Consensus in psychotherapy

Citation for published version (APA):

Dormaar, J. M. M. (1989). *Consensus in psychotherapy*. [Doctoral Thesis, Maastricht University]. Maastricht University. <https://doi.org/10.26481/dis.19900119jd>

Document status and date:

Published: 01/01/1989

DOI:

[10.26481/dis.19900119jd](https://doi.org/10.26481/dis.19900119jd)

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.

Consensus in psychotherapy

Proefschrift

ter verkrijging van de graad van doctor
aan de Rijksuniversiteit Limburg te Maastricht,
op gezag van de Rector Magnificus, prof. dr. F.I.M. Bonke,
volgens het besluit van het College van Dekanen,
in het openbaar te verdedigen op
vrijdag, 19 januari 1990
om 16.00 uur

door

Jacobus Marius Maarten Dormaar
geboren te Bloemendaal in 1939

Promotores: Prof. dr. M.W. DeVries
 Prof. dr. M.A.J. Romme

Beoordelingscommissie: Prof. dr. H. Philipsen (voorzitter)
 Prof. dr. J.J.C.B. Bremer
 Prof. dr. R. van Dyck
 Dr. O van der Hart
 Prof. dr. G.H. DeVries

Consensus in psychotherapy

To: Adanech, Saskia, Judith and Elizabeth

© 1989 Maarten Dormaar

Chapters 4, 5, 6, 7 and 9 were reprinted with kind permission of the publishers of the original publications.

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Dormaar, Jacobus Marius Maarten

Consensus in psychotherapy / Jacobus Marius Maarten

Dormaar. - Maastricht : Datawyse. - Ill.

Thesis Maastricht. - With ref. - With summary in Dutch.

ISBN 90-5291-016-2

SISO 607.2 UDC 615.851(043.3)

Subject headings: psychotherapy / cybernetics /
hypnotherapy.

Cover design: Judith Dormaar

Production: Datawyse Maastricht / Piet Portegijs

Printing: Krips Repro Meppel

Publication of this book was sponsored in part by Schering Nederland bv,
Duphar Nederland bv, Organon bv (part of AKZO Pharma Division),
the IPSE foundation and RIAGG Maastricht

Contents

	Acknowledgements	8
1	Introduction	11
2	Basic considerations connected with consensus	17
3	Consensus in philosophy of science, sociology and psychology	23
	Comment	58
4	Exposure and leave-taking rituals	61
	Comment	70
5	February man revisited	71
	Comment	80
6	Psychotherapeutic outcome: Measuring reliable clinical change	83
	Comment	92
7	Consensus in patient-therapist interactions	95
	Comment	105
8	Consensus in terms of similarity	107
	Comment	118
9	Individualized metaphors in hypnotherapy	121
10	Conclusion	129
	References	144
	Index	159
	Appendix: Instruments for consensus measurement	163
	Summary	172
	Samenvatting	174
	Curriculum vitae	176

Acknowledgements

Longstanding grappling with the difficulties of genuine mutual understanding, so vital to maintaining a personal relationship, spurred my interest into the issue of consensus and focused my professional attention since 1981 toward the possible relevance of consensus in patient-therapist interactions. The subject of this thesis testifies, in a way how I have tried to succeed in theory and research where I have failed in practice. Living on my own for the past year was a blessing in disguise since I had ample opportunity for basic theoretical reflection and commitment to the task that I might not have had amidst a busy family life.

A dissertation should be proof of the doctoral student's ability to do research on his own. I have done the present work independently in the sense of self-directed or autonomous, but certainly not in the sense of self-reliant or unattached. As in any research project, many persons were involved. First, my former co-worker Chantal Dijkman. She was - and continues to be - a highly reliable and inventive colleague. Jos Slangen and Juni Daalmans were the dedicated research assistants at the time of data-gathering. Second, I am indebted to my promotor Marten deVries in many ways. Every idea that I have tried out has passed his scrutiny and I have profited a lot from his stimulating and wide interests. I wish for any doctoral student a collaboration with his promotor like that I was so fortunate to enjoy. To my second promotor Marius Romme I am especially grateful for putting me on the right track and for his unabating support, both personally and professionally. The research detailed in this thesis arose from the results of his pilot study in 1979. Both of my promoters have generously created an atmosphere of academic freedom to pursue my interests. Maybe this is becoming a rare luxury nowadays, and I acknowledge the risks they have taken with me.

Of all others who have helped me to complete this work I will only mention a few. First, there is my secretary in the department of Social Psychiatry, Trees Soute. Promptly and without my noticing any complaint, she has prepared my entire manuscript and has kept order in my creative chaos. Nancy Nicolson's careful proofreading provided the finishing touch to make a readable text from the manuscript. Wybrandt van Schuur of the department of Statistics and Measurement Theory at Groningen University is not only a staunch and long-time friend, but also an erudite and widely interested scientist with invaluable advice on methodology. I owe him the idea of similarity coefficients for measuring consensus. I needed prolonged discussions between the two of us to make my point on significance, but then Wijbrandt called in Tom Snijders, professor of Mathematical Sociology at Utrecht University, who solved

my problem elegantly in a mathematical way. The result of our collaborative effort is reviewed in this thesis.

At the start of my research efforts Max Beekers, now head of the RIAGG Midden-Limburg, and Ad Appels, professor of Medical Psychology at our University, have been of great help in formulating the design. I also thank the Foundation 'Nationaal Fonds voor de Geestelijke Volksgezondheid' for their financial support in the early years. I suspect by this time they have lost all hope that their subsidizing would still materialize in published results.

I am further indebted to all my colleagues of the department of Social Psychiatry, who - perhaps unwittingly - have sharpened my ideas during those years of congeniality. I cannot fail, however, to single out Egbert van der Poel and Jan Scholte, who were always there if I needed to check the soundness of flashing ideas. Together with Margot Breuls, Riek Wijsen and Jan Olthof, I have studied the works of Bateson and other authors in cybernetics and family therapy. Our stimulating discussions have greatly influenced my approach to psychiatry. Shared interests in the 'ecology of mind' and hypnotherapy have resulted in a paper co-authored with Jan Olthof which has been included in the present thesis.

The research on consensus was carried out at the RIAGG Maastricht. Thanks to being employed to work both at the University and this comprehensive Community Mental Health Centre, I could reflect on our common treatment efforts both as an insider and as a researcher. The investigations have been supported by all my colleagues of the teams 'West' and 'Gulpen', but most notably by the members of the advising committee, Jos Weyers, Jo Smeets, Mieps van Hest, Leon Ghijsen and Louk Maas. I also want to express my gratitude to the former and present director of the 'RIAGG', Ton Hilberink and Harry Pomerantz, for approving the involvement of patients and therapists of their institution in my research.

To do psychiatric research and clinical practice requires a personal commitment. A clear division between personal and professional relationships I have therefore always deemed artificial. I will remember the love, kindness and support that were so vital to me, especially while carrying on in a time of trials and tribulations. A number of unnamed friends, in- and outside my professional life, will know that my gratitude goes to them.

1 Introduction

In the daily practice of psychiatry, treatment is mainly comprised of verbal interactions between patient and therapist, aimed at ameliorating the patient's suffering, removing the disorder or helping with problems in living. In the broadest sense these more or less systematic and planned conversations are labeled psychotherapy. Research attempting to measure the efficacy and describe the process of psychotherapy can be categorized as research examining specific and nonspecific factors. Research into specific factors is mainly focused on the isolation of effective components of treatment techniques, but also investigates the processes and effects of treatment with certain disorders or patient groups. The interest in nonspecific or common factors (Frank, 1974) is based on the idea that all forms of psychotherapy -whatever the characteristics of the patients and the problems to be treated- have some general features that may be influential and therefore are worthy of investigation.

This thesis aims to make a contribution to research into common factors affecting the outcome of psychotherapy by investigating the clinical importance of *consensus* between patient and therapist. As the concept of consensus has not previously been used in psychotherapy research, the relevance of the concept to the study of psychotherapy will be extensively reviewed and discussed. In doing so, I will first highlight the background of this study and present an overview of the different chapters. These include an elementary theoretical exploration, a review of studies on consensus in different fields, case material and own research findings on consensus, a validation of the operationalizing of consensus and finally, to demonstrate the wider relevance of the principle behind the operationalization, a case description and a suggested research application.

During psychotherapeutic training and subsequent practice I have learned a number of theories and models such as psychodynamic, client-centered, behavior and family therapy and finally hypnotherapy. All of them try to account for those types of human suffering that typically bring people to a psychiatrist's office and the techniques or approaches based on them tell the therapist what he can and should do in helping his patients to overcome their handicapped functioning. Practical experience and research finding (e.g. Smith, Glass and Miller, 1980), taught me, like every beginning psychiatrist or psychotherapist, that training and growing experience still do not guarantee successful outcomes.

