Ass.*, 7, 21-33.
- Evans, I.M. (1986). Response structure and the Triple-Response-Mode concept. In R.O. Nelson & S.C. Hayes (Eds.), *Conceptual Foundations of Behavioral Assessment*, (pp. 131-155). New York: The Guilford Press.
- Ey, H. (1973). *Traité des Hallucinations, Tome I.* (Vol. 1). Paris: Masson et Cie.
- Ey, H. (1973). *Traité des Hallucinations, Tome II.* (Vol. 2). Paris: Masson et Cie.
- Eysenck, H.J. (1947). *Dimensions of Personality*. London: Kegan Paul.
- Eysenck, H.J. (Ed.). (1960). *Handbook of Abnormal Psychology: an Experimental Approach*. London: Pitman Medical Publishing Co., Ltd.
- Eysenck, H.J. (1987). Behavior Therapy. In H.J. Eysenck & I. Martin (Eds.), *Theoretical Foundations of Behavior Therapy*, (pp. 3-36). New York: Plenum Press.
- Eysenck, H.J. & Eysenck, S.B.G. (1976). *Psychoticism as a dimension of personality*. London: Hodder & Stoughton.
- Eysenck, H.J. & Martin, I. (Eds.). (1987). *Theoretical Foundations of Behavior Therapy*. New York: Plenum Press.
- Fairbairn, W.R.D. (1952). *An Object-Relations Theory of the Personality*. New York: Basic Books.
- Falloon, I.H., Boyd, J.L. & McGill, C.W. (1984). *Family Care of Schizophrenia*. New York: The Guilford Press.
- Falloon, I.H., Laporta, M., Fadden, G. & Graham-Hole, V. (1993). *Managing Stress in Families: Cognitive and Behavioural Strategies for Enhancing Coping Skills*. London: Routledge.
- Falloon, I.R. & Fadden, G. (1993). *Integrated Mental Health Care: a Comprehensive Community-based Approach*. Cambridge: University Press.
- Falloon, I.R.H. (Ed.). (1985). *Family Management of Schizophrenia: a study of Clinical, Social, Family, and Economic Benefits*. Baltimore: The John Hopkins University Press.
- Falloon, I.R.H. (Ed.). (1988). *Handbook of Behavioural Family Therapy*. New York: The Guilford Press.
- Federn, P. (1952). *Ego Psychology and the Psychoses*. New York: Basic Books.
- Feighner, J.P., Robin, J.E., Gruze, S.B., Woodruff, R.A., Winokur, G. & Munoz, R. (1972). Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry*, 26(1), 57-62.
- Feinsilver, D.B. (1986). Continuing towards a comprehensive model for schizophrenic disorders: Conclusions and implications. In D.B. Feinsilver (Ed.), *Towards a Comprehensive Model for Schizophrenic Disorders: Psychoanalytic Essays in Memory of Ping-Nie Pao, M.D.* (Vol. 5, pp. 347-380). Hillsdale: The Analytic Press.
- Feinsilver, D.B. (1986). Pao's Telescopic Overview of Treatment. In D.B. Feinsilver (Ed.), *Towards a Comprehensive Model for Schizophrenic Disorders: Psychoanalytic Essays in Memory of Ping-Nie Pao, M.D.* (Vol. 5, pp. 237-258). Hillsdale: The Analytic Press.
- Feldman, J.M. & Lynch, J.G.j. (1988). Self-generated validity and other effects of measurement on belief, attitude, intention and behavior. *Journal of Applied Psychology*, 73(3), 421-435.
- Fenichel, O. (1945). *The Psychoanalytic Theory of Neurosis*. New York: W.W. Norton & Company.
- Fenton, W.S. & McGlashan, T.H. (1991). Natural history of schizophrenic subtypes: I. Longitudinal study of paranoid, hebephrenic and undifferentiated schizophrenia. *Archives of General Psychiatry*, 48, 969-977.
- Fenton, W.S. & McGlashan, T.H. (1991). Natural history of schizophrenic subtypes: II. Positive and negative symptoms and long-term course. *Archives of General Psychiatry*, 48, 978-986.
- Ferguson, G.A. (1959). *Statistical Analysis in Psychology & Education*. Tokyo: McGraw-Hill Kogakusha Ltd.
- Ferrero, F.P., Haynal, A.E. & Sartorius, N. (Eds.). (1992). *Schizophrenia and Affective Psychoses, Nosology in Contemporary Psychiatry*. Rome: John Libbey CIC srl.
- Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Stanford: Stanford University Press.
- Figurski, T.J. (1992). Everyday self-awareness: Implications for self-esteem, depression and resistance to therapy. In M.W. deVries (Ed.), *The Experience of Psychopathology*, (pp. 304-313). Cambridge: Cambridge University Press.
- Filstead, W.J. (1988). Monitoring the process of recovery using electronic pagers as a treatment intervention. In M. Galanter (Ed.), *Recent Developments in Alcoholism*, (Vol. 6, pp. 181-191). New York: Plenum Press.
- Fink, M. (1985). *Convulsive therapy: Theory and practice*. New York: Raven Press.
- Fink, W.L. (1986). Microcomputers and phylogenetic analysis. *Science*, 734, 1133-1139.
- Fishbein, M. & Ajzen, I. (1975). *Belief, Attitude, Intention and Behavior*. Reading, MA: Addison-Wesley.

- Fisher, S. & Greenberg, R.P. (Eds.). (1978). *The Scientific Evaluation of Freud's Theories and Therapies*. New York: Basic Books.
- Fiske, (1971). *Measuring the concept of personality*. New York: Aldine.
- Flaum, M. & Andreasen, N.C. (1991). Diagnostic criteria for schizophrenia and related disorders: Options for DSM-IV. *Schizophrenia Bulletin*, 17(1), 133-142.
- Foa, E. & Kozak, M. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin*, 99(1), 22-35.
- Foa, E.B. & Kozak, M.J. (1985). Treatment of anxiety disorders: Implications for psychopathology. In H.A. Tuma & J. Maser (Eds.), *Anxiety and the Anxiety Disorders*, (pp. 421-452). Hillsdale: Lawrence Erlbaum Associates Publishers.
- Folkard, S. (1976). Diurnal variations and individual differences in the perception of intractable pain. *Journal of Psychosomatic Research*, 20, 289-301.
- Foucault, M. (1972). *Histoire de la folie à l'âge classique*. Paris: Éditions Gallimard.
- Foulds, G. & Bedford, A. (1975). Hierarchy of classes of personal illness. *Psychological Medicine*, 5, 181-192.
- Frances, A.J., Widiger, T.A. & Sabshin, M. (1991). Psychiatric diagnosis and normality. In D. Offer & M. Sabshin (Eds.), *The Diversity of Normal Behavior. Further Contributions to Normatology*, (pp. 3-38). New York: Basic Books.
- Frederiksen, L.W., Epstein, L.H. & Kosevsky, B.P. (1975). Reliability and controlling effects of three procedures for self-monitoring smoking. *The Psychological Record*, 25, 255-264.
- Freud, S. (1911 (1984)). Formulations on the two principles of mental functioning. In A. Richards (Ed.), *On Metapsychology: the Theory of Psychoanalysis: Beyond the Pleasure Principle, the Ego and the ID and other Works*, (Vol. 11, pp. 29-44). Middlesex: Penguin Books.
- Freud, S. (1917 (1984)). A metapsychological supplement to the theory of dreams. In A. Richards (Ed.), *On Metapsychology: the Theory of Psychoanalysis: Beyond the Pleasure Principle, the Ego and the ID and other Works*, (Vol. 11, pp. 222-244). Middlesex: Penguin Books.
- Freud, S. (1923 (1984)). The Ego and the ID. In A. Richards (Ed.), *On Metapsychology: the Theory of Psychoanalysis: Beyond the Pleasure Principle, the Ego and the ID and other Works*, (Vol. 11, pp. 339-340). Middlesex: Penguin Books.
- Freud, S. (1924 (1961)). The loss of reality in neurosis and psychosis. In A. Richards (Ed.), *Standard Edition*, (Vol. 19, pp. 183-187). London: Hogarth Press.
- Freud, S. (1924 (1961)). Neurosis and psychosis. In A. Richards (Ed.), *Standard Edition*, (Vol. 19, pp. 150-152). London: Hogarth Press.
- Frijia, N. (1985). *The Emotions*. Cambridge: Cambridge University Press.
- Frith, C.D. (1992). *The Cognitive Neuropsychology of Schizophrenia*. Hillsdale: Lawrence Erlbaum Associates.
- Frith, C.D. & Done, D.J. (1989). Experiences of Alien Control in Schizophrenia Reflect a Disorder in the Central Monitoring of Action. *Psychological Medicine*, 19, 359-363.
- Frith, U. (Ed.). (1991). *Autism and Asperger Syndrome*. Cambridge: Cambridge University Press.
- Fromm-Reichmann, F. (1950). *Principles of Intensive Psychotherapy*. Chicago: University of Chicago Press.
- Gaebel, W. (1993). The importance of non-biological factors influencing the outcome of clinical trials. *British Journal of Psychiatry*, 163(Supplement 22), 45-50.
- Gibbons, R.D., Hedeker, D., Elkin, I., Waternaux, C., Kraemer, H.C., Greenhouse, J.B., Shea, M.T., Imber, S.D., Sotsky, S.M. & Watkins, J.T. (1993). Some conceptual and statistical issues in analysis of longitudinal psychiatric data: Application to the NIMH treatment of depression collaborative research program dataset. *Archives of General Psychiatry*, 50(9), 739-750.
- Gibbons, R.D., Hedeker, D., Waternaux, C. & Davis, J.M. (1988). Random regression models: A comprehensive approach to the analysis of longitudinal psychiatric data. *Psychopharmacology Bulletin*, 24(438-443).
- Gibson, J.J. (1960). The concept of the stimulus in psychology. *American Journal of Psychology*, 16, 694-703.
- Gibson, J.J. (1967). On the meaning of the term 'stimulus'. *Psychological Review*, 74, 533-534.
- Gibson, J.J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton-Mifflin.
- Glaeser, B. (Ed.). (1989). *Humanökologie*. Opladen: Westdeutscher Verlag.
- Glass, G.V., McGaw, B. & Smith, M.L. (1981). *Meta-Analysis in Social Research*. Beverly Hills, CA: Sage Publications.
- Goffman, E. (1961). *Asylums. Essays on the Social Situation of Mental Patients and Other Inmates*. London: Anchor Books, Doubleday & Co.
- Goldstein, A.P. (1973). *Structured Learning Therapy. Towards a Psychachotherapy for the Poor*. New York: Academic Press.

- Goldstein, A.P., Sprafkin, R.P. & Gershaw, N.J. (1976). *Skill Training for Community Living: Applying Structured Learning Therapy*. New York: Pergamon Press.
- Goldstein, H. (1987). *Multilevel models in educational and social research*. New York: Oxford University Press.
- Goldstein, M.J. (1987). Psychosocial issues. *Schizophrenia Bulletin*, 13(1), 157-172.
- Gottman, J.M. (1981). *Time-Series Analysis. A Comprehensive Introduction for Social Scientists*. (1 ed.). Cambridge: Cambridge University Press.
- Gould, L.N. (1949). Auditory hallucinations and subvocal speech. *Journal of Nervous and Mental Disease*, 109, 418-427.
- Gould, L.N. (1950). Verbal hallucinations as automatic speech. *American Journal of Psychiatry*, 107, 110-119.
- Graef, R. (1979). *Behavioral Consistency: an Analysis through repeated Measures*. Unpublished Dissertation, University of Chicago, Chicago.
- Graef, R., Gianinno, S. & Csikszentmihalyi, M. (1981). Energy consumption in leisure and perceived happiness. In J. Claxton, C.D. Anderson, J.R.B. Ritchie & McDougall (Eds.), *Consumers and Energy Conservation: International Perspectives on Research and Policy Options*, New York: Praeger.
- Graham, J.R. (1987). *The MMPI (2nd Ed.)*. New York: Oxford University Press.
- Graham, J.R. (1990). *The MMPI-2: Assessing Personality and Psychopathology*. New York: Oxford University Press.
- Green, M.F. (1993). Cognitive remediation in schizophrenia. Is it time yet? *American Journal of Psychiatry*, 150(2), 178-187.
- Green, M.F., Satz, P. & Christenson, C. (1994). Minor physical anomalies in schizophrenia patients, bipolar patients and their siblings. *Schizophrenia Bulletin*, 20(3), 433-440.
- Grencavage, L., Bootzin, R.R. & Shoham, V. (1993). Specific and nonspecific effects in psychological treatments. In C.G. Costello (Ed.), *Basic Issues in Psychopathology*, (pp. 359-376). New York: The Guilford Press.
- Grønmo, S. (1982). Use of time and quality of life: Dimensions of welfare related to significant activities. In International Research Group on Time Budgets and Social Activities (Z. Staikov) (Ed.), *It's About Time*, (pp. 97-117). Sofia: The Institute of Sociology at the Bulgarian Academy of Science & The Bulgarian Sociological Association.
- Gross, D.R. (1984). Time allocation. A tool for the study of cultural behavior. *American Review of Anthropology*, 13, 519-558.
- Gruenberg, E.M. (1967). The "Social Breakdown Syndrome" - Some origins. *American Journal of Psychiatry*, 123(12), 1481-....
- Guadagnoli, E. & Mor, V. (1990). Social interaction tests and scales. In B. Spilker (Ed.), *Quality of Life Assessment in Clinical Trials*, New York: Raven Press.
- Gur, R.E. & Pearlson, G.D. (1993). Neuroimaging in schizophrenia research. *Schizophrenia Bulletin*, 19(2), 337-354.
- Guyatt, G.H., Townsend, M., Berman, L.B. & Keller, J.L. (1987). A comparison of Likert and visual analogue scales for measuring change in function. *Journal of Chronic Diseases*, 40(12), 1129-33.
- Habets, (1990).
- Haddock, G., Bentall, R.P. & Slade, P.D. (1993). Psychological treatment of chronic auditory hallucinations: Two case studies. *Behavioural and Cognitive Psychotherapy*, 21, 335-346.
- Hamilton, N.G., Ponzoha, C.A., Cutler, D.L. & Wiegel, R.M. (1989). Social networks and negative versus positive symptoms of schizophrenia. *Neurosurgical Bulletin*, 15(4), 625-633.
- Hammersley, M. & Atkinson, P. (1983 (1993)). *Ethnography: Principles and Practice*. London: Routledge.
- Hand, D.J. & Taylor, C.C. (1987). *Multivariate analysis of variance and repeated measures. A practical approach for behavioral scientists*. London: Chapman & Hall.
- Happé, F. & Frith, U. (1994). Theory of mind in autism. In E. Schopler & G.B. Mesibov (Eds.), *Learning and Cognition in Autism*, New York: Plenum Press.
- Hardy, M.A. (1993). *Regression with dummy variables*. (Vol. 93). Beverly Hills: Sage.
- Harré, R. & Gillett, G. (1994). *The Discursive Mind*. Thousand Oaks: Sage.
- Harré, R. & Secord, P.F. (1972). *The Explanation of Social Behavior*. Oxford, UK: Blackwell.
- Harrow, M. & Quinlan, D.M. (1985). *Disordered Thinking and Schizophrenic Psychopathology*. New York: Gardner Press, Inc.
- Hartmann, H. (1953). Contribution to the development of Ego and Id. In H. Hartmann (Ed.), *Essays on Ego Psychology*, New York: International university Press.
- Hatfield, A.B. (1987). Coping and adaptation: a conceptual framework for understanding families. In A.B. Hatfield & H.P. Lefley (Eds.), *Families of the Mentally Ill: Coping and Adaptation*, (pp. 60-84). New York: The Guilford Press.