I think I went through the same ordeal as most of my colleagues by learning that you may very well fail in a case you considered similar to the one you had previously treated successfully and therefore decided to approach

along the same lines. Although we all have to settle for the general unpredictability of human nature and conduct, I felt that part of my limited success as a therapist was due to the fact that, with most attention going into matters of therapeutic techniques, of understanding and diagnosing disorders, I had left out or at least underestimated, something of vital importance. I thus came to realize more fully that no therapy works unless patient and therapist cooperate. This means that they should be willing and able to share outlooks on the problems to be tackled and to agree on the mode of treatment to some extent. Phrased somewhat more technically, I sensed that the *consensus* between patient and therapist might be of influence on their collaborative efforts, needed to achieve a successful outcome. Studying consensus in the psychotherapeutic setting belongs to the type of psychotherapy outcome research that studies general treatment conditions, like exemplified by the work of Jerome Frank and his research group on so-called common factors (e.g. Frank et al., 1978).

For a true understanding of the way in which I have extended this research tradition with the study of consensus in psychotherapy, it is important to state here that systems theory -also known as cybernetics in the English language domain (Lilienfeld, 1978)- has been the scientific approach that inspired and guided my interest in the subject.

Within systems theory itself however I discern a basic dichotomy between a so-called 'empirical position' and an 'observer-dependent position' regarding social science in general. The empirical position is exemplified by Parsons (1960, 1964) who sees socio-cultural systems as organized according to role and structure and characterized by stability, hierarchy, power and control. 'At the heart of this empirical position is a firm commitment to a hypothetical-deductive model of explanation that is based on appeal to external scientific law. Thus, social systems are objectively defined and exist independently of the observer. The observer can only inquire as to the representational adequacy of the language used to describe the system. Knowing how a system should be, therefore, places the observer, the therapist, in the independent position of determining how the system is (pathology/ diagnosis) and fixing the system as it should be (normalcy/treatment). All this is according to the therapist's objective descriptions' (Anderson & Goolishian, 1988, p. 376).

The observer-dependent position on the other hand is characterized by the idea that in doing psychotherapy or scientific research we should acknowledge more fully that every human observer is always actively constructing, altering and validating reality through his operations, which also include the mere acts of perception. The 'increasing dissatisfaction with the view that social sciences and psychotherapy are to be thought of in the terms of the logical empiricism of the so-called natural sciences' (Anderson & Goolishian, 1988, p. 376), has led clinicians and researchers who employ a cybernetic framework to

the realization that as observers we are always intricately connected with our observations. I therefore see as the axiomatic motto of observer-dependence Maturana's statement that 'anything said is said by an observer' (Maturana & Varela, 1980, p. 8). Many authors in the field of cybernetics and family therapy, notably those adhering to so-called second cybernetics, strategic and ecosystemic therapies, have formulated the theoretical tenets that guide practice and research from this position (Auerswald, 1985, 1986, 1987; Bateson, 1972; Bogdan, 1984; Dell, 1982, 1985; Doherty, 1986; Hoffman, 1981; Keeney, 1979, 1983; Schwartzman, 1984; Watzlawick et al., 1974).

I believe it has foremost been clinical practice that showed me the usefulness of this observer-dependent stance. Although it makes research and the reporting of it more complicated, I have gradually come to adhere to this rather unorthodox view in my scientific work as well.

From this position and in accordance with its axiomatic proposition, it follows that I first have to explicate some of the basic premises that have guided my observations and that may be deemed rather distinct from the implicit propositions that often are taken for granted when engaging in social scientific work. Such an explication also means that I think there is an inevitable and necessary coherence between the basic theoretical assumptions to be formulated in **chapter 2** and my view on consensus in psychotherapeutic practice and research.

I realize that statements that touch on the field of philosophy of science and even epistemology, surpass the true field of psychiatry in a traditional sense. On the other hand I feel - like the authors in the field of family therapy referred to above - that such an excursion is justified, when dealing with matters that have not been addressed directly in one's own field before. In such cases it is good practice first to consult neighbouring disciplines. The selective overview of sociological and psychological models and research on consensus as presented in **chapter 3** testifies to how I have sought advice on the subject of consensus from these fields that already have a longstanding tradition of supporting and inspiring psychiatry proper. But still, the fundamental criticism regarding social science that is inherent to the observer-dependent view within systems theory, spurred my curiosity into the philosophical disciplines, where I see the critical stance toward science and claims of knowledge as the central concern. Since I am presenting myself primarily as a practicing psychotherapist and clinical researcher, my theoretical orientation in chapter 2 does not represent an in-depth argument for the observer-dependent position in psychiatric research. It serves however my explicit purpose that other observers -whom I consider as just as dependent as myself but more versed in the field- can comment on the merits of my studies and improve their point of departure.

To further orient the reader to the pattern that connects the different chapters of the book, I will present some general ideas on the relevance of consensus to psychotherapy in connection with some statements on the next four chapters.

As said before, my interest in consensus grew from my practical experience as a therapist. Although **chapters 4 and 5** mainly deal with the application of specific treatment techniques, they properly belong to this thesis since they may show how I have learned to see the consensual aspects of the psychotherapeutic process. The comments on the chapters may serve to make this clear to the reader. The second reason to include these case studies is, that too often research findings seem to have little bearing on clinical practice. The integrated approach of research and practice advocated by me, may enhance the pertinence of the subject of consensus for the reader with foremost clinical interests.

As a prelude to these clinical illustrations I also want to offer some general remarks on the importance of consensus to successful psychotherapy. What will strike a naive observer when watching and listening to a psychotherapeutic encounter, is that there are just two or more persons talking with each other. This means that the practice of psychotherapy can be seen as a special kind of *conversation* in the first place. Common to all forms of conversation is that the participants form a temporary socio-cultural system that is foremost a *meaning-processing system* in which 'the interacting participants maintain and transform the identity of themselves and of their network through a more or less shared understanding of both themselves and the world' (Braten, 1984, p. 193). Shared understanding -and concomitant agreement- are forms of consensus according to the co-orientation model of Scheff (1967), presented in chapter 3, and used in my research.

In ordinary conversations, like small talk, consultation, advice and telling stories to each other, consensus is more or less presupposed in the same way as Farrell (1976) describes for rhetoric discourse. This means, that the participants assume that each of them know more or less where they are talking about and what the words said really mean, i.e. to what contexts the words refer or belong.

In psychotherapy however, consensus is not taken for granted. The therapist typically does not assume beforehand that he understands his patient, but he asks for details and clarifications, even when matters seem sufficiently clear for regular communication. Because of the nature of psychotherapy, the therapist has to strive for consensus with the patient on the naming of quite personal and private experiences of one of them. Mostly the patient is not used to share these experiences fully and he is certainly not inclined to let another person intrude in such a way that his personal feelings and beliefs are commented on and relabeled. Nevertheless, this is exactly what is going on in the practical execution of psychotherapy.

Furthermore, achieving consensus regarding the private experiences of the patient, is feasible only if there also exists a different level of consensus between patient and therapist. That is the kind of consensus that can be called the mutual understanding and tacit agreement that it is possible at all that two persons -or even more- can share (at least in language) the private world of one of them and that this may be beneficial. When working with psychotic patients, I think therapists can become aware that this level of consensus is not self-evident to all of us.

Granted that consensus may be an important aspect of the therapeutic process, on what topics has consensus to be reached by patient and therapist? Considering the phases and tasks that generally belong to such a process, I think it useful to discern consensus on problems, treatment procedures or techniques and goals. For practical purposes I have limited the clinical research to the measurement of consensus regarding problems. This research has been reported in chapter 6 and 7.

Chapter 6 presents the results of a statistical procedure for differentiating improved and unimproved patients. The prediction of improved and unimproved outcome by means of consensus is reported in **chapter 7**.

The final part of the thesis, consisting of chapters 8, 9 and 10, does not deal directly with consensus. **Chapter 8** presents the logic of a similarity coefficient to compute the relational measures of consensus and the manner to test the reliability and validity of such measures. In **chapter 9** I show that similarity is also the basic principle behind metaphor, which is a powerful tool in psychotherapy.

After a short overview of the preceding chapters and some final comments, the concluding **chapter 10** introduces the possible use of similarity coefficients for an alternative approach to psychiatric diagnosis. This last part of the thesis is intended to convey the integration of the leading concepts of this thesis, namely the relational nature of psychiatric data, like embodied in the relational aspects of consensus, the observer-dependence of our statements -e.g. diagnostic- and the benefits of an interdependent view on clinical practice and research in psychiatry.

2 Basic considerations connected with consensus

2.1 Science, facts and relationships

Although science is occupied with providing definitions, science itself is difficult to define. Dynamically, science is best described as an activity. Science is what scientists do. This view emphasizes the heuristic quality of science, meaning that science serves to discover and reveal, to pose and solve problems apart from merely adding to existing knowledge. In this heuristic sense, established facts are important insofar as they further the development of theory, discovery and investigation. So we certainly need facts, but what do we mean by a fact, scientifically speaking?

All scientific facts are essentially relationships (Kerlinger, 1973). In psychiatry as a scientific discipline this relational character of facts is quite obvious. By studying for instance a disease concept like schizophrenia, we relate disorder variables to other variables like age of onset, social class or a certain treatment regimen. The relationships that are discovered constitute the facts we were looking for. This example also points out that every concept embodies relationships, such as the concept schizophrenia embodies relationships between hallucinations, thought disorders and biochemical imbalances.

The acknowledged definition of a scientific relation is a set of ordered pairs of variables. Variables, in turn, are deduced from attributes of an individual or some other kind of unit (Cronbach, 1970). When we speak of attributes, we likewise refer to sets of interactions or relations. Calling a stone 'hard' means that when poking into it, it resists penetration and secondly that there exist strong and continuous interactions between the molecules of the stone (Bateson, 1979).

I emphasize the relational characteristics of facts and attributes, to highlight the idea that man perceives and understands reality basically in terms of relationships. This view is essential to understanding consensus as a direct *relational* variable.

By nature, social scientists are interested in men as *social* beings, that is in their interrelationships. Referring to the earlier proposition that relationships are the essence of science, it is quite remarkable that psychiatry - or even psychology and sociology for that matter - seldom investigate human relationships in a direct, quantitative way. While love, hate, respect or dependency are written about extensively, these core indicators of human relationships have rarely reached the status of a scientific variable.

I have blurred here intentionally the distinction between scientific and human relationships. Instead of presenting all possible similarities between both types of relationship that warrant this view, I will limit my argument as follows.

In usual R-methodology a scientific relationship is the operationalization of the linking of two concepts as a set of ordered pairs in the sense that the set of values of the first and second variable is ordered by means of all the units from which the variables are deduced. If however we want to describe and measure a human relationship from the standpoint of a scientific outside observer, this means that we are interested in the links between two units regarding a set of variables. In this case we can likewise define a relationship between units - instead of concepts - as a set of ordered pairs. This time the scientific relationship is the operationalization of the linking of two *persons* as a set of ordered pairs in the sense that the set of values on all relevant variables derived from the first and second person is ordered by means of these variables. This modification in establishing scientific relationships is fundamental to Q-methodology (Stephenson, 1953).

The basic mathematical expression of a relationship is a ratio, like in a coefficient. Put differently, a ratio is the primary embodiment of a relationship in numbers that denote or are derived from variables. As pointed out above, it is well-known that every relationship can figure in turn as a variable. So indeed, once established as a relationship in the scientific sense, a human relationship can be treated as a variable. Accordingly, in subsequent sections of this thesis consensus will mainly be presented as a *relational* variable, meaning that the variable has different values that denote the degrees of the consensual relationship between two persons.

Returning to the argument on the lack of variables denoting human relationships, I first acknowledge that the inevitable restrictions of a scientific model partially accounts for the difficulties in capturing adequately these everyday social phenomena in a quantitative manner. Models in physics and other 'hard' sciences however, explicitly include the same concepts such as time, mass and length, that guide our daily observation of physical reality. Therefore other factors may also contribute to this lack of correspondence between scientific investigation and ordinary descriptions of social reality. For instance, the selection of variables is often restricted to the realm of attributes of those entities that we - based on Western subject-predicate modes of thinking - perceive as units, namely individual persons (Bateson, 1979; Bateson & Bateson, 1986).

It's obvious indeed that in fields like sociology and social psychology, which explicitly deal with man as a social being - that is interactional and relational - many variables have been conceived as attributes of rather atomistic individuals. To illustrate the point in a way that is quite pertinent to psychiatry, we may note that most words that describe emotions - like love or

shame - are also used to indicate human relationships. emotions then, are to a large extent regarded as 'something' of an individual, at the expense of the interactive, communicative and intersubjective side of the affair. When speaking of 'being in love', we still keep intact that intersubjective and relational nature of emotions. But we lack the convention of speaking of persons in terms of being 'in shame' or 'in hate', thereby encumbering us from attending to the intersubjective side of emotional states (Rabkin, 1985). In psychiatry - or, more properly, in interactional psychotherapy - the fundamental interrelatedness of emotions and human relationships has been clearly articulated in the observation that emotions are man's - mainly nonverbal - means to communicate his definition of the relationship with his fellow-men (Watzlawick et al., 1967). In this view, emotional responses are called 'meta-communicative'. They *qualify* the verbal communications, which means that emotions form the relationship-context that largely determines the actual meaning of the words said, like in the desperate or angry husband who yells to his wife 'I love you!'.

In the field of psychiatry, where human emotions and relationships are a central concern, we may draw the conclusion that the lack of direct measurement of feelings and relationships, in such a way as to justify the relational character of emotions, severely limits the investigation of this core issue. But we can do better. Since emotions and relationships obviously exist in degrees of intensity, we can conceptualize these concepts in principle as variables. To prove this last assertion for the total range of relationships lies beyond the scope of this thesis. In this thesis I will therefore focus on the interactional or relational operationalization of consensus.

One difficulty in defining the concept of consensus is, that it cannot be equated with other, more familiar words that denote a specific relationship. We must therefore settle for statements such as that consensus connotes agreement, understanding and realization of being understood (Scheff, 1967; Laing et al., 1966) and that, affectively, consensus is associated with empathy, and an accepting attitude together with a feeling of being accepted or confirmed in one's personal experience. Behaviorally, consensus implies coordination and cooperation (O'Keefe, 1973). Colloquially, consensus is used rather loosely in phrases like 'there is little consensus on this theory' or 'a consensus-meeting on treatment of renal failure'. We then assume that people implicitly know what consensus is and that it is a relationship worth aiming for.

2.2 Consensus and reality

In the first place, consensus is an ubiquitous and fundamental human experience. A certain degree of consensus is a basic condition in man's social life. When taking consensus as an object of scientific inquiry however, we

largely abstract from this personal, experiential nature of consensus and are able to distinguish different types and levels of consensus as befit the specific interests in question. So we might differentiate consensus between scientific observers, like in interrater reliability, from consensus as an element of social reality like in sociology, and this kind from the interest of consensus to questions of epistemology and the philosophy or sociology of science, especially since Kuhn (1970).

I will not pursue this line, since I just want to highlight here the centrality of the concept and the commonalities across different fields and subjects. Thus, far from pretending to present a thorough overview that would surpass my competence and the goals of this introduction, I will offer a seemingly arbitrary set of observations and selected literature, which have in common that they guided my line of thought during and after the research period.

So I first note that a common orientation and success in 'mind reading' are needed for cooperation, division of labor and reliable expectations about our fellow-men. But consensus is also vital to the maintenance of our daily selves and the world as we perceive it.

When Thomas and Thomas (1924) concluded '*... if man define situations as real, they are real in their consequences*', they already subscribed to the social nature of reality by indicating that man does not define reality in solitude. More recently, the neurobiologist Maturana (1978, 1988) has argued at length that reality, explanations, language and accepted claims of knowledge depend on the constitution of a '*consensual domain*'. With this term he refers to a field of observations, propositions and the like that are put into language and are accepted as true, real and objective by those who share the domain. Without it, no reality exists that can be objectively - that is independent of the observer - named, pointed out or described. Therefore, Maturana speaks of '*objectivity-in-parenthesis*'. Consensus is central in his model of observer-dependent reality: *Language as a domain of recursive consensual coordinations of actions does not operate with symbols, yet symbols arise in language as distinctions of relations of distinctions. Also, according to this, words are no symbolic entities, nor do they connote or denote independent objects. They are distinctions of consensual coordinations of actions in the flow of consensual coordinations of actions* (Maturana, 1988; p. 47).

In his '*ontology of reality*' Maturana claims that we have actually created reality ourselves through our rational verbal exchanges with others and ourselves. Consequently we are able to refer to that self-made reality as something that we deem to exist independently of how we act in it or talk about it.

Authors in the field of cybernetics and family therapy, have stressed the recursive and observer-dependent nature of our perceptions (Von Foerster, 1981; Varela, 1976; Keeney, 1983), and the constructive aspect of reality (von Glaserfeld, 1984; Segal, 1986). In daily life however, we are not constantly

aware of the interactional nature of our world-view. We tend to rely on premises of perception that are largely unconscious (Bateson, 1972, 1979) in order to maintain a predictable interaction with reality, that results in a firm conviction in the objectivity of the world, 'out there'. The rocklike belief in a reliable reality independent of ourselves is built up and maintained thanks to others. Developmentally it is axiomatic, that the external reality of the person becomes a social reality. In his novel 'Vendredi ou les limbes du Pacifique', Michel Tournier (1967) shows that this construction of reality remains fragile. In a sensitive remake of the classic tale of Robinson Crusoe he relates how a man, living alone for decades, is desperately trying to save 'reality' while his perception and subsequent 'naming' of even the most plain things is no longer sustained, affirmed and corrected by others. Alternatively he is minister and repenting believer. As the mayor he sets the rules, but he also takes care to behave as a citizen who trespasses the law. By inflicting punishment on himself he makes the rules very real. The novel thus illustrates how the constant checking of our sensations by means of language and the attribution of meaning is a social, consensual process.