- Hatfield, A.B., Spaniol, L. & Zipple, A.M. (1987). "Expressed Emotion": A family perspective. *Schizophrenia Bulletin*, 13(2), 221-226.
- Hawking, S. (1993). *Black Holes and Baby Universes*. New York: Bantam Books.
- Hawking, S.W. (1988). *A Brief History of Time: from the Big Bang to Black Holes*. New York: Bantam Books.
- Hay, D.P., Peterson, G., Endler, N.S., Spiro, H. & Mills, M.J. (1992). The stigma of electroconvulsive therapy: A workshop. In P.J. Fink & A. Tasman (Eds.), *Stigma and Mental Illness*, (pp. 189-202). Washington: American Psychiatric Press, Inc.
- Hayes, S.C. & Cavoar, N. (1977). Multiple tracking and the reactivity of self-monitoring. *Behavior Therapy*, 8, 819-831.
- Haynes, S.N. & Wilson, C.C. (1979). *Behavioral Assessment: Recent Advances in Methods, Concepts, and Applications*. San Francisco: Jossey-Bass Publishers.
- Hedeker, D. (1993). *MIXREG. A Fortran Program for Mixed-Effects Linear Regression Models* : UIC Prevention Research Center: School of Public Health.
- Hedeker, D. & Gibbons, R.D. (1993). *A Random-effects Ordinal Regression Model for Multilevel Analysis* : UIC Prevention Research Center: School of Public Health.
- Hedeker, D., Gibbons, R.D. & Flay, B.R. (1993). *Random-effects Regression Models for Clustered Data: With and Example from Smoking Prevention Research* : UIC Prevention Research Center: School of Public Health.
- Hedeker, D., Gibbons, R.D. & Flay, B.R. (1994). Random-effects regression models for clustered data with an example from smoking prevention research. *Journal of Consulting Clinical Psychology*, 62(4), 757-765.
- Hedges, L.V. & Olkin, I. (1985). *Statistical Methods for Meta-Analysis*. Orlando, FL: Academic Press.
- Heidegger, M. (1962). *Being and Time* (Macquarrie, J. & Robinson, E.S., Trans.). New York: Harper & Row.
- Heilbrun, A.B. (1973). *Aversive maternal control: A theory of schizophrenic development*. New York: Wiley.
- Heilbrun, A.B. (1984). The adaptive style theory of schizophrenic development: Current research with schizophrenics. *Genetic Psychology Monographs*, 110, 229-255.
- Heilbrun, A.B. & Blum, N.A. (1984). Cognitive vulnerability to auditory hallucination: Impaired perception of meaning. *British Journal of Psychiatry*, 144, 508-512.
- Heilbrun, A.B., Blum, N.A. & Haas, M. (1983). Cognitive vulnerability to auditory hallucination: Preferred imagery mode and spatial location of sounds. *British Journal of Psychiatry*, 143, 294-299.
- Heilbrun, A.B.j. (1993). Hallucinations. In C.G. Costello (Ed.), *Symptoms of Schizophrenia*, (pp. 56-91). New York: John Wiley & Sons, Inc.
- Heinrichs, D., Hanlon, T. & Carpenter, W. (1984). The quality of life scale: An instrument for rating the schizophrenic deficit syndrome. *Schizophrenia Bulletin*, 10, 388-398.
- Hersen, M. & Barlow, D.H. (Eds.). (1976). *Single Case Experimental Designs: Strategies for Studying Behavior Change* (1 ed.). (Vol. 56). New York: Pergamon Press.
- Hettema, P.J. (Ed.). (1979). *Personality and Adaptation*. Amsterdam: North-Holland.
- Hettema, P.J. & Kenrick, D.T. (1989). Biosocial interaction and individual adaptation. In P.J. Hettema (Ed.), *Personality and Environment: Assessment of Human Adaptation*, (pp. 3-30). Chichester: John Wiley & Sons.
- Himaldi, B., Osteen, F. & Crawford, E. (1993). Delusional verbalizations and beliefs. *Behavioral Residential Treatment*, 8(4), 229-242.
- Hirsch, S.R. & Leff, J.P. (1975). *Abnormalities in Parents of Schizophrenics*. (Vol. 22). London: Oxford University Press.
- Hogarty, G.E., Goldberg, S.C., Schooler, N.R. & al., e. (1974). Drugs and sociotherapy in the aftercare of schizophrenic subjects. II: Two-year relapse rates. *Archives of General Psychiatry*, 31 (603-608).
- Hollingshead, A.B. & Redlich, F.C. (1958). *Social Class and Mental Illness*. New York: John Wiley & Sons.
- Honig, A. (1989). *Mind the Body*. Unpublished Proefschrift, Maastricht.
- Hormuth, S.E. (1986). The sampling of experience in situ. *Journal of Personality*, 54(1, March 1986), 262-293.
- Hormuth, S.E. (1990). *The Ecology of the Self: Relocation and Self-Concept Change*. Cambridge: Cambridge University Press.
- Howells, J.G. (Ed.). (1991). *The Concept of Schizophrenia: Historical Perspectives*. Washington: American Psychiatric Press, Inc.
- Howells, J.G. (1991). Schizophrenia in the Medieval period. In J.G. Howells (Ed.), *The Concept of Schizophrenia: Historical Perspectives*, (pp. 29-46). Washington DC: American Psychiatric Press.
- Hunter, J.E. (1990). *Methods of Meta-Analysis*. Beverly Hills, CA: Sage.

- Hunter, J.E., Schmidt, F.L. & Jackson, B.G. (1982). *Meta-Analysis: Cumulating Research Findings Across Studies*. Beverly Hills, CA: Sage Publications.
- Hurlburt, R.T. (1980). Validation and correlation of thought sampling with retrospective measures. *Cognitive Therapy Research*, 4, 235-238.
- Hurlburt, R.T. (1990). *Sampling Normal and Schizophrenic Inner Experience*. New York: Plenum Press.
- Hurlburt, R.T. (1993). *Sampling Inner Experience in Disturbed Affect*. New York: Plenum Press.
- Hurlburt, R.T., Happé, F. & Frith, U. (1994). Sampling the form of inner experience in three adults with Asperger syndrome. *Psychological Medicine*, 24, 385-395.
- Hurlburt, R.T. & Melacon, S.M. (1987). P-technique factor analysis of individual thought sampling data. *Cogn Ther Res*, 11(4), 487-500.
- Hurlburt, R.T. & Melancon, S.M. (1987). Goofed-up images: Thought Sampling with a schizophrenic woman. *Journal of Nervous and Mental Disease*, 17(9), 575-578.
- Immink, W. (1986). *Parameter Estimation in Markov Models and Dynamic Factor Analysis*. Unpublished PhD thesis, Rijksuniversiteit Utrecht.
- Izard, C.E. (1972). *Patterns of Emotions: A new analysis of anxiety and depression*. New York: Academic Press.
- Izard, C.E. (1991). *The Psychology of Emotions*. New York: Plenum Press.
- Izard, C.E. (1993). Organizational and motivational functions of discrete emotions. In M. Lewis & J.M. Haviland (Eds.), *Handbook of Emotions*, (pp. 631-642). New York: The Guilford Press.
- Jaccard, J. & Wan, C.K. (1993). Statistical analysis of temporal data with many observations: Issues for behavioral medicine data. *Annals of Behavioral Medicine*, 15(1), 41-50.
- Jaeschke, R., Singer, J. & Guyatt, G.H. (1989). Measurement of health status. Ascertaining the minimal clinically important difference. *Controlled Clin Trials*, 10(4), 407-15.
- Jaquish, G. & Savin-Williams, R. (1981). Biological and ecological factors in the expression of adolescent self-esteem. *Journal of Youth and Adolescence*, 10, 473-485.
- Johnson, C. & Larson, R. (1982). Bulimia: An analysis of moods and behavior. *Psychosomatic Medicine*, 44, 341-351.
- Johnson, F.H. (1978). *The Anatomy of Hallucinations*. Chicago: Nelson-Hall.
- Johnstone, E.C. (1991). Heterogeneity in schizophrenia: Clinical and biological aspects. In J.-P. Lindenmayer & S.R. Kay (Eds.), *New Biological Vistas on Schizophrenia*, (1 ed., Vol. 6, pp. 110-128). New York: Brunner/Mazel.
- Jones, J.M. (1988). Cultural differences in temporal perspectives: Instrumental and expressive behaviors in time. In J.E. McGrath (Ed.), *The Social Psychology of Time: New Perspectives*, (pp. 21-38). Newbury Park: Sage Publications.
- Jones, K. (1972). *A History of Mental Health Services*. London: Routledge & Kegan Paul.
- Jones, P.B., Murray, R.M. & Bodgers, B. (1994). Child development preceding adult schizophrenia in a national birth cohort. *Schizophrenia Research*, 11(2), 97.
- Jöreskog, K.G. (1990). New developments in LISREL: Analysis of ordinal variables using polychoric correlations and weighted least squares. *Quality and Quantity*, 24, 387-404.
- Jöreskog, K.G. & Sörbom, D. (1989). LISREL 7: A guide to the program and applications (Version 2). Chicago: SPSS.
- Judd, L.J. (1994). Introduction. In N.C. Andreasen (Ed.), *Schizophrenia: From Mind to Molecule*, (pp. xiii-xv). Washington: American Psychiatric Press, Inc.
- Kafka, J.S. (1993). The schizophrenic object, insight and consciousness. In G. Benedetti & P.M. Furlan (Eds.), *The Psychotherapy of Schizophrenia: Effective Clinical Approaches - Controversies, Critiques and Recommendations*, (pp. 87-94). Seattle: Hogrefe & Huber Publishers.
- Kahlbaum, K.L. (1863). *Grupierung der Psychische Krankheiten*. Danzig: Kafeman.
- Kane, J., Honigfeld, G., Singer, J. & Meltzer, H. (1988). Clozapine for the treatment resistant schizophrenic: A double-blind comparison versus Chlorpromazine/Benzotropine. *Archive of General Psychiatry*, 45, 789-796.
- Kane, J.M. (1985). Compliance issues in outpatient treatment. *Journal of Clinical Psychopharmacology*, 5, 22s-27s.
- Kane, J.M. (1987). Treatment of schizophrenia. *Schizophrenia Bulletin*, 13, 133-156.
- Kane, J.M. (1991). New developments in the pharmacological treatment of schizophrenia: Editor's introduction. *Schizophrenia Bulletin*, 17(2), 193-196.
- Kane, J.M. (1991). Obstacles to clinical research and new drug development in schizophrenia. *Schizophrenia Bulletin*, 17(2), 353-360.
- Kane, J.M. (1994). Efficacy, mechanisms, and side effects of typical and atypical neuroleptics. In N.C. Andreasen (Ed.), *Schizophrenia: From Mind to Molecule*, (pp. 173-188). Washington: American Psychiatric Press, Inc.

- Kane, J.M. & Marder, S.R. (1993). Psychopharmacologic treatment of schizophrenia. *Schizophrenia Bulletin*, 19(2), 287-302.
- Kanfer, F.H. & Philips, J.S. (1970). *Learning Foundations of Behavior Therapy*. New York: John Wiley & Sons, Inc.
- Kaplan, C.D. (1992). Drug craving and drug use in daily life of heroin addicts. In M.W. deVries (Ed.), *The Experience of Psychopathology: Investigating Mental Disorders in their Natural Settings*, (pp. 193-218). Cambridge: Cambridge University Press.
- Kaplan, C.D., Korf, D. & Sterk, C. (1987). Temporal and social contexts of heroin-using populations: An illustration of the "Snowball Sampling" technique. *Journal of Nervous and Mental Disease*, 175, 566-574.
- Kaplan, H.I. & Sadock, B.J. (1988). *Synopsis of Psychiatry: Behavioral Sciences - Clinical Psychiatry*. (5 ed.). Baltimore: Williams & Wilkins.
- Kay, S.R. (1991). New perspectives on the positive-negative distinction in schizophrenia. In J.-P. Lindenmayer & S.R. Kay (Eds.), *New Biological Vistas on Schizophrenia*, (Vol. 6, pp. 3-27). New York: Brunner/Mazel.
- Kay, S.R. (1991). *Positive and Negative Syndromes in Schizophrenia: Assessment and Research*. (1 ed.). (Vol. 5). New York: Brunner/Mazel.
- Kay, S.R. & Lindenmayer, J.P. (1987). Outcome predictors in acute schizophrenia: Prospective significance of background and clinical dimension. *Journal of Nervous and Mental Disease*, 175, 152-160.
- Kay, S.R. & Murrill, L.M. (1990). Predicting outcome of schizophrenia: Significance of symptom profiles and outcome dimension. *Comprehensive Psychiatry*, 31, 91-102.
- Kay, S.R., Opler, L.A. & Fiszbein, A. (1992 (1986)). *Positive and Negative Syndrome Scale (PANSS +/-)*. Manual. New York: Multi-Health Systems, Inc.
- Kay, S.R. & Sevy, S. (1990). Pyramidal model of schizophrenia. *Schizophrenia Bulletin*, 16(3), 537-545.
- Kazdin, A.E. (1976). *Token Economies*. New York: Plenum Press.
- Kellert, S.H. (1993). *In the Wake of Chaos. Unpredictable Order in Dynamic Systems*. Chicago: University of Chicago Press.
- Kelly, G.A. (1955). *The Psychology of Personal Constructs*. New York: Norton.
- Kelly, J.R. (1988). Entrainment in individual and group behavior. In J.E. McGrath (Ed.), *The Social Psychology of Time: New Perspectives*, (pp. 89-112). Newbury Park: Sage Publications.
- Kelly, J.R. & McGrath, J.E. (1988). *On Time and Method*. (Vol. 13). Newbury Park: Sage.
- Kenardy, J., Evans, L. & Oei, T. (1988). The importance of cognitions in panic attacks. *Behavioral Therapy*, 19(3), 471-483.
- Kenardy, J., Evans, L. & Oei, T. (1989). Cognitions and heart rate in panic disorders during everyday activity. *J Anxiety Disord*, 3(1), 33-43.
- Kenardy, J., Fried, L., Kraemer, H.C. & Taylor, C.B. (1992). Psychological precursors of panic attacks [published erratum appears in British Journal of Psychiatry 1993 Jul;163:130]. *British Journal of Psychiatry*, 160(668), 668-73.
- Kendler, K.S. & Diehl, S.R. (1993). The genetics of schizophrenia: A current, genetic-epidemiologic perspective. *Schizophrenia Bulletin*, 19(2), 261-285.
- Kenny, D.A., Kashy, D.A. & Bolger, N. (1994). *Analysis Strategies for Multi-Level Data: An Overview and Evaluation of Approaches* (Mimeo): University of Connecticut.
- Kern, R.S., Green, M.F. & Satz, P. (1992). Neuropsychological predictors of skills training for chronic mental patients. *Psychiatry Research*, 43, 223-230.
- Kingdon, D.G. & Turkington, D. (1994). *Cognitive - Behavioral Therapy of Schizophrenia*. Hillsdale: Lawrence Erlbaum Associates.
- Kirch, D.G. (1993). Infections and autoimmunity as etiologic factors in schizophrenia: A review and reappraisal. *Schizophrenia Bulletin*, 19(2), 355-370.
- Klein, D. (1982). Relation between current diagnostic criteria for schizophrenia and the dimensions of premorbid adjustment paranoid symptomatology and chronicity. *Journal of Abnormal Psychology*, 91, 319-325.
- Klein, D.N. & Riso, L.P. (1993). Psychiatric disorders: Problems of boundaries and comorbidity. In C.G. Costello (Ed.), *Basic Issues in Psychopathology*, (pp. 19-66). New York: The Guilford Press.
- Klein, M. (1975). *Envy & Gratitude & Other Works 1946 -1963*. New York: Delacorte Press/Seymour Lawrence.
- Kleitman, N. (1963). *Sleep and Wakefulness*. (2 ed.). Chicago: University of Chicago Press.
- Klerman, G.L., Vaillant, G.E., Spitzer, R.L. & Michels, R. (1984). The advantages of DSM-III. *American Journal of Psychiatry*, 141(4), 539-542.
- Klinger, E., Barta, S. & Maxeimer, M. (1980). Motivational correlates of thought content, frequency and commitment. *Journal of Personality and Social Psychology*, 39, 1222-1237.

- Knight, R.A. (1987). Relating cognitive processes to symptoms: A strategy to counter methodological difficulties. In P.D. Harvey & E.E. Walker (Eds.), *Positive and Negative Symptoms in Psychosis: Description, Research and Future Directions.*, (pp. 1-29). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Koele, P. (1984). Enkele opmerkingen betreffende het gebruik van Fishers z-transformatie van product-moment correlatie-coëfficiënten. *Tijdschrift voor Onderwijsresearch*, 9(5), 235-242.
- Kohut, H. (1971). *The Analysis of the Self*. New York: International University Press.
- Kooyman, M. (1993). *The Therapeutic Community for Addicts: Intimacy, Parent Involvement and Treatment Success*. Amsterdam: Swets & Zeitlinger.
- Kosko, B. (1993). *Fuzzy Thinking. The New Science of Fuzzy Logic*. New York: Hyperion.
- Koukkou-Lehmann, M. (1987). *Hirnechanismen Normalen und Schizophrenen Denkens*. Berlin: Springer Verlag.
- Kraan, H., Meertens, H., Hilwig, M., Volovics, L., Dijkman-Caes, C. & Portegijs, P. (1992). Selecting measures, diagnostic validity and scaling in the study of depression. In M.W. deVries (Ed.), *The Experience of Psychopathology: Investigating Mental Disorders in their Natural Settings*, Cambridge: Cambridge University Press.
- Kraemer, H.C. (1992). *Evaluating Medical Tests. Objective and Quantitative Guidelines*. (1 ed.). Beverly Hills, CA: Sage.
- Kraepelin, E. (1896). *Psychiatrie: ein Lehrbuch für Studierende und Ärzte*. Leipzig: J.A.Barth.
- Kratochwill, T.R. (Ed.). (1978). *Single Subject Research: Strategies for Evaluating Change*. New York: Academic Press.
- Kripke, D.F. (1983). Phase advance theories for affective illnesses. In F. Goodwin & T. Wehr (Eds.), *Circadian Rhythms in Psychiatry: Basic and Clinical Studies*, California: Boxwood Press.
- Krueger, R.A. (1988). *Focus Groups*. London: Sage.
- Kruse, L., Graumann, C.F. & Lantermann, E.D. (Eds.). (1990). *Ökologische Psychologie*. München: PVU.
- Kubey, R. (1990). Television and family harmony among children, adolescents and adults: Results from the Experience Sampling Method. In J. Bryant (Ed.), *Television and the American Family*, Hillsdale: Lawrence Erlbaum.
- Kubey, R. & Csikszentmihalyi, M. (1990). *Television and the Quality of Life: How Viewing Shapes Everyday Experience*. (1 ed.). Hillsdale, N.J.: Lawrence Erlbaum Associates, Publishers.
- Kupfer, D.J. & L.L., C. (1990). Clinical evidence of comorbidity: a critique of treated samples and longitudinal studies. In J.D. Maser & C.R. Cloninger (Eds.), *Comorbidity of Mood and Anxiety Disorders*, (pp. 231-238). Washington: American Psychiatric Press.
- Kupfer, D.J., Ritenour, A. & Monk, T. (1990). *The Social Rhythm Metric (SRM)* : University of Pittsburgh School of Medicine - Department of Psychiatry.
- Lacan, J. (1966 (1971)). *Écrits I*. Paris: Éditions du Seuil.
- Lake, L.C. (1993). On neural darwinism. *Social Neuroscience Bulletin*, 6(3), 42-47.
- Lamb, H.R. (Ed.). (1984). *The Homeless Mentally Ill*. Washington DC: APA Press.
- Lamiell, J. (1981). Toward an idiopathic psychology of personality. *American Psychologist*, 36, 276-289.
- Landes, D.S. (1983). *Revolution in Time. Clocks and the Making of the Modern World*. Cambridge: Harvard University Press.
- Lang, P.J. (1979). A bio-informational theory of emotional imagery. *Psychophysiology*, 16, 495-512.
- Lang, P.J. (1985). The cognitive psychophysiology of emotion: Fear and anxiety. In A. Tuma & J.D. Maser (Eds.), *Anxiety and the Anxiety Disorders*, (pp. 131-170). Hillsdale, NJ: Lawrence Erlbaum.
- Larson, R. (1979). *The Significance of Time Alone in Adolescents' Lives*. Unpublished Dissertation, University of Chicago, Chicago.
- Larson, R. (1983). Adolescents' daily experience with family and friends: Contrasting opportunity systems. *Journal of Marriage and the Family*, 45, 739-750.
- Larson, R. (1987). On the independence of positive and negative affect within hour-to-hour experience. *Motivation and Emotion*, 11, 145-156.
- Larson, R. (1989). Beeping children and adolescents: A method for studying time use and daily experience. *Journal of Youth and Adolescence*, 18(6), 511-530.
- Larson, R. (1989). Is feeling "in control", related to happiness in daily life? *Psychological Reports*, 64, 775-784.
- Larson, R. (1990). The solitary side of life: An examination of the time people spent alone from childhood to old age. *Developmental Review*, 10(2), 155-183.
- Larson, R. & Asmussen, L. (1992). Bulimia in daily life: a context-bound syndrome. In M.W. deVries (Ed.), *The Experience of Psychopathology: Investigating Mental Disorders in their Natural Settings*, (pp. 167-179). Cambridge: Cambridge University Press.

- Larson, R., Csikszentmihalyi, M. & Freeman, M. (1984). Alcohol and marijuana use in adolescents' lives: A random sample of experiences. *International Journal of Addictions*, 19, 367-381.
- Larson, R., Csikszentmihalyi, M. & Freeman, M. (1992). Alcohol and marijuana use in adolescents' daily lives. In M.W. deVries (Ed.), *The Experience of Psychopathology: Investigating Mental Disorders in their Natural Settings*, (pp. 180-192). Cambridge: Cambridge University Press.
- Larson, R., Csikszentmihalyi, M. & Graef, R. (1980). Mood variability in the psychological adjustment of adolescents. *Journal of Youth and Adolescence*, 9, 469-490.
- Larson, R., Csikszentmihalyi, M. & Graef, R. (1982). Time alone in daily experience: Loneliness or renewal? In L.A. Peplau & D. Perlman (Eds.), *Loneliness: A Sourcebook of Current Theory, Research and Therapy*, New York: Wiley.
- Larson, R. & Delespaul, P.A.E.G. (1992). Analyzing Experience Sampling data: A guidebook for the perplexed. In M.W. deVries (Ed.), *The Experience of Psychopathology: Investigating Mental Disorders in their Natural Settings*, (pp. 58-78). Cambridge: Cambridge University Press.
- Larson, R. & Johnson, C. (1985). Bulimia: Disturbed patterns of solitude. *Addictive Behaviors*, 10, 281-290.
- Larson, R. & Lampman-Petratis, C. (1989). Daily emotional states as reported by children and adolescents. *Child Development*, 60, 1250-1260.
- Larson, R. & Richards, M. (1989). The changing life space of early adolescence. *Journal of Youth and Adolescence*, 18(6 (special)), 501-626.
- Larson, R. & Richards, M.H. (1994). *Divergent Realities. The Emotional Life of Mothers, Fathers and Adolescents*. New York: Basic Books.
- Laughlin, C., McManus, J. & d'Aquili, E. (1990 (1992)). *Brain, Symbol and Experience. Towards a Neuropsychology of Human Consciousness*. New York: Columbia University Press.
- Launay, G. & Slade, P.D. (1981). The measurement of hallucinatory predisposition in male and female prisoners. *Personality and Individual Differences*, 2, 221-234.
- Lazarus, R. (1991). *Emotions and Adaptation*. New York: Oxford University Press.
- Leff, J. (1990). Depressive symptoms in the course of schizophrenia. In L.E. DeLisi (Ed.), *Depression in Schizophrenia*, (pp. 1-24). Washington: American Psychiatric Press, Inc.
- Leff, J. & Vaughn, C. (1985). *Expressed Emotion in Families: its Significance for Mental Illness*. New York: The Guilford Press.
- Lehman, A.F. (1983). The effects of psychiatric symptoms on quality of life assessments among the chronic mentally ill. *Evaluation and Program Planning*, 6, 143-151.
- Lehman, A.F. (1983). The well-being of chronic mental patients: Assessing their quality of life. *Archives of General Psychiatry*, 40, 369-373.
- Lehman, A.F. (1988). A quality of life interview for the chronically mentally ill. *Evaluation and Program Planning*, 11, 51-62.
- Lehman, A.F. & Burns, B.J. (1990). Severe mental illness in the community. In B. Spilker (Ed.), *Quality of Life Assessments in Clinical Trials*, New York: Raven Press.
- Lehmann, H.E. & Cancro, R. (1985). Schizophrenia: Clinical features. In H.I. Kaplan & B.J. Sadock (Eds.), *Comprehensive Textbook of Psychiatry*, (4 ed., pp. 680-713). Baltimore: Williams & Wilkins.
- Leighton, D.C., Harding, J.S., Macklin, D.B., MacMillan, A.M. & Leighton, A.H. (1963). *The Character of Danger: Psychiatric Symptoms in Selected Communities*. New York: Basic Books.
- Lendrem, D. (1986). *Modelling in Behavioural Ecology: An Introductory Text*. London: Chroom Helm.
- Lewin, K. (1936). *Principles of Topological Psychology*. New York: McGraw-Hill.
- Leysen, J.E., Schotte, A.G., Janssen, P.M.F. & Megens, A.A.H.P. (1994). The Balance of 5HT₂/D₂ Antagonism, a crucial point for the treatment of schizophrenia. *Schizophrenia Research*, 11(2), 114.
- Lieberman, R.P. (Ed.). (1988). *Psychiatric Rehabilitation of Chronic Mental Patients*. Washington: Americal Psychiatric Press, Inc.
- Lieberman, R.P. (1988). *Social and Independent Living Skills: Symptom Management Module - Trainer's Manual*. (1 ed.). Los Angeles: Clinical Research Center for Psychiatric Rehabilitation/Camarillo State Hospital.
- Lieberman, R.P. (1988). Social skills training. In R.P. Lieberman (Ed.), *Psychiatric Rehabilitation of Chronic Mental Patients*, (pp. 147-198). Washington: Americal Psychiatric Press, Inc.
- Lieberman, R.P., DeRisi, W.J. & Mueser, K.T. (1989). *Social Skills Training for Psychiatric Patients*. New York: Pergamon Press.
- Lieberman, R.P. & Eckman, T.A. (1989). Dissimination of skills training modules to psychiatric facilities. Overcoming obstacles to the utilization of a rehabilitation innovation. *British Journal of Psychiatry*, 155(Suppl.5), 117-122.
- Lieberman, R.P., King, L.W., DeRisi, W.J. & McCann, M. (1975). *Personal Effectiveness. Guiding People to Assert themselves and Improve Their Social Skills*. Champaign, Ill.: Research Press Co.