For the 'objective' perception of ordinary things we need little confirmation, but even then the danger of destruction of our common ways of observing or experiencing looms under conditions of social isolation. In Sartre's 'La Nausée' (1938) the main character feels revolted and alienated when he is no longer able to create a social, objective experience from the sensations of seeing and holding a stone while he sits at the river-side. The development and maintenance of ideas and convictions, however, usually need stronger and more lasting reciprocal support. The more esoteric a certain knowledge or observation, the more restricted the consensual domain and we may thus find small clusters or niches of solid consensus like within small religious communities, as well as in subregions of specialized professional fields, notably within the social sciences.

3 Consensus in philosophy of science, sociology and psychology

This chapter will deal with current theories and models on consensus and present research findings from different fields. This limited review of the literature may accomplish the following:

First, I have stated already in chapter 1 that, to my knowledge, consensus has not been studied before in psychotherapy research. In addition to the general remarks of chapter 1, the relevance of such a study may be reinforced by showing the results of consensus research in settings different from psychotherapy.

Second, consensus is still a relatively neglected area of social scientific research. With Klapp (1957), whose ideas will be reviewed in detail, I deplore this since I deem it a very useful and basic concept. Although limited in scope and depth, the subsequent overview may stimulate further research on consensus in psychiatry.

Third, virtually all research that I came across concerns social settings where coordination and proper communication between members are highly important. Also, all of them assume, openly or in a veiled way, that consensus, communication, and social coordination or integration are closely tied. The main reason why most consensus research is done with the family and the marital pair may be that these are relatively stable social settings with these characteristics and rather accessible for research. It is obvious that communication and coordination are also vital to psychotherapy. It therefore follows naturally that the investigation of consensus in the therapeutic setting may be valuable as well.

The different theories and models introduced here foremost testify to the wide applicability of the concept. The sociological model of consensus by Scheff (1967) will be dealt with extensively, since it is the most definite on the relationships between consensus and coordination. A second reason to use this model in my own research is that it also lends itself best to operationalizations that can be used in the psychotherapeutic setting. I have restrained from drawing general conclusions on the research in relation to psychotherapy, nor have I pointed out which studies and research findings I deemed particularly relevant for the study of consensus in psychotherapy. At this stage, my own research only aimed to explore the general viability of such a study. Models and research presented here may however guide more specific interests in the future.

The overview that follows will show that consensus is not a unitary construct. It subsumes distinctive aspects, although partly related according to

the co-orientation models to be discussed. As a last preliminary remark, I therefore have to warn the reader that the presentation of research findings may be somewhat confusing. Because most research dealt with several aspects at a time, an arrangement according to discipline and setting was the only way to structure the survey.

3.1 Consensus in philosophy of science

In the book '*Rational consensus in science and society*', Lehrer and Wagner (1981) defend the form of consensus that is '*a rational amalgamation of the information individuals possess*'. They argue for a consensual basis of knowledge by contrasting it to Cartesian epistemology. According to Descartes we should start with sure, unquestionable propositions, clear and self-evident, without any chance of mistake. Thus, faulty or doubtful premises can be dispelled and science be built on rocklike certainty. But if we limit ourselves to that degree of certainty, we bog down in skepticism. In the same vein, the hypothetico-deductive method pictures science too simply. The idea is: from a theory T we deduce an observational consequence, O. The decision to maintain or reject T depends on the assessment of the truth of O. In this way the scientific method permits falsification, the rejection of a theory or hypothesis through testing. The inductionists' view, however, is that a theory can only be *endorsed* through observations and never be fully falsified. The conclusion of the dilemma is that scientists can - and in practice do - decide to reject a theory, an observation or the rules of inference that led from T to O. This means: no rocklike certainty, but three uncertainties or probabilities. The decision on assignment of probabilities depends on the consensual process of those involved.

The book deals with the rational ways to establish and assess such a process in science, but is also relevant to the fields of morals and politics, law, social policy and language. The central thesis is that the experts, involved in an undecided issue, try to reach a common decision in a well-informed manner and in mutual respect. To reach this goal, each scholar gives his opinion on the matter, assigns a weight to it and to the opinions of all relevant colleagues. The weights represent judgments and these are structured hierarchically in different rounds of assignment:

- 1 Judgement on the suitability of colleagues to give an opinion on the issue;
- 2 Judgement of colleagues regarding their ability to pass judgement on colleagues;
- 3 Judgement of colleagues as judges of judges.

In each round the judgements of others influence the judgement, i.e. weight-assignment, of the judge himself. If, for instance, an expert assigns a high weight to himself and a low weight to all others in round 1 and the

others assign a low weight to his opinion-giving suitability, this results in a correction of the weight of his opinion. In this manner an overall *consensual probability* gradually emerges that stands for the - at that moment - most probably true and correct decision regarding the issue. This method of reaching and assessing consensus is further elaborated in a mathematical model. The authors deal with consensus exclusively in the sense of 'agreement'. The rounds of weight-assignment result in corrected weights that are an amalgamation of judgements of several persons. This means that the weighing process amounts to something akin to 'understanding' and 'realization' (Scheff, 1967).

A related contribution to basic theory on consensus is Farrell's (1976) study of consensus in relationship to rhetorics and knowledge. Since Aristotle, rhetorics relies on 'social knowledge' that connotes, following Habermas, a field of collectively accepted knowledge concerning issues of public interest. Pleading for a renewed recognition of this kind of knowledge, Habermas (1975) states that the twentieth century witnessed ever more restrictive theories of knowledge aiming at the exclusion of practical questions from the discourse on truth and knowledge. *'An explosion of 'information' with a corresponding decline in public dialogue, seemed the paradoxical implication'* (Farrell, p. 3).

Still, all knowledge is founded on some kind of human consensus. Even the scientist cannot do without an exchange on criteria of truth and proof, 'his allegiance ... is towards the creation of a consensus' (Ziman, 1968; p. 78). His analytic rigor and synthetic precision vary in relation to two interdependent factors:

- 1 The degree of actual consensus on methods of investigation, procedures of analysis and operations of measurement.
- 2 The knower's degree of detachment from human interests related to the object of knowledge (Farrell, 1976; p. 4).

If one or both of these criteria are deficient, scientific argument becomes a rhetoric dispute that presupposes social knowledge. Specialized, technical knowledge results from *actual* consensus on methods and procedures. Social knowledge *presumes* that consensus or attributes it to the audience in order that the argumentation can result in the advocated choice or manner of conduct.

In line with Farrell we might add that social science in many respects relies on the kind of knowledge formulated by Aristotle and Habermas, among others. The two criteria for rigor and precision are largely lacking, and the practical intent of social science belongs in this case to the field of social knowledge. The classic objectives of science - description, prediction and control - have limited value, especially the two last ones. Emphasizing prediction and control - in so far as possible at all - would reduce the reality under study, i.e. our fellow men in their relationships and interactions, to an alienated reality to be manipulated, and detached from ourselves as researchers.

The persons investigated would be cut down to determined, i.e. non-living, mechanisms. A general, broad objective that implicitly motivates and guides social scientists can be described as a search for keys to the furthering of socially desirable actions or developments and the tracing of factors or situations that promote or hamper improvement of the human condition. What is desirable and what human improvement means is more ambiguous than the prediction and control of natural phenomena that have been made - rightly or wrongly - to objects outside of ourselves. The definition of the terms 'desirable' and 'improvement' requires the sort of knowledge that is 'attributed, audience dependent, potential in state, generative and normative in implication' (Farrell, 1976; p. 11). In pointing to the importance of social knowledge, the author correctly states that '*... Any sophisticated social system will be confronted throughout its existence by serious problems which require careful deliberation and concerted action. Imagine a society in which the knowledge required to deal with such problems is absent or confined to narrow quarters. Such a system may be a collectivity of individuals; but is far from a community of persons. The overarching function of social knowledge is to transform the society into a community*' (ibid, p. 11).

Where we are scientifically involved with ourselves as the object of knowledge, we include ourselves as the observed and cannot detach ourselves from the social-cultural context and its norms and values, which are constantly generated and adjusted as social knowledge by society in a consensual process of which we are part.

3.2 Consensus in sociology

Over 30 years ago Klapp (1957) gave a still valuable and convincing argument for consensus as a unifying concept in sociology. He reminds the reader of a similar plea by Louis Wirth in 1947 and by Park and Burgess still 30 years earlier in their textbook 'Introduction to the Science of Sociology'. In this book consensus was the central concept of their theory, indebted to seminal work by Cooley, Dewey and Durkheim, who conceived of communicating processes producing consensual like-mindedness as the condition for social differentiation, interdependence and cooperation. Klapp points out (p. 337): '*What is surprising is that in the decades following, sociology has made so little use of this concept supposedly so central to its theory*'.