- Lieberman, R.P., Kopelowicz, A. & Young, A.S. (1994). Biobehavioral treatment and rehabilitation of schizophrenia. *Behavior Therapy*, 25, 89-117.
- Lieberman, R.P., Van Putten, T., Marshall, B.D., Mintz, J., Bowen, L., Kuehnel, T.G., Aravagiri, M. & Marder, S.R. (1994). Optimal drug and behavior therapy for treatment-refractory schizophrenic patients. *American Journal of Psychiatry*, 151(5), 756-759.
- Liddle, P.F. (1987). The symptoms of chronic schizophrenia: A reexamination of the positive-negative dichotomy. *British Journal of Psychiatry*, 151, 145-151.
- Liddle, P.F. & Barnes, T.R.E. (1990). Syndromes of chronic schizophrenia. *British Journal of Psychiatry*, 157, 558-561.
- Lieberman, J.A. & Koreen, A.R. (1993). Neurochemistry and neuroendocrinology of schizophrenia. *Schizophrenia Bulletin*, 19(2), 371-430.
- Likert, R.A. (1932). Technique for the measurement of attitudes. *Archive of Psychology*, 140, 1-55.
- Lindenmayer, J.-P. (1992). Pharmacological strategies for the neuroleptic nonresponder. In J.-P. Lindenmayer & S.R. Kay (Eds.), *New Biological Vistas on Schizophrenia*, (1 ed., Vol. 6, pp. 241-258). New York: Brunner/Mazel.
- Lindsley, O.R. (1963). Direct measurement and functional definition of vocal hallucinatory symptoms. *Journal of Nervous and Mental Disorder*, 136, 293-297.
- Loewenstein, R.J., Hamilton, J., Alagna, S., Reid, D. & deVries, M.W. (1987). Experiential Sampling in the study of multiple personality disorder. *American Journal of Psychiatry*, 144, 19-24.
- Lombardo, T.J. (1987). *The Reciprocity of Perceiver and Environment: The Evolution of James J. Gibson's Ecological Psychology*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Lousberg, R., Schmidt, A.J.M., Groenman, N.H., Vendrig, L. & Dijkman-Caes, C.I.M. (1994). Validating the MPI-DLV with Experience Sampling data. In: Lousberg, H.B.R. *Chronic Pain: Multiaxial Assessment and Behavioral Mechanisms*. Unpublished Proefschrift, Maastricht.
- Lukoff, D., Nuechterlein, K.H. & Ventura, J. (1986). Manual for expanded Brief Psychiatric Rating Scale (BPRS). *Schizophrenia Bulletin*, 12(4), 594-602.
- Lurija, A.R. (1965 (1990)). *Een teveel aan geheugen. Een Neuropsychologisch Geval*. Amsterdam: Bert Bakker.
- Machleidt, W. (1992). Typology of functional psychoses: A new model on basic emotions. In F.P. Ferrero, A.E. Haynal & N. Sartorius (Eds.), *Schizophrenia and Affective Psychoses, Nosology in Contemporary Psychiatry*, (pp. 97-104). Rome: John Libbey CIC srl.
- Maher, B.A. (1991). Deception, rational man, and other rocks on the road to a personality psychology of real people. In W.M. Grove & D. Cicchetti (Eds.), *Thinking Clearly about Psychology: Vol.2. Personality and Psychopathology*, (pp. 72-88). Minneapolis: University of Minneapolis Press.
- Malik, S. (1981). *Psychological Modernity: a Comparative Study of African and American Graduate Students in the Midwest*. Unpublished Dissertation, University of Chicago, Chicago.
- Malinowski, B. (1935). *Coral Gardens and their Magic. Soil Tilling and Agricultural Rites*. (Vol. 1). Bloomington: Indiana University Press.
- Mann, J., Ten Have, T., Plunkett, J.W. & Meisels, S.J. (1991). Time sampling: A methodological critique. *Child Development*, 62(1), 227-241.
- Margraf, J. (1990). Ambulatory psychophysiological monitoring of panic attacks. *Journal of Psychophysiology*, 4, 319-330.
- Margraf, J., Taylor, C.B., Ehlers, A., Roth, W.T. & Agras, W.S. (1987). Panic attacks in the natural environment. *Journal of Nervous and Mental Disease*, 175(9), 558-565.
- Maser, J.D. & Cloninger, C.R. (Eds.). (1990). *Comorbidity of Mood and Anxiety Disorders*. Washington: American Psychiatric Press.
- Massimini, F., Csikszentmihalyi, M. & Carli, M. (1987). The monitoring of optimal experience. A tool for psychiatric rehabilitation. *Journal of Nervous and Mental Disease*, 175(9), 545-549.
- Massimini, F., Csikszentmihalyi, M. & Carli, M. (1992). The monitoring of optimal experience: a tool for psychiatric rehabilitation. In M.W. deVries (Ed.), *The Experience of Psychopathology: Investigating Mental Disorders in their Natural Settings*, (pp. 270-279). Cambridge: Cambridge University Press.
- Massimini, F. & Massimo, C. (1988). The systematic assessment of Flow in daily experience. In M. Csikszentmihalyi & I. Selega-Csikszentmihalyi (Eds.), *Optimal Experience: Psychological Studies of Flow in Consciousness*, (pp. 266-287). Cambridge: Cambridge University Press.
- Mayers, P. (1978). *Flow in Adolescence and Its Relation to School Experience*. Unpublished Dissertation, University of Chicago, Chicago.
- McDonald, M.J. & House, D.V. (1992). Positivist influences in "environment-behavior" studies: Egon Brunswik and contemporary social ecology. In C.W. Tolman (Ed.), *Positivism in Psychology: Historical and Contemporary Problems*, (pp. 155-184). New York: Springer-Verlag.
- McFall, R.M. (1977). Parameters of self-monitoring. In R.B. Stuart (Ed.), *Behavioral Self-Management: Strategies, Techniques and Outcomes*, New York: Brunner/Mazel.

- McFarlane, W.R. (Ed.). (1983). *Family Therapy in Schizophrenia*. New York: The Guilford Press.
- McGlashan, T.H. & Fenton, W.S. (1991). Classical subtypes for schizophrenia: Literature Review for DSM-IV. *Schizophrenia Bulletin*, 17(4), 609-632.
- McGrath, J.E. (1988). Introduction: The place of time in social psychology. In J.E. McGrath (Ed.), *The Social Psychology of Time: New Perspectives*, (pp. 7-20). Newbury Park: Sage Publications.
- McGuffin, P., Farmer, A.E. & Gottesman, I.I. (1987). Modern diagnostic criteria and genetic studies of schizophrenia. In H. Häfner, W.F. Gattaz & W. Janzarik (Eds.), *Search for the Causes of Schizophrenia*, (pp. 143-156). New York: Springer Verlag.
- McGuffin, P., Farmer, A.E., Gottesman, I.I., Murray, R.M. & Reveley, A.M. (1984). Twin concordance for operationally defined schizophrenia: Confirmation of familiarity and heritability. *Archives of General Psychiatry*, 41, 541-545.
- McGuire, M.T. & Fairbanks, L.A. (1977). Ethology: Psychiatry's bridge to behavior. In M.T. McGuire & L.A. Fairbanks (Eds.), *Ethological Psychiatry: Psychopathology in the Context of Evolution Biology*, (pp. 1-40). New York: Grune & Stratton.
- McGuire, M.T. & Polsky, R.H. (1979). Behavioral changes in hospitalized acute schizophrenics. An Ethological Perspective. *Journal of Nervous and Mental Disease*, 167(II), 651-7.
- McGuire, P.K., Shah, G.M.S. & Murray, R.M. (1993). Increased blood flow in Broca's area during auditory hallucinations in schizophrenia. *Schizophrenia Research*, 11(2), 169.
- McNally, R.J. (1994). Introduction to the special series: Innovations in cognitive-behavioral approaches to schizophrenia. *Behavior Therapy*, 25, 1-4.
- McNeil, T.F., Contor-Graae, E., Sjöström, K., Nordström, L. & Rosenlund, T. (1994). Differential etiological factors in schizophrenia: The relationship of obstetric complications to family history and season of birth. *Schizophrenia Research*, 11(2), 92.
- McNeill, D. & Freiburger, P. (1993). *Fuzzy Logic: the revolutionary Computer Technology that is Changing our World*. New York: Touchstone.
- Mednick, S.A., Huttunen, M.O. & Machon, R.A. (1994). Prenatal influenza infections and adult schizophrenia. *Schizophrenia Bulletin*, 20(2), 263-268.
- Mehan, H. & Wood, H. (1975). *The Reality of Ethnomethodology*. New York: John Wiley.
- Meise, U., Kruz, M. & Fleischhacker, W.W. (1994). Antipsychotic maintenance treatment of schizophrenia patients: Is there a consensus? *Schizophrenia Bulletin*, 20(1), 215-226.
- Meldrum, S.J. (1988). Ambulatory monitoring: A bibliography of techniques and technology published in world literature in 1985. *J Ambul Monitoring*, 1(No. 4), 329-332.
- Meldrum, S.J. (1988). Ambulatory monitoring: A bibliography of techniques and technology published in world literature in the third quarter of 1988. *J Ambul Monitoring*, 1(No. 4), 333-334.
- Meltzer, H.Y. (1991). The mechanism of action of novel antipsychotic drugs. *Schizophrenia Bulletin*, 17(2), 263-288.
- Meltzer, H.Y., Cola, P., Way, L., Thompson, P.A., Bastani, B., Davies, M.A. & Snitz, B. (1993). Cost effectiveness of Clozapine in neuroleptic-resistant schizophrenia. *American Journal of Psychiatry*, 150(11), 1630-1638.
- Merckelbach, H.L.G.J. (1989). *Preparedness and Classical Conditioning of fear: A Critical Inquiry*. Maastricht.
- Merrick, W.A. (1992). Dysphoric moods in depressed and non-depressed adolescents. In M.W. deVries (Ed.), *The Experience of Psychopathology: Investigating Mental Disorders in their Natural Settings*, (pp. 148-156). Cambridge: Cambridge University Press.
- Miller, N.E. & Cohen, G.D. (Eds.). (1987). *Schizophrenia and Aging: Schizophrenia, Paranoia, and Schizophreniform Disorders in later Life*. New York: The Guilford Press.
- Minors, D.S. & Waterhouse, J.M. (1981). *Circadian Rhythms and the Human*. Bristol: Wright, J. & Sons, Stonebridge Press.
- Mischel, W. (1968). *Personality and Assessment*. New York: John Wiley & Sons.
- Moldin, S.O. & Erlenmeyer-Kimling, L. (1994). Measuring liability to schizophrenia: Progress report 1994: Editor's introduction. *Schizophrenia Bulletin*, 20(1), 25-29.
- Monk, A. (1991). *Exploring Statistics with Minitab. A Workbook for the Behavioral Sciences*. Chichester: John Wiley & Sons.
- Monk, T.H., Flaherty, J.F., Frank, E., Hoskinson, K. & Kupfer, D.J. (1990). The "Social Rhythm Metric". An instrument to quantify the daily rhythms of life. *Journal of Nervous and Mental Disease*.
- Monk, T.H., Kupfer, D.J., Frank, E. & Ritenour, A.M. (1991). The "Social Rhythm Metric" (SRM): Measuring daily social rhythms over 12 weeks. *Psychiatry Research*, 36, 195-207.
- Monson, C.K. & Hurlburt, R.T. (1993). A commentary to suspend the introspection controversy: Introspecting subjects did agree about "imageless thought". In R.T. Hurlburt (Ed.), *Sampling Inner Experience in Disturbed Affect*, (1 ed., pp. 15-26). New York: Plenum Press.
- Mowrer, O.H. (1950). *Learning Theory and Personality Dynamics*. New York: Ronald Press.

- Mujeeb-ur-Rahman, M. (Ed.). (1977). *The Freudian Paradigm: Psychoanalysis and Scientific Thought*. Chicago: Nelson-Hall.
- Mulder, W.R.E.H. (1994). *Categories or Dimensions? The Use of diagnostic Models in Psychiatry*. Groningen, Enschede.
- Munroe, R.H. & Munroe, R.L. (1971). Effects of environmental experience on spatiability in an East African society. *Journal of Social Psychology*, 83, 15-22.
- Munroe, R.H. & Munroe, R.L. (1971). Household density and infant care in an East African society. *Journal of Social Psychology*, 83, 9-13.
- Musès, C. (1985). *Destiny and control in Human Systems: Studies in the Interactive Connectedness of Time (Chronotopology)*. Boston: Kluwer-Nijhoff Publishing.
- Nelson, R.O. (1977). Assessment and therapeutic functions of self-monitoring. In M. Hersen, R.M. Eisler & P.M. Miller (Eds.), *Progress in Behavior Modification*, (Vol. 5,). New York: Academic Press.
- Nelson, R.O., Boykin, R.A. & Hayes, S.C. (1982). Long-term effects of self-monitoring and reactivity and accuracy. *Behavior Research and Therapy*, 20, 357-363.
- Nelson, R.O., Hay, L.R., Devany, J. & Koslow-Green, L. (1980). The reactivity and accuracy of children's self-monitoring: Three experiments. *Child Behavior Therapy*, 2, 1-24.
- Nelson, R.O. & Hayes, S.C. (Eds.). (1986). *Conceptual Foundations of Behavioral Assessment*. New York: The Guilford Press.
- Nelson, R.O. & Hayes, S.C. (1986). The nature of behavioral assessment. In R.O. Nelson & S.C. Hayes (Eds.), *Conceptual Foundations of Behavioral Assessment*, (pp. 3-41). New York: The Guilford Press.
- Nelson, R.O., Lipinski, D.P. & Boykin, R.A. (1978). The effect of self-recorders training and the obtrusiveness of the self-recording device on the accuracy and reactivity of self-monitoring. *Behavioral Therapy*, 9, 200-208.
- Nicolson, N.A. (1992). Stress, coping and cortisol dynamics in daily life. In M.W. deVries (Ed.), *The Experience of Psychopathology: Investigating Mental Disorders in their Natural Settings*, (pp. 219-232). Cambridge: Cambridge University Press.
- Noorthoorn van der Kruijff, E.O. (1994). *Social Roles and Lifetime Development in the Chronic Mentally Ill*. Unpublished PhD, Rijksuniversiteit Limburg, Maastricht.
- Norman, R.M.G. & Malla, A.K. (1991). Subjective stress in schizophrenic patients. *Social Psychiatry and Psychiatric Epidemiology*, 26, 212-216.
- Norman, W.T. (1967). On estimating psychological relationships: Social desirability and self-report. *Psychol Bull*, 67, 273-293.
- Nuechterlein, K.H. (1983). Signal detection in vigilance tasks and behavioral attributes among offspring of schizophrenic mothers and among hyperactive children. *Journal of Abnormal Psychology*, 92, 4-28.
- Nuechterlein, K.H. (1987). Vulnerability models for schizophrenia: State of the art. In H. Hafner, W.F. Gattaz & W. Janarzik (Eds.), *Search for the Causes of Schizophrenia*, (Vol. 10, pp. 160-203). Berlin: Springer.
- Nuechterlein, K.H. & Dawson, M.E. (1984). A heuristic vulnerability/stress model of schizophrenic episodes. *Schizophrenia Bulletin*, 10, 300-312.
- Nuechterlein, K.H. & Dawson, M.E. (1984). Information processing and attentional functioning in the developmental course of schizophrenic disorders. *Schizophrenia Bulletin*, 10, 160-203.
- Nuttin, J.M. (1974). *The Illusion of Attitude Change, Towards a Response Contagion Theory of Persuasion*. Leuven: Leuven University Press and Academic Press.
- Oades, R.D. (1982). *Attention and Schizophrenia: Neurobiological Bases*. Boston: Pitman Advanced Publishing Program.
- Oddone, E., Guarisco, S. & Simel, D. (1993). Comparison of housestaff's estimates of their workday activities with results of a random work-sampling study. *Academic Medicine*, 68(11), 859-861.
- Offer, D. & Sabshin, M. (Eds.). (1991). *The Diversity of Normal Behavior. Further Contributions to Normatology*. New York: Basic Books.
- Ortony, A., Clore, G.L. & Collins, A. (1988). *The Cognitive Structure of Emotions*. Cambridge: Cambridge University Press.
- Overall, J.E. & Gorham, D.R. (1962). The "Brief Psychiatric Rating Scale" (BPRS). *Psychological Report*, 10, 799-812.
- Pao, P.N. (1979). *Schizophrenic Disorders: Theory and Treatment from a Psychodynamic Point of View*. New York: International Universities Press, Inc.
- Patnoe, S. (1988). *A Narrative History of Experimental Social Psychology. The Kurt Lewin Tradition*. New York: Springer Verlag.
- Paty, J., Kassel, J. & Shiffman, S. (1992). The importance of assessing base rates for clinical studies: an example of stimulus control of smoking. In M.W. deVries (Ed.), *The Experience of*

- Psychopathology: Investigating Mental Disorders in their Natural Settings*, (pp. 347-352). Cambridge: Cambridge University Press.
- Paul, G.L. & Lentz, R.J. (1977). *Psychosocial Treatment of Chronic Mental Patients - Milieu versus Social-Learning Programs*. Cambridge, Mass.: Harvard University Press.
- Paul, G.L., Licht, M.H., Mariotto, M.J., Power, C.T. & Engel, K.L. (Eds.). (1987). *Observational Assessment Instrumentation for Service and Research: The Time-Sample Behavioral Checklist*. (Vol. 2). Illinois: Research Press.
- Paul, G.L., Licht, M.H., Mariotto, M.J., Power, C.T. & Engel, K.L. (Eds.). (1987). *The Time-Sample Behavioral Checklist: Observational Assessment Instrumentation for Service and Research*. (Vol. 2). Illinois: Research Press.
- Paul, G.L., Licht, M.H., Mariotto, M.J., Power, C.T. & Engel, K.L. (Eds.). (1988). *Observational Assessment Instrumentation for Service and Research: The Staff-Resident Interaction Chronograph*. (Vol. 3). Illinois: Research Press.
- Pavlov, I.P. (1927 (1960)). *Conditioned Reflexes. An Investigation of the physiological Activity of the Cerebral Cortex*. New York: Dover Publishing Inc.
- Peck, D.F. & Shapiro, C.M. (Eds.). (1990). *Measuring Human Problems. A Practical Guide to the Assessment of Adult Psychological Problems* (1 ed.). Chichester: John Wiley & Sons.
- Penrose, R. (1989 (1990)). *The Emperor's New Mind. Concerning Computers, Minds, and the Laws of Physics*. London: Vintage.
- Penrose, R. (1994). *Shadows of the Mind. A search for the missing science of consciousness*. Oxford: Oxford University Press.
- Pervin, L.A. (1985). Personality: Current controversies, issues and directions. *An Rev Psychol*, 36, 83-114.
- Peters, C.P. (1991). Concepts of Schizophrenia after Kraepelin and Bleuler. In J.G. Howells (Ed.), *The Concept of Schizophrenia: Historical Perspectives*, (pp. 93-108). Washington: American Psychiatric Press, Inc.
- Peters, U.H. (1991). The German classical concept of schizophrenia. In J.G. Howells (Ed.), *The Concept of Schizophrenia: Historical Perspectives*, (pp. 59-74). Washington: American Psychiatric Press, Inc.
- Petitti, D.B. (1994). *Meta-Analysis, Decision Analysis and Cost-Effectiveness Analysis. Methods for Quantitative Synthesis in Medicine*. (Vol. 24). New York: Oxford University Press.
- Piaget, J. (1966). Time Perception in Children. In J. Fraser (Ed.), *The Voices of Time*, (pp. 202-216). New York: Braziller.
- Piaget, J. (1973). *Main Trends in Psychology*. London: Allen.
- Pickenhain, L. (1970). The higher nervous activity of man. In W.H. Gantt, L. Pickenhain & C. Zwingmann (Eds.), *Pavlovian Approach to Psychopathology: History and Perspectives*, (pp. 126-149). Leipzig: Edition Leipzig.
- Pittam, J. & Scherer, K.R. (1993). Vocal expression and the communication of emotions. In M. Lewis & J.M. Haviland (Eds.), *Handbook of Emotions*, (1 ed., pp. 185-198). New York: The Guilford Press.
- Plomin, R. (1994). *Genetics and Experience. The Interplay between Nature and Nurture*. (Vol. 6). Thousand Oaks: Sage.
- Plutchik, R. (1993). Emotions and their vicissitudes: Emotions and psychopathology. In M. Lewis & J. Haviland (Eds.), *Handbook of Emotions*, (1 ed., pp. 53-66). New York: The Guilford Press.
- Pockock, S.J. (1983). *Clinical Trials: A Practical Approach*. Chichester: Wiley.
- Pogue-Geile, M.F. & Harrow, M. (1984). Negative and positive symptoms in schizophrenia and depression: A follow up. *Schizophrenia Bulletin*, 10, 371-387.
- Pogue-Geile, M.F. & Harrow, M. (1985). Negative symptoms in schizophrenia: Longitudinal course and prognostic significance. *Schizophrenia Bulletin*, 11, 427-439.
- Pogue-Geile, M.F. & Harrow, M. (1987). Negative symptoms in schizophrenia: Longitudinal characteristics and etiological hypotheses. In P.D. Harvey & E.E. Walker (Eds.), *Positive and Negative Symptoms in Psychosis: Description, Research, and Future Directions*, (pp. 94-123). Hillsdale: Lawrence Erlbaum Associates, Publishers.
- Polstra, L. (1988). Chronici-tijd: Een literatuuronderzoek naar de rol van het fenomeen tijd in de rehabilitatie van de chronisch psychiatrische cliënt, Androgogisch Instituut, Groningen.
- Pope, K.S. & Johnson, P.B. (1987). Psychological and psychiatric diagnosis: Theoretical foundations, empirical research, and clinical practice. In D.M. Levin (Ed.), *Pathologies of the Modern Self: Postmodern Studies on Narcissism, Schizophrenia, and Depression*, (pp. 385-404). New York: New York University Press.
- Popper, K. (1956). *The Open Universe*. Totowa, N.J.: Rowman & Littlefield.
- Portegijs, P. (1988). *Wanneer heeft de acute rugpijnpatiënt rugpijn?*. Maastricht: Department of Social Psychiatry.

- Post, R.M. (1992). Transduction of psychosocial stress into the neurobiology of recurrent affective disorder. *American Journal of Psychiatry*, 149(8), 999-1010.
- Prigerson, H.G., Reynolds III, C.F., Frank, E., Kupfer, D.J., George, C.J. & Houck, P.R. (1994). Stressful life events, social rhythms, and depressive symptoms among the elderly: An examination of hypothesized causal linkages. *Psychiatry Research*, 51(1), 33-50.
- Prigogine, I. (1980). *From Being to Becoming*. San Francisco: W.H. Freeman & Company.
- Prigogine, I. & Stengers, I. (1984). *Order out of Chaos. Man's new dialogue with nature*. London: Flamingo.
- Professional Staff of the US-UK Cross-National Project. (1974). The diagnosis and psychopathology of schizophrenia in New York and London. *Schizophrenia Bulletin*, 11(1), 80-102.
- Rabin, A.L., Doneson, S.L. & Jentons, R.L. (1979). Studies of psychological functions in schizophrenia. In L. Bellack (Ed.), *Disorders of the Schizophrenic Syndrome*, (pp. 181-231). New York: Basic Books Inc.
- Rabkin, J.G. (1980). Stressful life events and schizophrenia: A review of the literature. *Psychological Bulletin*, 87, 408-425.
- Racagni, G. & Smeraldi, E. (Eds.). (1987). *Anxious Depression: Assessment and Treatment*. New York: Raven Press.
- Rado, S. (1969). *Adaptational Psychodynamics: Motivation and Control*. New York: Science House.
- Raffaelli, M. & Duckett, E. (1989). We are just talking...: Conversations in early adolescence. *Journal of Youth and Adolescence*, 18(6), 567-582.
- Ray, M.L. (1982). *Advertising and Communication Management*. Engelwood Cliffs, NJ: Prentice-Hall.
- Reis, H.T. (1989). *The impact of social support on transition to adult life* (Final Report to the National Science Foundation). Rochester: University of Rochester - Department of Psychology.
- Reis, H.T. & Wheeler, L. (1991). Studying social interaction with the "Rochester Interaction Chronograph". *Advances in Experimental Social Psychology*.
- Reveley, A.M. (1994). Phenomenology, environmental risk, and genetics: Twin studies of schizophrenia. In N.C. Andreasen (Ed.), *Schizophrenia: From Mind to Molecule*, (pp. 105-120). Washington: American Psychiatric Press, Inc.
- Reynolds, T.D. (1965). Fluctuations in schizophrenic behavior. *Medical Annals of the District of Columbia*, 34(II), 520-49.
- Richards, M.H. & Duckett, E. (1994). The relationship of maternal employment to early adolescents daily experience with and without parents. *Child Dev.*, 65(1), 225-236.
- Ries, R.K., Wilson, L., Bokan, J. & Chiles, J.A. (1981). ECT in medication resistant schizoaffective disorder. *Comprehensive Psychiatry*, 22, 167-173.
- Robbins, M. (1993). *Experiences of Schizophrenia: an Integration of the Personal, Scientific, and Therapeutic*. New York: The Guilford Press.
- Roberts, G.W. & Bruton, C.J. (1990). Notes from the graveyard: Schizophrenia and neuropathology. *Neuropathology and Applied Neurobiology*, 16(1), 3-16.
- Robinson, J.P. (1977). *How Americans use Time: A Social-Psychological Analysis of Everyday Behavior*. New York: Praeger.
- Robinson, J.P. (1987). Microbehavioral approaches to monitoring human experience. *J Nerv Ment Dis*, 175(9), 514-518.
- Robinson, J.P. (1988). Time-diary evidence about the social psychology of everyday life. In J.E. McGrath (Ed.), *The Social Psychology of Time: New Perspectives*, (pp. 134-150). Newbury Park: Sage Publications.
- Roccatagliata, G. (1991). Classical concepts of schizophrenia. In J.G. Howells (Ed.), *The Concept of Schizophrenia: Historical Perspectives*, (pp. 1-27). Washington DC: American Psychiatric Press.
- Roder, V., Brenner, H.D., Kienzle, N. & Hodel, B. (1992). *Integriertes Psychologisches Therapieprogramm für schizophrene Patienten (IPT)*. Weinheim: Psychologie Verlags Union.
- Rogers, K.L. & Winokur, G. (1990). Diagnostic separateness of schizophrenia and affective disorder. In L.E. DeLisi (Ed.), *Depression in Schizophrenia*, (pp. 59-78). Washington: American Psychiatric Press, Inc.
- Romme, M.A.J. & Escher, A.D.M.A.C. (1989). Hearing Voices. *Schizophrenia Bulletin*, 15(2), 209-216.
- Romme, M.A.J. & Escher, A.D.M.A.C. (Eds.). (1993). *Accepting Voices*. London: Mind Publications.
- Rorer, L.G. (1965). The great response-style myth. *Psych Bull*, 63, 129.
- Rosenthal, R. (1978). Combining results from independent studies. *Psychological Bulletin*, 85, 185-193.
- Rosenthal, R. (1979). The 'file drawer problem' and tolerance for null results. *Psychological Bulletin*, 86, 638-641.
- Rosenthal, R. (1984). *Meta-Analytic Procedures for Social Research*. (1 ed.). (Vol. 6). Beverly Hills: Sage.

- Russell, D., Peplau, L.A. & Cutrona, C.E. (1980). The revised "UCLA Loneliness Scale": Concurrent and discriminant validity evidence. *Journal of Personality and Social Psychology*, 39, 472-480.
- Rutner, I.T. & Bugle, C. (1969). An experimental procedure for the modification of psychotic behavior. *Journal of Consulting and Clinical Psychology*, 33, 651-653.
- Rutter, M. & Brown, G.W. (1966). The reliability and validity of measures of family life and relationships in families containing a psychiatric patient. *Social Psychiatry*, 1, 38.
- Sarbin, T.R. (1967). The concept of hallucinations. *Journal of Personality*, 35, 359-380.
- Saugstad, L.F. (1989). Social class, marriage, and fertility in schizophrenia. *Schizophrenia Bulletin*, 15(1), 9-44.
- Savin-Williams, R. & Jaquish, G. (1981). The assessment of adolescent self-esteem: A comparison of methods. *Journal of Personality*, 49, 324-336.
- Schene, A. (1986). *Thuis Bezorgd: een literatuurstudie naar het verschijnsel 'Burden on the Family'*. (Vol. reeks 90). Utrecht: NcGv.
- Schipper, H., Clinch, J. & Powell, V. (1990). Definitions and conceptual issues. In B. Spilker (Ed.), *Quality of Life Assessments in Clinical Trials*, (pp. 11-24). New York: Raven Press.
- Schneider, K. (1939). *Psychischer Befund und Psychiatrische Diagnose*. Leipzig: Thieme.
- Schneider, K. (1959). *Klinische Psychopathologie*. (5 ed.). Stuttgart: Thieme.
- Schooler, N.R. (1991). Maintenance medication for schizophrenia: strategies for dose reduction. *Schizophrenia Bulletin*, 17(2), 311-324.
- Scotti, J.R., McMorrow, M.J. & Trawitzki, A.L. (1993). Behavioral treatment of chronic psychiatric disorders: Publication trends and future directions. *Behavior Therapy*, 24(527-550).
- Scull, A.T. (1977). *Decarceration: Community Treatment and the Deviant*. New York: Prentice-Hall.
- Searle, S.R. (1987). *Linear Models for Unbalanced Data*. New York: Wiley.
- Searles, H.F. (1965). *Collected Papers on Schizophrenia and Related Subjects*. (Vol. 63). London: The Hogarth Press.
- Sedler, M.J. (1991). Concepts of schizophrenia: 1600-1860. In J.G. Howells (Ed.), *The Concept of Schizophrenia: Historical Perspectives*, (pp. 47-57). Washington DC: American Psychiatric Press.
- Sedvall, G. (1994). Positron-emission tomography as a metabolic and neurochemical probe. In N.C. Andreasen (Ed.), *Schizophrenia: From Mind to Molecule*, (pp. 147-158). Washington: American Psychiatric Press, Inc.
- Selten, J.P.C.J. & Slaets, J.P.J. (1994). Second-trimester exposure to 1957 A2 influenza epidemic is not a risk factor in schizophrenia. *Schizophrenia Research*, 11(2), 95.
- Selzer, M.A. & Carsky, M. (1990). Treatment alliance and the chronic schizophrenic. *American Journal of Psychotherapy*, 44(4), 506-515.
- Selzer, M.A., Sullivan, T.B., Carsky, M. & Terkelsen, K.G. (1989). *Working with the Person with Schizophrenia: The Treatment Alliance*. New York: New York University Press.
- Shepherd, G. (1986). Social Skills Training and Schizophrenia. In C.R. Hollin & P. Trower (Eds.), *Handbook of Social Skills Training: Clinical Applications and New Directions*, (Vol. 12/2, pp. 9-38). Oxford: Pergamon Press.
- Shumaker, S.A., Anderson, R.T. & Czajkowski, S.M. (1990). Psychological tests and scales. In B. Spilker (Ed.), *Quality of Life Assessments in Clinical Trials*, New York: Raven Press.
- Siegel, R. & West, L.J. (Eds.). (1975). *Hallucination: Theory and Research*. New York: Wiley.
- Silva, F. (1993). *Psychometric Foundations and Behavioral Assessment*.
- Silver, N.C. & Dunlap, W.P. (1987). Averaging correlation coefficients: Should Fisher's z transformation be used? *Journal of Applied Psychology*, 72(1), 146-148.
- Simon, H.A. (1969). *Sciences of the Artificial*. Cambridge, MA: MIT Press.
- Sjoberg, L. & Magneberg, R. (1990). Action and emotion in everyday life. *Scand. Journal of Psychology*, 31(1), 9-27.
- Skultans, V. (1979). *English Madness. Ideas on Insanity 1580-1890*. London: Routledge & Kegan Paul.
- Slade, P.D. (1976). An investigation of psychological factors involved in the predisposition to auditory hallucinations. *Psychological Medicine*, 6, 123-132.
- Slade, P.D. & Bentall, R.P. (1988). *Sensory Deception: A Scientific Analysis of Hallucinations*. London: Croom Helm.
- Slife, B.D. (1993). *Time and Psychological Explanation*. (1 ed.). New York: State University of New York Press.
- Slobodchikoff, C.N. (Ed.). (1988). *The Ecology of Social Behavior* (1 ed.). San Diego: Academic Press.
- Smeraldi, E. (1988). Genetics and psychopathology for schizophrenic disorders. In E. Smeraldi & L. Bellodi (Eds.), *A Genetic Perspective for Schizophrenic and Related Disorders*, (pp. 1-8). Milan: Edi.Ernes.
- Smeraldi, E. & Bellodi, L. (Eds.). (1988). *A Genetic Perspective for Schizophrenic and Related Disorders*. Milan: Edi.Ernes.