This has not changed much since 1957. Klapp continues with presenting some studies that show that consensual relationships are different from highly cohesive, symbiotic relationships, that consensus entails role differentiation and promotes organization and development of a group. He further shows that theorists like Parsons and Shils (1951), Rose (1954) and Newcomb (1955)

actively avoid the term consensus and use concepts like solidarity, 'internalized meanings and values ... which permit predictions of another's behavior' and 'symmetry of co-orientation'. For this 'apparent neglect' of the term consensus he offers as an explanation that 'some sociologists ... deny the importance of communications ... as central data', that consensus is hard to observe and measure and that other concepts have taken its place. He refers to a lot of more specific concepts including culture and tradition, rules, codes and norms, '*persuasion, propaganda, decision-making, discussion, compromise, and the like as efforts to reach or create consensus*' (p. 340), and to role, status and self as consensual aspects of personality. He likewise mentions '*frame of reference and universe of discourse as a framework of consensus within which communication and role-playing occur*' (p. 340). He ventures that all of these and many more terms 'add up to the same thing as consensus' (p. 340). Klapp believes that the loss of the word consensus should not be considered simply as a matter of 'a rose by other names' and '*that consensus should have an importance in sociology comparable to that of energy in physics - namely as a unifying concept, [like] ... light, heat, sound and electromagnetism are forms of energy; so, I think, culture, structure, norm, role, symbol, and so on, should be treated as forms of consensus*' (p. 340-341).

Apart from the theoretical work to be done, he points to the need for '*measurement and validation of the kinds and degrees of consensus*' (p. 341). It is especially this last problem that still concerns us and that will be the leading topic in the review of more recent literature and also in our own contribution. Carrying the physical - science metaphor a little further, it can be concluded that solving the problem of measurement adequately would be a step toward a general model that should hold, and consequently can be tested, in instances where some of the specific concepts referred to are used now.

To Klapp's admonishment I would like to add that in present times the importance of the idea of coordination of actions, of co-orientation and therefore consensus is obvious. There seems to be a growing awareness of the importance of consultation, cooperation, interdependence, participation and the like, as opposed to attaining goals and decisions through power, hierarchy, authority and force. The shifting character of psychotherapy and psychiatric treatment toward responsibility and participation of the patient mirrors a general tendency in Western culture. In social science - as a culturally bound phenomenon - a counterpart of such social development would be the study of kinds and degrees of consensus.

Before proceeding with the research literature I would also like to point out that highlighting the relevance of the concept may obscure an important feature of consensus. I herewith refer to the fact of the constant change in *content*, that is the variability and transience of the topics about which consensus tends to exist between individuals or within groups and the corresponding

variations in individuals and groups that exhibit consensus at any given moment. Our main interest is in the continuity of the dynamic *process* of consensus seeking. This allusion to the feature of variability and transience is simply meant to preclude the idea that consensus could be seen as static or constant.

3.2.1 The co-orientation model of Scheff

Since we have used it partly for our own research, the consensus model by Scheff (1967) will now first be reviewed in detail. An individual agreement-definition of consensus, as in most public opinion-research, is of limited use. The usual assumption is that consensus co-determines behavior. With low consensus, in the sense of agreement, on a specific norm, that norm would not be in effect. But in a state of 'pluralistic ignorance'^{*}, when everybody *thinks* that all others agree, such a norm may be just as influential as with high agreement. Scheff therefore adheres to the so-called co-orientation definition of consensus, based on earlier work by interactionist social psychologists like Mead (1934) and Newcomb (1966) and by balance- or consistency theorists like Heider and Festinger (1962). This definition implies a social system model of consensus. The level of perceiving and conceptualizing is at the level of relationships and interactions, not at the individual level as in the first definition.

Co-orientation is achieved through communication and, therefore, consensus and communication are interdependent. communication leads to consensus and consensus makes communication possible. In the interactionist's view, communication is a collective process, not a conglomerate of individual 'sender-receiver' processes. Human communication is an 'interpenetration of perspectives' (Dewey, 1938). Individuals share - at least for the moment - a part of each other's point of view. communication also involves 'mind reading', a mental activity likened to role-taking behavior (Mead, 1934) and necessary to reach intersubjectivity: the shared consciousness of communicating individuals (Schutz, 1962). communication is, however, not just a goal in itself, but the vehicle for regular and reliable interactions that are basic to cooperation and other forms of reciprocal, socially integrated behaviors (like in psychotherapy) that are summed up in the term social coordination. The relationship of coordination to consensus is again stressed in Schelling's (1963) basic proposition that the necessity of social coordination gives rise to the seeking of consensus (through communication). If actors do not desire to coordinate their activities,

^{*} Coined by Allport (1924) as 'the phenomenon that individuals underestimate the degree to which others share their beliefs and sentiments or hold unwarranted assumptions about the thoughts, feelings and behavior of other people'.

no seeking of consensus will occur. *'The theory fails to deal with situations where motivation (for coordination) is low or absent'* (Scheff, p. 35; parentheses added).

Co-orientation is a multi-leveled phenomenon. *Agreement* is called the zero-level of co-orientation: the two (or more) individuals agree on a certain statement X. *Understanding*, the first level of co-orientation, means that person A is aware of the opinion of person B regarding issue X, and B of A's opinion. The second level is called *realization*: person A knows that B is aware of his (A's) opinion. Schelling (1962) mentions that there are still higher levels of mutual understandings and that these 'collective representations' - a term of Durkheim - or the ensuing reciprocity of perspectives (Schutz, 1962) create strong constraints on our ways of perceiving from which we cannot easily and unilaterally detach ourselves by our own force or will-power. Moreover, we are not capable of examining critically level by level these co-orientations in a concrete situation.

The foregoing leads Scheff to a formal definition of a social-system model of consensus which says that *'complete consensus on an issue exists in a group when there is an infinite series of reciprocating understandings between the members of the group concerning the issue'*. The terms for the first three levels of co-orientation and their operationalizations are from Laing et al. (1966) who developed a technique for measuring interpersonal perception among marital partners. The shorthand notation for characterizing the profile of mutuality between the marital pair is A for agreement, D for disagreement, U for understanding, M for misunderstanding, R for realization and F for failure to realize. The profile (RUAUR) thus means that husband and wife agree and that each understands the other and realizes that (s)he is understood. Similarly (FMDMF) means that husband and wife disagree, that each misunderstands and fails to realize that (s)he is misunderstood. The profile husband(FMAUR)wife is indicative of unilateral understanding and realization by the wife. Husband and wife agree, but the husband thinks that they disagree and fails to realize that he is understood. The wife, however, understands that they agree, and realizes that she is misunderstood.

Extending this approach from a dyad to a group raises the question at what degree of shared opinions consensus should be defined. Scheff proposes to fix that degree at the simple majority: if a majority agrees with an issue there is agreement, if a majority understands that there is agreement or disagreement, there is understanding, and so on. Although the problem of assessing consensus in a group (i.e. more than two persons) did not occur in our own research, we will go into this issue after the presentation of some other models of consensus to formulate an alternative approach for measurement and decisions on the question of degree.

3.2.2 Balance model in social psychology

The consistency model or balance model, developed by interactionist social psychologists like Newcomb, Heider and Festinger (1962), has been worked out for consensus by White (1982) and has often been used in research.

Newcomb, in describing the development of college-friendships, was the first to distinguish the affective and the cognitive sides of the relationship. This enabled him to delineate which relationships evolved and endured. The basic principle of the model is that perceptions - i.e. opinions, beliefs - of a person are consistent or in balance with the person's affective attitude toward the other person with whom those beliefs or opinions are shared. This implies that, in the case of a positive affective stance of person A toward person B and a positive (i.e. agreeing) cognition of A regarding statement X (e.g.: 'violence on TV has a bad influence on children'), person A will tend to suppose a same positive cognition about X with person B. Person A is likewise consistent or in balance if he dislikes B and supposes disagreement with him and, when liking B, supposes that both of them disagree with (reject) statement X. As both persons will strive for consistency, communication will serve to persuade each other in the case that one or both persons are in imbalance.

The following diagram serves to make clear the principle of balance-theory (see fig. 1-3).

Figure 1: affective-cognitive triangle for person A

Figure 2: affective-cognitive triangle for person B

Figure 3: affective-cognitive triangle for the dyad A-B

When the product of signs of the triangle sides is positive, the person is in balance. When the product is negative, the person is in imbalance. The same rule applies to dyadic structures. In the example illustrated in fig. 1-3, both A and B are individually in a balanced state. A sees his relationship with B as positive (liking, acceptance) and he supposes B's opinion to be the same as his own, namely positive (agreeing, endorsing). B however sees his relationship with A as negative (disliking, rejecting) and he perceives A's opinion to be at odds with his own. As a dyad they are in balance as well.

The dyadic structure - the product of the signs of the corresponding sides of the two individual triangles - shows that A perceives B incorrectly ($B \rightarrow X: -$), that B perceives A correctly ($A \rightarrow X: +$) and that they perceive their relationship differently. Different perceptions are indicated with a negative sign, corresponding perceptions with a positive sign. The example in fig. 1-3 is one out of many possible configurations. Proceeding from eight individual orientation systems for each person, there are $8 \times 8 = 64$ dyadic combinations which possess one of the six possible dyadic structures (see figure 4). The structures

$+\triangle -$ and $-\triangle +$ and the structures $-\triangle +$ and $+\triangle -$

are equivalent on the dyadic level.