- Smith, J. & Birchwood, M. (1990). Relatives and patients as partners in the management of schizophrenia. *British Journal of Psychiatry*, 156, 654-660.
- Spaulding, W., Garbin, C.P. & Crinean, W.J. (1989). The logical and psychometric prerequisites for cognitive therapy of schizophrenia. *British Journal of Psychiatry*, 155(Suppl.5), 69-73.
- Spilker, B. (1990). Introduction. In B. Spilker (Ed.), *Quality of Life Assessment in Clinical Trials*, (pp. 3-10). New York: Raven Press.
- Spilker, B. (Ed.). (1990). *Quality of Life Assessments in Clinical Trials*. New York: Raven Press.
- Spitzer, R.L., Andreasen, N., Endicott, J. & Woodruff, R.A.J. (1978). Proposed classification of schizophrenia in DSM-III. In L.C. Wynne, R.L. Cromwell & S. Matthysse (Eds.), *The Nature of Schizophrenia*, (pp. 670-685). New York: John Wiley & Sons.
- Spitzer, R.L., Endicott, J. & Robins, E. (1975). Research diagnostic criteria. *Psychopharmacological Bulletin*, 11, 22-25.
- Spring, B. & Ravdin, L. (1992). Cognitive remediation in schizophrenia: Should we attempt it? *Schizophrenia Bulletin*, 18(1), 15-20.
- Srole, L., Langner, T.S., Michael, S.T., Opler, M.K. & Rennie, T.A.C. (1962). *Mental Health in the Metropolis: The Midtown Manhattan study*. New York: McGraw-Hill.
- Steiner, J. (1993). *Psychic Retreats: Pathological Organizations in Psychotic, Neurotic, and Borderline Patients*. (Vol. 19). London: Routledge.
- Stephens, M.A.P., Norris-Baker, C. & Willems, E.P. (1984). Data quality in self-observation and report of behavior. *Behavioral Assessment*, 6, 237-252.
- Stokols, D. (1988). Transformational processes in people-environment relations. In J.E. McGrath (Ed.), *The Social Psychology of Time. New Perspectives*, (pp. 233-252). Newbury Park: Sage.
- Stone, A.A. (1981). The association between perceptions of daily life experiences and self- and spouse-rated mood. *Journal of Research in Personality*, 15, 520-522.
- Stone, A.A., Hedges, S.M., Neale, J.M. & Satin, M.S. (1985). Prospective and cross-sectional mood reports offer no evidence of a "blue monday" phenomenon. *Journal of Personality and Social Psychology*, 47, 129-134.
- Stone, A.A. & Neale, J.M. (1982). Development of a methodology for assessing daily experiences. In A. Baum & J. Singer (Eds.), *Environment and Health*, (Vol. IV, pp. 49-83). New York: Erlbaum.
- Stone, A.A. & Neale, J.M. (1984). Effects of Severe Daily Events on Mood. *Journal of Personality and Social Psychology*, 46, 137-144.
- Stone, A.A. & Neale, J.M. (1984). A new measure for daily coping: Development and preliminary results. *Journal of Personality and Social Psychology*, 46, 892-906.
- Stone, A.A., Neale, J.M. & Shiffman, S. (1993). Daily assessments of stress and coping and their association with mood. *Annals of Behavioral Medicine*, 15(1), 8-16.
- Stone, A.A., Reed, B.R. & Neale, J.M. (1987). Changes in daily event frequency precede episodes of physical symptoms. *Journal of Human Stress*, 13(2), 70-74.
- Strauss, J., Carpenter, W. & Bartko, J. (1974). The diagnosis and understanding of schizophrenia: II. Speculations on the processes that underlie schizophrenic symptoms and signs. *Schizophrenia Bulletin*, 11, 61-76.
- Strauss, J.S. (1989). Mediating processes in schizophrenia. Towards a new dynamic psychiatry. *British Journal of Psychiatry*, 155(suppl.5), 22-28.
- Strauss, J.S. (1994). Is biological psychiatry building on an adequate base? Clinical realities and underlying processes in schizophrenic disorders. In N.C. Andreasen (Ed.), *Schizophrenia: From Mind to Molecule*, (pp. 31-46). Washington: American Psychiatric Press, Inc.
- Strauss, J.S. & Carpenter, W.T.J. (1981). *Schizophrenia*. New York: Plenum Medical Book Company.
- Strauss, J.S., Hafez, H., Lieberman, P. & Harding, C.M. (1985). The course of psychiatric disorder: III. Longitudinal principles. *American Journal of Psychiatry*, 142, 289-296.
- Strongman, K.T. (1987). *The Psychology of Emotions*. Chichester: John Wiley & Sons.
- Stuve, P., Erickson, R.C. & Spaulding, W. (1991). Cognitive rehabilitation: The next step in psychiatric rehabilitation. *Psychological Rehabilitation Journal*, 15(1), 9-26.
- Sudman, S. & Bradburn, N.M. (1982). *Asking Questions. A Practical Guide to Questionnaire Design*. San Francisco: Jossey-Bass.
- Suen, H.K. (1990). *Principles of Test Theories*. (1 ed.). Hillsdale NJ: Lawrence Erlbaum Ass.
- Suen, H.K. & Ary, D. (1989). *Analyzing Quantitative Behavioural Observation Data*. Hillsdale NJ: Lawrence Erlbaum Associates.
- Sullivan, H.S. (1929). *The Interpersonal Theory of Psychiatry*. Washington DC: William Alanson White.
- Sullivan, H.S. (1956). *Clinical Studies in Psychiatry*. New York: Norton.
- Sullivan, H.S. (1962). *Schizophrenia as a Human Process*. New York: Norton.
- Suls, J. & Martin, R.E. (1993). Daily recording and ambulatory monitoring methodologies in behavioral medicine. *Annals of Behavioral Medicine*, 15(1), 3-7.

- Susser, E. & Struening, E.L. (1990). Diagnosing and screening for psychotic disorders in a study of the homeless. *Schizophrenia Bulletin*, 16, 133-145.
- Susser, E.S. & Lin, S.P. (1992). Schizophrenia after prenatal exposure to the Dutch hunger winter of 1944-1945. *Archives of General Psychiatry*, 49, 983-988.
- Szalai, A., Converse, P., Scheuch, E. & Stone, P. (1972). *The Use of Time*. The Hague: Mouton.
- Tarrier, N. (1992). Psychological treatment of positive schizophrenia. In D.J. Kavanagh (Ed.), *Schizophrenia: an Overview and Practical Handbook*, (pp. 356-373). London: Chapman & Hall.
- Taylor, C.B., Fried, L. & Kenardy, J. (1990). The use of a real-time computer diary for data acquisition and processing. *Behav Res Ther*, 28(1), 93-7.
- Tellegen, A. (1982). *The Differential Personality Questionnaire*. Minneapolis: University of Minnesota.
- Tennant, C. (1985). Stress and schizophrenia: A review. *Integral Psychiatry*, 3, 248-261.
- Tennen, H., Suls, J. & Affleck, G. (1991). Personality and daily experience: The promise and the challenge. *Journal of Personality*, 59(September), 313-335.
- Thakor, N.V., Yang, M., Amerasan, M., Reiter, E., Hoehn-Saric, R. & McLeod, D.R. (1989). Micro-computer-based ambulatory monitoring for vital signs of anxiety disorders. *J Ambul Monitoring*, 2(No. 4), 277-294.
- Timpanaro, S. (1976 (1974)). *The Freudian Slip: Psychoanalysis and Textual Criticism* (Soper, K., Trans.). London: NLB.
- Tinberger, N. (1965). *Social Behavior in Animals*. London: Methuen & Co Ltd.
- Tinberger, N. (1972). *The Animal in its World. Volume 1. Field Studies*. London: George Allen & Unwin Ltd.
- Torrey, E.F. (1983). *Surviving Schizophrenia: a Family Manual*. New York: Harper & Row, Publishers.
- Torrey, E.F. (1988). Stalking the Schizo-virus. *Schizophrenia Bulletin*, 14, 223-229.
- Torrey, E.F., Bowler, A.E., Taylor, E.H. & Gottesman, I.I. (1994). *Schizophrenia and Manic-Depressive Disorders: the Biological Roots of Mental Illness as Revealed by the Landmark Study of Identical Twins*. New York: BasicBooks, HarperCollins Publ.
- Torrey, E.F. & Kaufmann, C.A. (1986). Schizophrenia and neuroviruses. In H.A. Nasrallah & D.R. Weinberger (Eds.), *Handbook of Schizophrenia: The Neurology of Schizophrenia Vol. I*, (pp. 361-376). Amsterdam: Elsevier.
- Torrey, E.F., Taylor, E.H., Bracha, H.S., Bowler, A.E., McNeil, T.F., Rawlings, R.R., Quin, P.O., Bigelow, L.B., Rickler, K., Sjostrom, K., Higgins, E.S. & Gottesman, I.I. (1994). Prenatal origin of schizophrenia in a subgroup of discordant monozygotic twins. *Schizophrenia Bulletin*, 20(3), 423-432.
- Toulmin, S. & Goodfield, J. (1965 (1982)). *The Discovery of Time*. Chicago: University of Chicago Press.
- Tschacher, W., Schiepek, G. & Brunner, E.J. (Eds.). (1992). *Self-Organization and Clinical Psychology*. Berlin: Singer.
- Ullmann, L.P. & Krasner, L. (1975). *A Psychological Approach to Abnormal Behavior*. (2 ed.). Englewood Cliffs, NJ: Prentice-Hall.
- van den Bosch, R.J. (1982). Attentional Correlates of Schizophrenia and Related Disorders. In *Search for Objectives Indicators of Psychopathology*. Lisse: Swets & Zeitlinger.
- van der Gaag, M. (1992). *The Results of Cognitive Training in Schizophrenic Patients*. Delft: Eburon.
- van der Gaag, M. (1993). *The Results of Cognitive Training in Schizophrenic Patients*. Delft: Eburon.
- van der Gaag, M. & Haenen, M.A. (1990). Schizophrenia and the Impaired Evaluation of Facial Affect. In P. Drenth (Ed.), *European Perspectives on Psychology, Vol.2*, (pp. 81-89). Chichester: Wiley & Sons.
- Van der Poel, E.G.T. & Delespaul, P.A.E.G. (1992). The applicability of ESM in personalized rehabilitation. In M.W. deVries (Ed.), *The Experience of Psychopathology: Investigating Mental Disorders in their Natural Settings*, (pp. 290-303). Cambridge: Cambridge University Press.
- Van Diest, R. (1992). Vital exhaustion or depression: a study of daily mood in exhausted male subjects at risk for myocardial infarction. In M.W. deVries (Ed.), *The Experience of Psychopathology: Investigating Mental Disorders in their Natural Settings*, (pp. 233-239). Cambridge: Cambridge University Press.
- Van Diest, R. (1993). *Sleep and Vital Exhaustion*. Unpublished Proefschrift, Maastricht.
- van Eck, M.M. & Nicolson, N.A. (1994). Perceived stress and salivary cortisol in daily life. *Annals of Behavioral Medicine*, 16(3), 211-227.
- Van Egeren, L.F. & Madarasmis, S. (1988). A computer-assisted diary (CAD) for ambulatory blood pressure monitoring. *American Journal of Hypertension*, 1, 179s-185s.

- van Heck, G.L. (1989). Situation concepts: definitions and classification. In P.J. Hettema (Ed.), *Personality and Environment: Assessment of Human Adaptation*, (pp. 53-69). Chichester: John Wiley & Sons.
- van Meer, C.R. (1991). *Schizofrenie van nabij*. Delft: Eburon.
- van Meter, K.M. (1990). Sampling and cross-classification analysis in international social research. In E. Øyen (Ed.), *Comparative Methodology. Theory and Practice in International Social Research*, London: Sage.
- van Meter, K.M., deVries, M.W., Kaplan, C.D. & Dijkman-Caes, C.I.M. (1992). States, syndromes, and polythetic classes: developing a classification system for ESM data using the ascending and cross-classification method. In M.W. deVries (Ed.), *The Experience of Psychopathology: Investigating Mental Disorders in their Natural Settings*, (pp. 79-96). Cambridge: Cambridge University Press.
- Van Praag, H.M. (1992). *"Make-Believes" in Psychiatry or The Perils of Progress*. (1 ed.). New York: Brunner Mazel.
- Van Putten, T., Marder, S.R., Wirshing, W.C., Aravagiri, M. & Chabert, N. (1991). Neuroleptic plasma levels. *Schizophrenia Bulletin*, 17(2), 197-216.
- Ventura, J., Green, M.F., Shaner, A. & Liberman, R.P. (1993). Training and quality assurance with the "Brief Psychiatric Rating Scale" (BPRS): "The drift busters". *International Journal of Methods in Psychiatric Research*, 3, 221-244.
- Ventura, J., Lukoff, D., Nuechterlein, K.H., Liberman, R.P., Green, M.F. & Shaner, A. (1993). "Brief Psychiatric Rating Scale" (BPRS) — Expanded version (4.0): Scales, anchor points and administration manual. *International Journal of Methods in Psychiatric Research*, 3, 227-244.
- Visser, H.G. (1993). *Ik zie, Ik zie wat jij niet ziet ?!* (Verslag Wetenschapsstage). Rijkuniversiteit Limburg, Maastricht.
- Wada, T. (1922). An experimental study of hunger and its relation to activity. *Archives of Psychological Monographs*, 8, 1-65.
- Walker, E.E. (1987). Validating and conceptualizing positive and negative symptoms. In P.D. Harvey & E.E. Walker (Eds.), *Positive and Negative Symptoms in Psychosis: Description, Research, and Future Directions*, (pp. 30-49). Hillsdale: Lawrence Erlbaum Associates, Publishers.
- Walker, E.F. (Ed.). (1991). *Schizophrenia: A Life-Course Developmental Perspective*. New York, NY: Academic Press.
- Walker, E.F. (1994). Developmental moderated expressions of the neuropathology underlying schizophrenia. *Schizophrenia Bulletin*, 20(3), 453-480.
- Walker, E.F., Savoie, T. & Davis, D. (1994). Neuromotor precursors of schizophrenia. *Schizophrenia Bulletin*, 20(3), 441-452.
- Walter, E.V. (1988). *Placeways: a Theory of the Human Environment*. Chapel Hill: The University of North Carolina Press.
- Warner, R. (1988). Rhythm in Social Interaction. In J.E. McGrath (Ed.), *The Social Psychology of Time. New Perspectives*, (pp. 63-88). Newbury Park: Sage.
- Watson, J.B. (1924 (1970)). *Behaviorism*. New York: W.W. Norton & Company.
- Watts, F.N. & Bennett, D.H. (Eds.). (1983). *Theory and Practice of Psychiatric Rehabilitation*. Chichester: John Wiley & Sons.
- Wearden, J.H. (1988). Some neglected problems in the analysis of human operant behavior. In G. Davey & C. Cullen (Eds.), *Human Operant Conditioning and Behavior Modification*, (pp. 197-224). Chichester: John Wiley & Sons.
- Wegner, D.M. & Vallacher, R.R. (1980). *The Self in Social Psychology*. New York: Oxford University Press.
- Wells, A.J. (1985). *Variations in Self-Esteem in the Daily Life of Mothers*. University of Chicago, Chicago.
- Wells, A.J. (1992). Variations in self-esteem in daily life. Methodological and development issues. In R.P. Lipka & T.M. Brinthaup (Eds.), *Self-Perspectives across the Life Span*, (pp. 282). Albany: State University of New York Press.
- Wentland, E.J. & Smith, K.W. (1993). *Survey Responses: An Evaluation of Their Validity*. San Diego: Academic Press.
- Wheeler, L. & Reis, H.T. (1991). Self-recording of everyday life events: Origins, types and uses. *Journal of Personality*, 59(September), 339-354.
- Whitehead, A.N. (1925). *Science and the Modern World*. New York: Macmillan.
- Whiting, B. (1977). Changing lifestyles in Kenya. *Daedalus*, 106, 211-225.
- Wicker, A.W. (1979 (1982)). *An Introduction to Ecological Psychology*. Cambridge: Cambridge University Press.
- Wicker, A.W. & King, J.C. (1988). Life cycles of behavior settings. In J.E. McGrath (Ed.), *The Social Psychology of Time. New Perspectives*, (pp. 182-200). Newbury Park: Sage.

- Wiggins, J.S. (1964). Convergence among stylistic response measures from objective personality tests. *Educ Psychol Meas*, 24, 551-562.
- Willems, E. (1969). Planning a rationale for naturalistic research. In E. Willems & H. Rausch (Eds.), *Naturalistic viewpoints in psychological research*, New York: Holt, Rinehart & Winston.
- Williams, J.E. (1976). Self-monitoring of paranoid behavior. *Behavior Therapy*, 7, 562.
- Wilson, K.C.M., Hopkins, R., deVries, M.W. & Copeland, J.R.M. (1992). Research alliance and the limit of compliance: Experience Sampling with the depressed elderly. In M.W. deVries (Ed.), *The Experience of Psychopathology: Investigating Mental Disorders in their Natural Settings*, (pp. 339-346). Cambridge: Cambridge University Press.
- Wirshing, W., Eckman, T.A., Liberman, R.P. & Marder, S. (1990). Management of risk of relapse in schizophrenia through training chronic patients in self-management skills. In C. Tamminga & S.C. Schulz (Eds.), *Schizophrenia Research*, New York: Raven Press.
- Wolf, F.M. (1986). *Meta-Analysis: Quantitative Methods for Research Synthesis*. Newbury Park, CA: Sage Publications.
- Wong, S.E., Martinez-Diaz, J.A., Massel, H.K., Edelstein, B.A., Wiegand, W., Bowen, L. & Liberman, R.P. (1993). Conversational skills training with schizophrenic inpatients: A study of generalization across settings and conversants. *Behavior Therapy*, 24, 285-304.
- Wood, R.L. (1988). Clinical constraints affecting human conditioning. In G. Davey & C. Cullen (Eds.), *Human Operant Conditioning and Behavior Modification*, (pp. 87-118). Chichester: John Wiley & Sons.
- World Health Organization (WHO). (1973). *International Pilot Study on schizophrenia. Vol.1*. Geneva: WHO Press.
- World Health Organization (WHO). (1975). *International Pilot Study on schizophrenia: A Multi National Study*. Geneva: WHO Press.
- World Health Organization (WHO). (1992). *The ICD-10 Classification of Mental and Behavioral Disorders: Clinical Descriptions and Diagnostic Guidelines*. Geneva: WHO.
- Wright, J. (1989). *Address Unknown. Homeless in America*. New York: NY Aldine de Gruyter.
- Wundt, W. (1896 (1911)). *Grundriss der Psychologie*. (10 ed.). Leipzig: Verlag van Wilhelm Engelmann.
- Wynne, L.C., Cromwell, R.L. & Matthysse, S. (Eds.). (1978). *The Nature of Schizophrenia: New Approaches to Research and Treatment*. New York: John Wiley & Sons.
- Yarmey, D. (1979). *The Psychology of Eyewitness Testimony*. New York: Free Press.
- Zigler, E. & Phillips, L. (1961). Psychiatric diagnosis and symptomatology. *Journal of Abnormal and Social Psychiatry*, 63, 69-75.
- Zimmerman, J. & Levitt, E.E. (1975). Why not give your client a counter: A survey of what happened when we did. *Behavior Research and Therapy*, 13, 333-337.
- Zubin, J. (1989). Suiting therapeutic interventions to the scientific models of aetiology. *British Journal of Psychiatry*, 155(suppl.5), 9-14.
- Zubin, J. & Spring, B. (1977). Vulnerability: a new view of schizophrenia. *Journal of Abnormal Psychology*, 86, 260-266.
- Zuzanek, J. (1980). *Work and Leisure in the Soviet Union: A Time Budget Analysis*. New York: Praeger.

Summaries

Summary

The study of daily life adaptation is central for our understanding of psychopathology. In the flow of daily life and in the context of their emotions, thoughts and behaviors, patients experience their illness. In the medical model for somatic illness this experiential process may be considered secondary to the illness itself, but in psychiatry it is central. The alternative modes of giving meaning to the environment and the dynamic process of coping with the changing constraints imposed by the environment reflect the underlying mechanisms that characterize psychopathology. “Assessing Schizophrenia in Daily Life — The Experience Sampling Method” is a monograph on the assessment of daily life adaptation and its relevance for psychiatric research and clinical practice.

Contemporary enhanced assessment strategies of psychiatric nosology — especially in the context of schizophrenia — have not resulted in a better understanding of the psychopathological process and its etiologic mechanisms, nor in improvements of our therapeutic techniques. Even with the most enhanced nosologic instruments — structured interviews and standardized questionnaires — the clustering of cases within the schizophrenic syndrome created no homogeneous disease entity. The heterogeneity is held responsible for the inconclusive research findings into the etiology and optimized treatment of schizophrenia. The improved sophistication of etiologic studies — primarily related to brain research — has not been fed with equally improved phenomenological descriptions of the diversity of schizophrenic subjective experiences. As a consequence, the current research capitalizes on poor diagnostic tools and yields inconclusive results (Chapter 1).

Psychopathology is reflected in the daily life adaptation process and expressed in the subject’s cognitions, behaviors and emotions. Therefore, we need an in-depth study of the basic properties of moment-to-moment changes in these cognitions, behaviors and emotions within and in relation to the natural context in which they occur. This will allow us to understand within subject changes in the illness process and will hopefully lead to a better comprehension of the schizophrenic course, resulting in improved strategies to influence it in treatment (Chapter 2).

For the study of this continual adaptation process we need innovative assessment techniques that reliably and validly describe the variability of the daily life experience. Traditional assessment strategies — such as clinical interviews and standardized questionnaires — fall short for this purpose. They only assess one dimension of the subjective experience: the *subject’s self-narrative* — a more or less consistent self-image that the patient creates over time. In contrast, the *immediate emotional responses* — the reports of dynamic changes from moment to moment in cognitions, emotions and behaviors — should also be assessed. Their momentary states in relation to context as well as the process of change over time are invaluable

indicators of psychopathology. Measurement techniques should: 1) make assessments in the subject's natural everyday environment (ecological validity); 2) reduce the period between the assessed moment and the report herewith avoiding memory distortions; 3) restrict the report to one moment minimizing aggregation; and 4) collect observations repeatedly. To allow the assessment of daily life functioning in its dynamic interaction with the environment, the Experience Sampling Method (ESM) was developed. ESM is a random signal contingent sampling technique that allows the non-reactive assessment of moment-to-moment changes in cognitions, emotions and behaviors as well as the context in which they naturally occur. Subjects carry a digital watch that signals them randomly 10 times each day over the course of a week. At that moment, they fill in an assessment form (Chapter 3).

The study presents the development of the ESM methodology and its standardization for psychiatric research. Over the years we have collected extensive experience with different operationalizations of the ESM methodology. We present the technical arguments that have led to our choices for specific signaling devices and beeping schedules (Chapter 4).

The place of ESM methodology in psychiatric assessment is discussed. Some authors have taken ESM validity and reliability for granted. ESM strategies have often been used to validate traditional cross-sectional assessment methods such as standardized questionnaires. We have argued that ESM reliability and validity are not self-evident and that these psychometric issues should be considered. Using an analysis of the complex ESM dataset we have described which kind of reliability and validity analysis strategy might apply for each of the ESM data types (Chapters 5).

We have conducted empirical studies that illustrate the ESM reliability and validity. The first one demonstrated the test-retest stability of experiential ESM summary scores over long periods of time. We found qualitative differences in the answers that were given within 15 minutes compared to the answers given later. This shows — very powerfully — that ESM is a sensitive method to assess momentary emotions even in schizophrenia. A next study compares ESM to unstructured and structured diaries. While there was an important overlap between both instruments, subjects reported that the cueing of reports with a signal, as well as the structured assessment forms, made the ESM assessment easier to do. The fourth study is an extensive comparison between the ESM time sampling and the RIR event sampling strategies for the assessment of social interaction. Unexpectedly, ESM assessments yielded better ratings of the amount of social interaction as well as of the moods during the interaction. RIR assessments, in contrast, gave better indicators of the personality related characteristics of social interaction. The last validity study compared ESM ratings with structured guesses of time use and experience in context. This demonstrated the utility of the ESM methodology in the population of chronic mental patients. Salient information on time use and experience in context could not be collected less expensively by asking the patient or his case-manager. In a final section we discussed the relevance of these findings related to the ESM methodology for future assessments of psychopathology in clinical trials of new neuroleptics (Chapter 6).

ESM is a powerful, but complex assessment strategy. The reliability and validity of ESM data do not result automatically in valid conclusions from ESM analyses. To

avoid getting lost in the ESM data analysis process we need to be conscious about the questions we try to answer. The distinction between questions about “persons” and questions about “moments” is especially important. Moreover, the multilevel structure of the ESM data has consequences for the appropriate statistical analysis. We argue that elementary statistical techniques are the most appropriate. Only if enough expertise is available, complex statistical methods such as multilevel random regression techniques and meta-analytical strategies can be used to answer ESM questions (Chapter 7).

In the third part ESM is applied to the study of schizophrenia in daily life. The first chapter offers additional circumstantial evidence of ESM reliability and validity in the study of moment-to-moment schizophrenic experience. If the answers of schizophrenic patients were random, we would not have found systematic mental state differences between schizophrenic and normal subjects. The chapter also contains a preliminary assessment of “experience in context” leading to a discussion of the optimal living environment for schizophrenic patients (Chapter 8).

The analyses of the previous chapter were based on a small sample size. We therefore reran the analyses on a larger patient population with more adequate control subjects — depressive chronic mental patients and other mentally ill control subjects. We were able to reconfirm the differences between the samples on mental state dimensions and time use patterns. The amount of time patients spent “alone” or “with their family” was modeled using subject characteristics. Differences were found between chronic depressive and chronic schizophrenic subjects. Overall, social relations tend to restore by age in schizophrenic patients, but deteriorate in depressive subjects. We also demonstrated how — as a group — chronic mental patients felt differently when being in the presence of a different number of persons. The optimal situation — according to experiential ratings — was when between 3 to 7 persons were present for schizophrenia but only 1 person in chronic depression. The optimal working environment, however, was different: one to 2 persons for schizophrenia and alone for depression (Chapter 9).

The next chapter assesses schizophrenic symptomatology in daily life. We found that more patients in our sample hallucinated visually compared to auditively. The intensity within the sample of hallucinators was, however, more extreme for auditive hallucinations. Subject characteristics of hallucinators were related to illness severity, positive illness type, but also averaged moment-to-moment anxiety. Concurrent with a hallucination report, subjects also experienced other psychotic symptoms. Overall, visual hallucinations were more terrifying. During hallucinatory moments subjects were also more lonely and, unexpectedly, were more hungry. Reports of concurrent consumption of alcohol and medication were higher when no hallucinations were reported. We finally modeled the hallucinatory experiences over time. No contextual predictors could be found. Before reporting auditive hallucinations, subjects were, however, more anxious. During the hallucinatory episode, anxiety levels were not related to auditory hallucination intensity, but they were for visual hallucinations, especially in the beginning and at the end of the episode. We finally assessed the influence of concurrent activities on the intensity of hallucinations over the course of

an episode. Withdrawal and engaging in concentrating activities were both found to reduce the moment-to-moment hallucinatory intensity (Chapter 10).

The successfulness of the daily life adaptation — the “Quality of Life” (QoL) — is an important outcome indicator, barely needed for conducting research that can help selecting the appropriate treatment options. We explored how ESM measures can be combined into measures of “QoL” in schizophrenic subjects. We defined a set of ESM-“QoL” indicators. One group is based on productive activities, another group on the “diversity”, “variability” and “stability” of time use characteristics over time. We also assessed QoL-indicators combining time use and concurrent mental state. We assumed that QoL is related to “doing what you like to do most of the time” and the “Flow” (Optimal Experience) model. We found that ESM-QoL indicators were multi-dimensional. Some indicators had more affinity with overall social functioning, while others were related to cross-sectional pathology. Also, indicators using moment-to-moment mood assessments were different from the indicators that used assessments of psychopathology. Further research with these indicators is necessary (Chapter 11).

The fourth and last part of this book assesses the clinical use of ESM data. ESM is not designed to focus on psychopathology alone. The content of ESM assessment-forms also contains the mundane aspects of daily life adaptation. The development of diagnostic tools that customize interventions in the care for chronic mental patients, is warranted. ESM is an important innovative assessment tool that allows individualizing rehabilitative interventions. It can be very valuable in bridging the gap between the clinician “who knows” and the patient “who does not know.” Through ESM the daily life reality of the patient is brought into the clinician’s office. The dialogue between the professional, the patient and the care givers is therefore build on the same information base. This results in a “collaborative effort” in which all the persons concerned regard themselves as equal partners in the management of schizophrenia. The motivational potentials of a shared responsibility in the treatment plan cannot be overestimated. There is a great demand for communication and assessment strategies that foster this shared responsibility. This purposeful collaboration allows a “*personalized rehabilitation*” process. As a consequence, ESM offers a base for a true *negotiated mental health care* (Chapter 12).

The cases treated with the help of ESM demonstrate the need to customize interventions to the dynamics observed in specific patients. In ESM the patient is the specialist in his own life and the data, he has collected, targets the interventions. This process of treatment related assessments is discussed in three extensive case studies (Chapter 13).