Figure 4: The six dyadic structures in the ABX-diagram

Numbers 1, 2 and 3 are balanced (positive product of triangle-sides), numbers 4, 5 and 6 are imbalanced.

Balance theory now presumes that a person experiencing inconsistency (e.g. $+\triangle -$) has the choice between either winning over his partner so that he will also perceive X positively, or changing his own opinion on X or revising his view on the relationship. If however the definition of the relationship is mutually positive and therefore congruent, the two first possibilities are more obvious. This amounts to that such a dyad, when cognitive inconsistency has become apparent, will strive for more agreement. Also,

balance theory predicts a blocking of communication over controversial issues to keep the relationship positive without having to experience inconsistency. An example is depicted in figure 5.

Figure 5: Specimen of individual and dyadic balance with a positive relationship based on blocked communication

It also follows that holding a negative definition of the relationship favours the perception of cognitive incongruity (figure 6).

Figure 6: Balanced individual orientations allowing the perception of cognitive dissimilarity

With a positive affective attitude toward the other person, the perception of disagreement creates an imbalanced state. Avoidance of such a condition is recognized in the expression that 'love blinds' for differences in outlook. Also, if two persons both define the relationship negatively and perceive disagreement, no matter if this is correct or false, the dyad is in a balanced state (fig. 7) that may easily remain stable.

Figure 7: Balanced dyadic structures based on balanced individual orientations with dissimilar cognitive perceptions

Wackman (1973), however, assumes that the main function of communication is the exchange of information, thus decreasing the uncertainty regarding the cognitions of the other party. Wackman has founded his statement on studies in which agreement, congruency and accuracy were measured simultaneously. These showed the '*consistent result: the major outcome of interpersonal communication was increased accuracy, rather than increased agreement or increased congruency*' (p. 540). He also notes that most communication research has dealt with factors affecting persuasion and not variables that influence information exchange. This may be the reason the above results often have not been reported.

If information exchange has occurred, this will show in a change of state in the individual(s) and therefore in the dyad. To return to the example of figure 5 this exchange could lead to a change in both individual orientation systems to a state of $\begin{smallmatrix} + & \triangle & - \\ + & & \end{smallmatrix}$. Both are individually imbalanced or inconsistent, but well informed about the perception of the other. The dyadic structure will have been shifted to $\begin{smallmatrix} + & \triangle & + \\ + & & \end{smallmatrix}$. This is a more positive state that is better 'informed' and should therefore be more stable than $\begin{smallmatrix} - & \triangle & - \\ + & & \end{smallmatrix}$.

White (1982) has made predictions about changes in dyadic structures starting from different combinations of individual consistent and inconsistent orientation systems. His predictions were based on axioms of information theory and tested in a first degree Markov-chain model. In his pilot study of marital communication his predictions were only partly valid. It turned out that balanced but partially informed dyadic structures like $\begin{smallmatrix} - & \triangle & + \\ + & & \end{smallmatrix}$ recurred more often than was

predicted by the model. This might signify that such structures do not show transfer of information or that simultaneous but contrary change in individual orientations block a dyadic modification.

White concludes that an extension of the axioms is needed and that communicative interactions with a combination of negative affect - mostly communicated through analogue, non-verbal signals - and negative (rejecting) cognitions of one or both partners interfere strongly with the transfer of information and therefore block changes of dyadic structure. With a mutually positive affect, information exchange is most likely. It is well known that the importance of establishing such a positive bond is stressed in psychotherapy. This emphasis seems well founded in the light of this model.

3.3 Comment on the co-orientation model

One of the main points of Scheff's model is that consensus has clearly been posed as a multi-leveled phenomenon that is to be measured in a relational way. That is, the measure 'understanding' consists of a combination of a person's opinion and the other person's guess of that opinion. Agreement, understanding and realization are clearly related at the *conceptual* level (see chapter 5). But we do not know whether the different aspects of consensus come into being simultaneously or successively.

It might be possible, for instance, that under certain conditions consensus starts with some mutual understanding, that agreement develops on that basis, which gives rise to further understanding and a beginning of realization. Testing hypotheses about the evolution of the consensual process requires repeated tests over time. Spillman et al. (1979; see page 33) have devised a manner to measure agreement as a process, but unfortunately most operationalizations of consensus are not well-suited for repeated measurement.

In appraising Scheff's model critically, some shortcomings may further be pointed out:

- 1 Consensus can be assessed on different levels, but on only one issue at a time. Although it is Scheff's contention 'to formulate ... an operational definition of *degree* of consensus ...' (p. 32, italics added), his operationalizations only offer dichotomies in different levels.
- 2 The rule for agreement etc. in groups by means of simple majority (or two-third majority or unanimity) is arbitrary.
- 3 In assessing 'understanding' or 'realization', the majorities may be different (i.e. consisting of different persons) from the majority that is in agreement. Alternatively, if one uses the persons who are in agreement to judge the level of understanding and realization, the judgements of a considerable part of the group cannot be counted in the assessment.

- 4 When operationalizing 'realization' according to Scheff's own suggestions, respondents have to be asked confusing questions like: 'how would the average person think that the average person will have answered?'. Likely, the respondent will answer that he has no idea on the matter.

Taken together, these drawbacks call for a different computation method in group situations and another approach to frame questions, which are derived from those used for dyads in our own research.

First, we adopt Scheff's proposal to define agreement in terms of *endorsing* a statement or issue. This means that two persons both disagreeing with (rejecting) a certain statement are not considered to be in agreement. For most practical purposes a 'common negativism' will not be of interest and will falsely inflate the level of agreement. The Jaccard index - the similarity coefficient that we propose for computing consensus (see chapter 8) - takes this point well into account. We also approve of the recommendation of Scheff, and many others, to consider consensus *per issue* as a dichotomous variable.

Regarding the measurement of consensus, I first explicate the first objection. Measurement of a consensus-profile for each issue separately is psychometrically unsound and unrealistic. Consensus on related issues will tend to show cohesion. When investigating opinions on contraceptive methods within the marital pair (Neal & Groat, 1976), political problems, segregation (Fields & Schuman, 1977) or any other multi-faceted topic, it is obvious that the consensus (or dissensus) that really matters is the extent to which people agree and understand each other on the different aspects of the general theme.

3.3.1 Consensus within a group

The last three objections all concern the measurement of consensus at the group level. To neutralize these, the following operationalizations of successive levels of consensus within a group can be proposed.

3.3.1.1 Agreement

For the level of agreement, the individual opinions regarding specific issues are used to compute similarity coefficients like the Jaccard index (Dormaar et al., 1989; Snijders et al., 1989) for all dyads. The main reason is the fact, that agreement is essentially always an intersubjectivity between *two* persons. If all n members of a group agree, this indicates that all $\frac{1}{2}n(n-1)$ dyads agree. Also, for one issue only, if a certain person in a group of n persons agrees with 'a' other persons ($a=1,2,\dots,n-1$), then $\frac{1}{2}a(a+1)$ dyads agree. Another advantage of the computation of agreement with a similarity coefficient like Jaccard is that the proportion of dyads exhibiting real agreement can be defined as having Jaccard-values that are unlikely to be due to chance, e.g. beyond 2 standard

deviations from the mean of the so-called 'independent distribution'. This affords an in-depth investigation of the agreeing and the non-agreeing dyads.

As regards the independent distribution, it has been explained elsewhere (Snijders et al., 1989; see also chapter 8) that the mean and variance of the observed distribution of Jaccard values that represent 'agreement' should be compared to the distribution under statistical independence between members of the dyads. A significant difference between the distributions means that the observed distribution cannot be accounted for by chance and that, therefore, a special factor probably is responsible for it. According to the co-orientation theories presented here, that factor should be communication on the issues between members of the group.

In research on consensus in large anonymous groups the procedure outlined above is not valid, however, since the distribution of the observed consensus-values and the independent distribution will tend to be equal. This is due to the fact that both are derived from the observed frequencies of the issues involved. Instead, a comparison should be based on the specific research question. If, for instance, the research is focused on the differences in consensus between specific subgroups, the distribution of a subgroup (say, women or members of some political party) can be compared with the distribution of another subgroup or an aselect population-sample. In research on the influence of communication or information-exchange on agreement it would be appropriate to compare distributions before and after a communication experiment like a publicity campaign or a public debate.