Samenvatting

Het onderzoek naar de aanpassing van patiënten met psychische stoornissen aan de dagelijkse realiteit is belangrijk om een beter begrip te verkrijgen in psychopathologische processen. Het is in de stroom van de dagelijkse subjectieve ervaringen — in gevoelens, gedachten en gedragingen — dat de patiënten hun ziekte ervaren. In het medische model voor somatische ziekten zijn subjectieve ervaringen ondergeschikt aan de ziekte zelf, maar in de psychiatrie staan deze centraal. De uiteenlopende methoden om aan de omgeving betekenis te geven en het dynamische proces om te leren omgaan met veranderingen in de omgeving moeten beschreven worden om de onderliggende mechanismen die kenmerkend zijn voor de psychopathologie in kaart te brengen. “Onderzoek van Schizofrenie in het Dagelijkse Leven - De ‘Experience Sampling Method’ (methode voor het verzamelen van ervaringen)” is een studie over het onderzoek naar de omgang van patiënten met schizofrenie met hun dagelijkse omgeving en het belang ervan voor het psychiatrisch onderzoek en de klinische praktijk.

Verbeterde hedendaagse diagnostische instrumenten in de psychiatrie — met name voor schizofrenie — hebben niet geleid tot een beter begrip van het psychopathologisch proces en haar aetiologie, noch tot verbetering van onze therapeutische technieken. Zelfs met de meest geavanceerde nosologische instrumenten in de psychiatrie — gestructureerde interviews en gestandaardiseerde vragenlijsten — blijven de patiënten die de diagnose schizofrenie krijgen een heterogene groep. Dit gebrek aan homogeniteit is verantwoordelijk voor de stagnatie in het onderzoek naar de aetiologie van schizofrenie en de beperkte vooruitgang in het optimaliseren van behandelingen. De voortschrijdende complexiteit van de studies naar de aetiologie — voornamelijk hersenonderzoek — heeft geen gelijke tred gehouden met verbeterde fenomenologische beschrijvingen van de verscheidenheid in schizofrene subjectieve ervaringen. Als gevolg is het lopende onderzoek gebaseerd op een diagnostiek met beperkte waarde en leidt het tot weinig consistente resultaten (Hoofdstuk 1).

Psychopathologie wordt gereflecteerd in de dagelijkse aanpassingsprocessen en komt tot uiting in gedachten, gedragingen en gevoelens van de patiënten. Daarom is een studie vereist van de elementaire kenmerken van veranderingen in gedachten, gedragingen en gevoelens van moment tot moment, binnen de natuurlijke context waarin zij ontstaan. Deze studie zal ons toelaten het verloop in het ziekteproces binnen een patiënt over tijd beter te begrijpen. We hopen dat het zal leiden tot meer inzicht in schizofrenie en ons zal toelaten betere behandelingen op te zetten (Hoofdstuk 2).

Voor de studie van dit continu aanpassingsproces hebben wij nieuwe onderzoekstechnieken nodig die een betrouwbare en valide beschrijving kunnen geven van de fluctuaties in de subjectieve ervaringen zoals ze beleefd worden in het dagelijkse leven. Traditionele onderzoekstrategieën — zoals het klinisch interview en de gestandaardiseerde vragenlijsten — schieten hierin tekort. Zij onderzoeken slechts één

dimensie van de subjectieve ervaringen: het *eigen verhaal van subjecten* — het min of meer onveranderlijk zelfbeeld dat subjecten doorheen hun leven van zichzelf opbouwen. Aanvullend zouden de *onmiddellijke emotionele reacties* — de weerslag van dynamische veranderingen van moment tot moment in de gedachten, emoties en gedragingen van individuen — ook moeten worden onderzocht. Zowel hun status op een welbepaald moment in relatie tot de aanwezige context als hun verandering over tijd zijn waardevolle indicatoren voor psychopathology. De metingen ervan zouden: 1) moeten gebeuren in de natuurlijke dagelijkse omgeving van het subject (ecologische validiteit); 2) de periode moeten verkleinen tussen het observatiemoment en het moment van de rapportage, dit om herinneringsverstoringen te voorkomen; 3) het rapport tot één moment beperken, zodat niet moet worden samengevat; en 4) herhaaldelijk moeten gebeuren. De Experience Sampling Methode (ESM) — letterlijk vertaald een “methode voor het verzamelen van ervaringen” — werd ontwikkeld om het dagelijks functioneren in de dynamische interactie met de omgeving te onderzoeken. ESM is een zelfbeoordelingstechniek waarbij herhaalde observaties gekoppeld worden aan een random tijdssignaal. De subjecten dragen een digitaal uurwerk, dat gedurende één week 10 maal per dag een signaal geeft. Op dat moment vullen de subjecten een vragenlijst in (Hoofdstuk 3).

De studie beschrijft de ontwikkeling van de ESM methode en de standaardisering voor het psychiatrisch onderzoek. Gedurende jaren hebben we in verschillende projecten uitgebreid ervaring met de ESM methode opgedaan. Wij bespreken hier de technische argumenten die hebben geleid tot de keuze van specifieke signaleringsinstrumenten en schema's voor het bepalen van de observatiemomenten (Hoofdstuk 4).

Sommige auteurs hebben de betrouwbaarheid en validiteit van ESM voor waar aangenomen zonder verdere toetsing. ESM is vaak gebruikt om traditionele cross-sectionele instrumenten — zoals gestandaardiseerde vragenlijsten — te valideren. Wij hebben aangegeven dat de betrouwbaarheid en de validiteit niet vanzelfsprekend is en dat deze in psychometrisch opzicht bekeken moet worden. Door middel van een analyse van het complexe ESM databestand, hebben we beschreven wat de aard van het betrouwbaarheids- en validiteitsonderzoek zou kunnen zijn voor elk van de ESM data types (Hoofdstuk 5).

We hebben empirisch onderzoek uitgevoerd dat de betrouwbaarheid en validiteit van ESM kan illustreren. Het eerste onderzoek demonstreerde de test-retest betrouwbaarheid van de geaggregeerde ESM gegevens over een lange tijdperiode. We vonden kwalitatieve verschillen in de antwoorden die gegeven werden binnen 15 minuten in vergelijking tot de antwoorden die later gegeven werden. Dit toont overduidelijk aan dat ESM een gevoelige methode is om de “onmiddellijke emotionele reacties” in kaart te brengen, zelfs bij patiënten die lijden aan schizofrenie. Een volgende studie vergelijkt ESM met ongestructureerde en gestructureerde dagboeken. De resultaten verkregen met de verschillende technieken vertoonden een belangrijke overlap. Echter, de subjecten gaven aan dat ESM gemakkelijker was omdat ze door een signaal op de hoogte werden gesteld, dat ze een vragenlijst moesten invullen terwijl ook duidelijk aangegeven werd welke antwoorden van hen verwacht werden. De vierde studie is een uitgebreide vergelijking tussen de ESM “time sampling” en de RIR “event sampling” methodieken om sociale interacties in kaart te brengen. Hoewel de RIR speciaal op de

gebeurtenis “sociale interactie” gericht is, gaven de ESM metingen onverwachts toch betere schattingen van de hoeveelheid tijd die in sociale interacties werden doorgebracht zowel als van de stemming tijdens deze contacten. De RIR gaf echter een betere indicatie van, aan persoonlijkheid, gerelateerde subjectkenmerken voor interpersoonlijk handelen. De laatste validiteitsstudie vergeleek ESM metingen met gestructureerde schattingen van de tijdsbesteding en aan context gerelateerde stemming. Deze studie toonde aan dat ESM informatie genereert die niet kon worden verkregen door het de patiënt of zijn casemanager gewoon te vragen. Een afsluitende discussieparagraaf gaat in op de relevantie van de ESM methodologie voor het onderzoek naar psychopathologie in clinical trials bij nieuwe neuroleptica (Hoofdstuk 6).

ESM is een krachtige, maar complexe onderzoeksmethode. De betrouwbaarheid en validiteit van ESM gegevens leiden echter niet automatisch tot valide conclusies uit ESM studies. Om te voorkomen dat we verloren zouden lopen in de complexiteit van de ESM data analyses, moeten we ons bewust zijn van de vragen die we willen beantwoorden. Met name het verschil tussen vragen over “individuen” en vragen over “momenten” is uitermate belangrijk. We stellen dat elementaire statistische technieken vaak de meest aangewezen zijn. Slechts wanneer voldoende expertise aanwezig is, kan men overgaan tot het gebruik van complexe statistische methoden zoals de multi-level random regressie technieken en de meta-analytische strategieën (Hoofdstuk 7).

In het derde deel wordt de ESM methode toegepast in de studie van schizofrenie in het dagelijkse leven. Het eerste hoofdstuk voert additionele bewijzen aan voor de ESM betrouwbaarheid en validiteit in deze onderzoekspopulatie. Indien de antwoorden van schizofrene patiënten random gegeven waren, zouden we geen systematische verschillen hebben kunnen vinden tussen schizofrene en normale subjecten. Het hoofdstuk bevat tevens een pilot onderzoek naar “ervaringen in context”. Dit leidt tot een discussie over de optimale woonomgeving voor schizofrene patiënten (Hoofdstuk 8).

De analyses uit het vorige hoofdstuk werden gebaseerd op een kleine patiëntengroep. We hebben daarom de studie opnieuw uitgevoerd bij een grotere groep schizofrene patiënten met een meer aangepaste controlepopulatie — met name chronisch depressieve patiënten. We konden opnieuw verschillen vinden tussen de verschillende groepen met betrekking tot mentale status en de tijdsbesteding. De hoeveelheid tijd die de patiënten alleen of met hun familie doorbrachten, werd gemodelleerd op basis van subjectkenmerken. We vonden verschillen tussen chronisch depressieve en chronisch schizofrene patiënten. In het algemeen werden de sociale relaties van schizofrene patiënten hersteld wanneer ze ouder werden, maar verslechterden ze bij depressieve patiënten. We toonden ook aan hoe chronische psychiatrische patiënten (als een groep) zich anders voelden wanneer een verschillende aantal mensen in hun omgeving was. De optimale situatie, gebaseerd op stemmingsmetingen, werd gevonden wanneer drie tot zeven personen aanwezig waren bij een schizofrene patiënt, terwijl de gemiddelde depressieve patiënt zich enkel goed voelt wanneer niet meer dan één persoon aanwezig is. De optimale werkomgeving daarentegen week hiervan af. Schizofrenen zijn het meest gemotiveerd wanneer twee personen aanwezig zijn, terwijl depressieven het best werken wanneer ze alleen zijn (Hoofdstuk 9).

In het volgende hoofdstuk onderzochten we de schizofrene symptomatologie in het dagelijkse leven. We vonden dat meer patiënten visuele hallucinaties hadden, dan

auditieve. Bij de mensen die hallucineerden was de intensiteit meer extreem bij auditieve hallucinaties. Subjectkenmerken van mensen die hallucineerden waren verbonden met de ernst van de ziekte, een positief symptoomtype en een hoog gemiddeld moment tot moment angstniveau. Wanneer patiënten hallucineerden ervaarden ze ook andere psychotische symptomen. Op die momenten voelden patiënten zich ook meer eenzaam en, onverwachts, ook meer hongerig. In het algemeen waren visuele hallucinaties meer beangstigend. De patiënten gaven vaker aan alcohol en medicatie te hebben gebruikt wanneer geen hallucinaties werden gerapporteerd. Tenslotte hebben we geprobeerd de hallucinatie-episodes te modelleren over de tijd. We konden geen noemenswaardige veranderingen in de omgeving voorafgaand aan een episode vinden. Vóór auditieve hallucinaties echter gerapporteerd werden was het angstniveau bij de patiënten gestegen. Tijdens de episode vonden we echter geen relatie tussen angst en de intensiteit van de auditieve hallucinaties. Visuele hallucinaties waren vooral bij het begin en het einde van een episode gerelateerd aan angst. We onderzochten tenslotte de invloed van activiteiten tijdens de episode. Zich terugtrekken of geconcentreerd werken waren beide gerelateerd met een reductie van de intensiteit (Hoofdstuk 10). Een adequate adaptatie in het dagelijkse leven — een goede kwaliteit van het leven (Quality of Life/QoL) — is een belangrijke indicator in het onderzoek naar de meest aangewezen behandeling. We onderzochten hoe ESM metingen gecombineerd kunnen worden tot QoL-indicatoren. We definieerden er verschillende. Een groep is gebaseerd op produktieve activiteiten, een andere op de “diversiteit”, “variabiliteit” en “stabiliteit” van de tijdsbesteding. We onderzochten ook QoL-indicatoren waarbij tijdsbesteding en momentane stemming geïntegreerd werden. Zo stelden we dat de QoL hoog is wanneer je “de meeste tijd doet wat je graag doet” of “uitdagingen in je leven ervaart” (Flow-model). We vonden dat ESM QoL-indicatoren multidimensioneel waren. Sommige hadden kenmerken van sociaal functioneren en andere waren gerelateerd aan de cross-sectionele psychopathologie. Verder onderzoek is nodig (Hoofdstuk 11).

In het vierde en laatste deel wordt de ESM methodiek toegepast in de klinische praktijk. ESM werd niet ontworpen om slechts psychopathologie te meten. De inhoud van de ESM vragenlijsten bevat ook informatie over de meer gewone aspecten van het dagelijkse leven. We hebben nood aan diagnostische instrumenten die ons in staat stellen behandelingen aan de noden van individuele patiënten aan te passen. ESM is een belangrijk innovatief onderzoeksinstrument dat ons in staat stelt individuele interventies te ontwikkelen. Het kan waardevol zijn om de kloof te overbruggen tussen de behandelaar die “het weet” en de patiënt die “het niet weet”. Door middel van ESM komt de realiteit van het dagelijkse leven van de patiënt in het kantoor van de behandelaar binnen. De dialoog tussen de clinicus, de patiënt en zijn familie kan dan ook gebaseerd worden op dezelfde informatie. Hierdoor worden alle betrokken personen evenwaardige partners in het beheersen van de schizofrene problematiek. De motivationele kracht die hierdoor ontstaat kan niet worden overschat. Deze doelgerichte samenwerking leidt tot een *gepersonaliseerd rehabilitatieproces*. Hierdoor wordt een ware *onderhandelde psychiatrische zorg* mogelijk (Hoofdstuk 12).

De patiënten die behandeld werden met behulp van ESM, toonden aan hoe belangrijk het was interventies aan te passen aan de dynamiek die we bij hen konden observeren. Dit proces wordt geïllustreerd aan de hand van drie gevalstudies (Hoofdstuk 13).

Curriculum Vitae

PHILIPPE DELESPAUL is a Belgian psychologist (°1958). He received his professional degree at the University of Louvain where he studied clinical psychology, mathematical psychology and psychological methodology. He is a registered clinical psychology supervisor and a behavioral therapist who teaches in therapist training programs in Belgium and The Netherlands. Since 1982 he worked at the “Vijverdal” psychiatric hospital in Maastricht. Currently he heads a therapy and rehabilitation ward for chronic mental patients. He is a lecturer in the University of Limburg — in the Department of Psychiatry and Neuropsychology, Section of Social Psychiatry and Psychiatric Epidemiology. He is a research scientist at the International Institute for Psycho-Social and Socio-Ecological Research (IPSER) where he participates in the Ecological Psychiatry research program.

.....