3.3.1.2 Understanding

The hypotheses of the research in question also determine the computation of understanding, the phrasing of questions and fixing the cut-off level. Research into group-consensus may need, for instance, the distributions of persons and issues that show understanding. In that case a dataset of judges by targets may be made to perform an analysis of variance as presented by Kenny & Albright (1987) for accuracy-research. The judges are the individual group members, the targets are the items and the judgement can be scored positive for understanding when it is within a certain limit (e.g. one or two standard deviations) of the mean of the group opinion. If, however, the researcher is interested in the *persons* that can estimate correctly the majority's opinions, he will have to make a choice - as in Scheff's case - about the percentage-level he defines as majority. In the phrasing of his questions he should omit the confusing words 'majority' or 'average person'. He may ask the respondents directly to state the percentage of the group presumed to agree with each issue. The cut-off point where respondents suppose the endorsement of a majority or minority remains in the hands of the experimenter.

This research-strategy will be illustrated with a proposed manner of computation. As an example, suppose that in a certain public opinion research Jaccard-values per respondent (group-member) have been computed over the issues that are under consideration. This computation is performed by using the estimates of the respondent and the group's actual opinions to fill out the Jaccard-matrix. A concrete case might be as follows:

- Number of issues: 20.
- Cut-off for majority-definition: over 50% of all group members agree. For 12 issues the respondent 'A' expected an endorsement by 50% or more.
- For 10 of these his expectation was correct.
- The total number of issues that had at least 50% endorsement in the group was 14.

table 3.1: Contingency table for understanding of a group-member.

A's estimation of group-member's opinions	actual group opinions		
	over 50% agree	under 50% agree	
over 50% agree	10 (a)	2 (b)	12
under 50% agree	4 (c)	4 (d)	8
	14	6	20

Jaccard-value of understanding for this group member:

$$J = \frac{a}{a+b+c} = \frac{10}{10+2+4} = 0.625$$

Hereafter the understanding by group members can be compared to a distribution of understanding under suitable conditions of independence, in the same way as discussed in Snijders et al. (1989) for agreement. In this way the subgroup of 'good understanders' can be assessed. The binomial distribution of the ratio good versus poor understanders per issue can elucidate the issues about which the group members are well-informed regarding their fellow group member's positions.

3.3.1.3 Realization

Questions on the level of 'realization' in a group remain problematic. Realization on the group-level requires an ability of putting oneself in other people's position that may well be beyond conscious articulation for most respondents. An acceptable formulation could be as follows:

Issue: 'The costs of keeping our environment clean from now on, should be largely paid for by general tax-revenues'.

- a Agree/disagree (question to assess agreement);
- b What percentage of the group do you suppose to endorse this statement? (question to assess understanding).

We now would like to know what part of the group you think is aware of the group's opinion on this issue.

First you have stated your own opinion, secondly you answered what percentage of the group you suppose to agree with this issue. Call this percentage 'B'. If your own first answer was 'agree', then you also suppose a percentage 'B' to agree with you. If your first answer was 'disagree', then you expect a percentage of '100-B' to take your position.

Now, suppose you're right that indeed a percentage 'B' (resp. 100-B) agrees with this issue. Our third question now is:

- c What percentage of the group do you expect to judge that at least that percentage of the group shares the same opinion with you?

To compute realization over many issues with a Jaccard-index, the researcher has to take similar decisions as with understanding to dichotomize the raw data of percentages. He may want to know 'Does the respondent realize that a majority understands that a majority endorses this issue?'. In that case he can make a complex matrix (see table 3.2) to place the respondents' answers to the respective issues. In this way, again, a distribution of Jaccard values of respondent - group 'dyads' can be made.

It is unlikely that research-situations on this level of abstraction are desirable. It can be conceived, however, that it is important for strategic politicians to really know if ideas are correctly judged by others - e.g. the electorate or colleagues. His faulty judgements on this level like 'they certainly do know how I feel about these topics' or 'I'm glad that they do not really know what I have in mind' may have serious consequences.

We hope that our suggestions about how to measure the different levels of consensus in group-situations with a similarity coefficient like Jaccard may stimulate further research. The alternative operationalizations of consensus with correlation coefficients and analysis of variance techniques will be discussed later on.

table 3.2: Assignment of issues per respondent and group estimations of opinions

res- pondent's estimation	group's estima- tions	over 50% supposes that over 50% agrees		over 50% supposes that less than 50% agrees	
		over 50% agrees (=correct)	less than 50% agrees (=incorrect)	over 50% agrees (=incorrect)	less than 50% agrees (=correct)
respondent thinks that over 50% supposes that over 50% agree	a (a/a)	1	2	3	4 (a/b)
		a	b	c	d
respondent thinks that less than 50% supposes that over 50% agrees	c (c/a)	9	10	11	12 (c/b)
respondent thinks that less than 50% supposes that 50% agrees	(a/c)	17	18	19	20 (a/d)
respondent thinks that less than 50% supposes that less than 50% agrees	(c/c)	25	26	27	28 (c/d)

The characters in the cells in the 2x2 matrix of the Jaccard-matrix:
realization can be computed from the issues that fall into cells
1, 5, 9 and 13. Alternatively, one might decide to include
the issues that belong in cells 20, 24, 28 and 32

$\frac{a}{c} \frac{b}{d}$

to compute Jaccard-indices of realization

less 50% supposes that over 50% agrees		less 50% supposes that less than 50% agrees	
over 50% agrees (=correct)	less than 50% agrees (=incorrect)	over 50% agrees (=incorrect)	less than 50% agrees (=correct)
b (b/a)	5	6	7
			8 (b/b)
	a	b	c
d (d/a)	13	14	15
			16 (d/b)
	21	22	23
			24 (b/d)
	29	30	31
(d/c)			32 (d/d)

3.4 Consensus in psychology

In most of psychology the tradition of research on consensus differs considerably from sociology and interactionistic social psychology. In sociology and related fields consensus is conceived as a condition between two or more persons concerning issues outside the individuals involved, like opinions, attitudes and values. In personality research the emphasis is on 'person perception', on how people perceive each other. It can be rather confusing that the same terms, like consensus and accuracy are sometimes used for different purposes while on the other hand different terms are employed for the same themes. Before going into some of the psychological research on consensus we therefore first present the main terms for the aspects of consensus with their equivalents (see table 3.2).

Table 3.3: Main terms for levels of consensus in co-orientation theory and their equivalents.

Main term		Equivalent	
agreement	Scheff, 1967 Laing et al, 1966	consensus	(Lehrer & Wagner, 1981; Marks & Miller, 1987)
		cognitive overlap	(Wackman, 1973)
understanding	Laing et al., 1966 1987) 1987)	accuracy	McLeod & Chaffee, 1973; Kenny & Albright, 1987)
		sensitivity	vide Kenny & Albright,
		empathy	vide Kenny & Albright,
congruence	Wackman, 1973	perceived similarity	White, 1985
		assumed similarity	Marks & Miller, 1987
		self-generated comparison information	Sanders & Mullen, 1983)
realization	Laing et al, 1966		

The two mainstreams in 'person-perception' studies regarding consensus are on *accuracy* and *false consensus*. About the first topic Kenny and Abright (1987) have recently presented a 'nomothetic, interpersonal and componential' model for psychological accuracy research. Here, assessing accuracy entails a judgement process in which '... each judge rates a set of targets on a set of traits. For each judgement there is a criterion score. Accuracy is defined as the correspondence between the judgement and the criterion' (ibid, p. 390). Judgement and criterion can be divided in component parts (constant, trait, target and uniqueness) that can be represented in a two-way analysis of variance in which the rows are traits and the columns are targets. The combinations of these component parts of judgement and criterion yield four components of accuracy [Cronbach (1955) in Kenny and Albright, 1987; p. 391]. The authors extend this approach to the assessment of accuracy in the sense of correspondence between a person's prediction of the other as the judgement and the other's actual behavior as the criterion, as well as to studies of how accurate people are in knowing how others perceive them. Unwittingly, the authors herewith introduce the second order level of consensus, called 'realization' by Laing et al. (1966) and Scheff (1967). This topic of realization has been elaborated upon in an investigation of the correspondence between people's own 'impressions' of themselves and others of the relevant group on personality attributes and the ranks they predict as the group's opinion about themselves and others on these same attributes. This kind of approach is advocated for perception of liking, self-disclosure, lie-detection and pose-emotion studies.

The authors state clearly that their focus is on 'measuring when and how people are accurate and not on who is accurate' (ibid, p. 390). This departure from the original goal 'to differentiate people in their ability to know the social world surrounding them' (ibid, p. 390) has been prompted by serious criticism of Cronbach and Cage that the original measurement technique of discrepancy scores between judgement and criterion is a poor one. The main interest of the current psychological contribution to accuracy research has therefore become the investigation of the factors that influence people's accuracy like judge-target response-set correspondence and the correspondence between the actual behavior of individuals and the general rating by their judges. All four correspondence-factors presented by Kenny and Albright as components of accuracy are operationalized in correlation coefficients. Taken literally, correlation is however a poor definition of correspondence in the sense of *similarity*. Also, the authors do not mention the serious limitation to generalizability of their componential approach: that it can only offer insights into relevant components with respect to each trait separately. It is unrealistic that the components, defined as factors determining people's accuracy, would not have some general validity, instead of being dependent on the trait' - like intelligence or competitiveness - that are used in the assessment.

The second line of consensus research in the psychological domain concerns the so-called false-consensus effects, that is the tendency of people to perceive a too high agreement with their own beliefs, attributes and behaviors. Marks and Miller (1987) discuss the several theories that have been advanced to explain the phenomenon. The results of the 45 studies the authors have found on the false-consensus phenomenon over the last ten years all support one or more of these theoretical views, depending on the specific circumstances of the research. In the context of this introduction the relevance of the thorough review cited above is mainly the demonstration of how 'biases in social perception are affected by numerous factors ... depending on the situation' (ibid, p. 401).

3.5 Consensus and intersubjectivity

As was shown in the co-orientation model of consensus, intersubjectivity is fundamental to consensus. Because of their relevance to psychotherapeutic practice I finally present Scheff's (1973) ideas on the importance of emotions to intersubjectivity. He starts with the question of how words and gestures may have the same meaning for different individuals. This is not just a philosophical problem, but is also tackled in 'person perception' in psychology and in consensus research in social psychology and sociology. The problem of meaning is more complex than a conditioning model can account for, as is apparent in the problem of multiple meanings. A computer translation based on 'dictionary knowledge' of the saying 'out of sight, out of mind', might carry a meaning in Chinese like 'first blind, then insane'. Apart from the ambiguities of alternative meanings the *playfulness to communication - including irony, sarcasm, understatement, euphemism and the like - seems to require attention not just to the message, but to the complex situation in which communication takes place.* (p. 503) The fact that '*all messages are profoundly ambiguous*' (p. 503) leads us to Buber's invocation how from the process of gaining knowledge of the other only confusion ensues. Recalling that intersubjectivity and communication nevertheless do occur, Scheff returns to Schelling's and Mead's emphasis on role-taking as the prime condition for proper understanding of the other and to '*the central place of the recognition of emotion in role-taking*' (p. 506).

First he discusses Tomkin's argument on the expression and recognition of emotions as the direct access to the other person's intentions and studies that support the proposition that emotional expressions are biologically determined and '*make a common language that is universal to the human species*' (p. 508). Returning to verbal ambiguity, it is stated that emotions are '*a set of operators which transform the meaning of verbal statements. If it is true that a message may always mean the exact opposite of its literal meaning, as when an*

exasperated parent says to a plate-dropping child, 'that was really smart', then there must be some nonverbal signal that makes the intended meaning of the message understandable' (p. 508). This reasoning leads to the argument that 'intersubjectivity is based primarily on the expression of emotions' (p. 509), followed by a short account on the importance for the individual of being able to express and react to emotion and the consequences of living in social settings that hamper or encourage this. These ideas propounded by Scheff have been presented rather extensively as they conform so well to psychiatric practice, especially in the form of psychotherapy.

Next to this presentation of several theories and models pertaining to consensus, I will now offer a short review of actual research.

3.6 Consensus research in the sociology of science

Hargens (1975), studying consensus within different scientific disciplines, reasoned that if the profession shows little agreement on methods, relevance of problems, procedures and the like, a Ph.D.-student will have to explain more before he has wound up a dissertation in a scientifically justified way to the satisfaction of his prestigious colleagues. Also, he reasoned that with lower consensus, the proportion of manuscripts rejected by journal editors will be higher. The number of pages of dissertations and the rejection percentages thus should be indirect measures for the degree of consensus in a field of science. Indeed his hypothesis was confirmed (see fig 3.1). Length of dissertations and rejection percentages were much lower for mathematics, physics and other exact sciences than for sociology, political science, history and the like.

The differences in rejection percentages could not be explained by relative lack of space in or paucity of scientific journals. In a further interpretation of scientific consensus Hargens and Hagstrom (1982) advance that a poorer codification, i.e. a little developed formal system - if well developed, of a mathematical character - leads to elaborate language. If, conversely, the standards of proof and argumentation are more univocal, it will take less space to reach a solid conclusion.

With this last study, based on data from probability samples of US graduate faculty in five disciplines (from physics to political science), the authors also found that *'consensus and codification [like in physics and chemistry] make it easier for skilled academic scientists to achieve various kinds of success'*. (p. 194)

Figure 3.1: Consensus in various scientific disciplines

As to career patterns they note, however, that large differences in consensus appear to have a small effect. The authors conclude with the remark that *'...This is not surprising. The logic of discovery in all fields is less codified than the logic of demonstration, and it is difficult to identify and assess creative talent with confidence in even the most codified fields of science'*. (p. 194)

In connection to these studies, a hypothesis could also be forwarded about the number of references per article or page. Citing many authors in the introduction and discussion of an article is a means to establish one's professional identity vis à vis colleagues and serves to create a consensual domain (Maturana, 1988) of truth and reality. To the extent that codification, referred to by Hargens (1975), or rigor and precision, cited from Farrell (1975), within a discipline are less, it can be supposed that the mean number of references will be higher.

3.7 Public opinion research on consensus

The next contribution stems from public opinion research. In the Detroit area study, conducted in 1956, 1969 and 1971, Fields and Schuman (1977) investigated perceptions of racial opinions of 'most Detroiters'. Their 1969 survey research had been directed specifically to estimations by whites of white support for segregation. The 1969 survey also inquired specifically about the perceptions of opinions of fairly well-known others. When distinguishing answers as 'conservative' or 'liberal', over 70% of the respondents reacted liberal to indirectly phrased questions, like what a mother should do if her six-year old daughter asked permission to bring a negro-girl home to play. But also over 70% supposed that their direct social environment - the five closest neighbours - would react conservatively. The researchers could also reliably estimate the opinions of the neighbours, since most were respondents as well. It thus appeared that the assumed reactions of neighbours were much too conservative.

This research exemplifies, although not explicated as such by the authors, the interdependence of agreement (the correspondence of opinions), perceived consensus (the correspondence of own opinion and assumed opinion of other(s)), and understanding (the correspondence between others' real opinions and 'perceived opinions'). The authors conclude that interpersonal perception^{*} shows three sources of error:

- 1 'Looking-glass perception': the tendency to suppose that others believe the same as oneself.
- 2 Conservative bias: the prejudice that others are more conservative than oneself.
- 3 Reality constraints: denoting the difference between reaction to an imaginary situation, like the response to a question, and to a real situation. To make this last point clear: the researchers found that responses to questions, e.g. about an open-housing law, were less conservative than reactions as measured by the number of signatures on a list requesting the implementation of such a law.

An earlier report on 'pluralistic ignorance' (F.H. Allport, in Merton, 1968), also regarding attitudes toward racial segregation, is based on figures of a nationwide national election study conducted in 1968 (O'Gorman, 1973). In a detailed analysis O'Gorman remarks that response bias - i.e. social desirability - cannot explain away the differences between responders' own opinions and their overestimation of white support for racial segregation. Nevertheless, as

^{*} The word "perception" is commonly used for beliefs, opinions, attitudes, convictions in sociological and psychological interpersonal research. See Tagiuri in Fields & Schuman (1977).

already pointed out by Scheff, the data-sets and phrasing of questions preclude an exact assessment of the contribution of response bias-factors and true ignorance.

Another example of the use of the concepts of consensus has been provided by Guppy and Goyder (1984) in their study on the rating of occupations, as commonly applied in socio-economic status assignment. In a reanalysis of survey-data they found that the variation in ratings by individuals, usually interpreted as measurement error, was partly systematic. The results show that the consensus in rating among white and higher status respondents, measured as agreement in the form of average transformed correlation using Hotteling's formula, was higher than among black and lower class respondents. With regression analysis these differences were found to be significant. They conclude that *'... to label as 'random' a form of error which is acknowledged to correlate with socio-economic status is self-contradicting'* (p. 719)

and they object to the psychologizing of error that is a social status phenomenon into a low working class cognitive ability-problem. They therefore advocate the acknowledgement of the error as dissensus instead of dismissing the dissensus as error.

3.8 Sociological research in consensus in families and dyadic relationships

We now turn to research on dyadic interaction that touches specifically on consensus. In most cases, different aspects of consensus have been investigated concomitantly.

An early study on consensus in a psychiatric setting concerns the research by Coe et al. (1969). They have presented data from admitted patients and their families compared to normal families as controls. Mother-father and patient-spouse dyads in the patient-families supposed disagreement for a mere 4% of the 54 items of a questionnaire on family interactions in 4 areas of family-duties. Comparable dyads in the control families supposed disagreement for 28% of the items. The actual percentage of disagreement in the patient-families was however between 40 and 38%, whereas in the control families this was 32%.

In consensus-terms it can be said that dyads in both kinds of families had a comparable level of agreement (60 to 68%), but that dyads from patient-families showed a too high level of congruency and therefore a lower level of understanding than control-dyads. The authors conclude (p. 128) that *'An inability to admit to problems appears to be a characteristic of psychiatric families'* and that *'... in families where disagreement is unacceptable, ... the resentment which accumulates will require some form of release.'*

Many studies on consensus concern the marital dyad, some of which will be reviewed now. First, Schulman (1974, p. 140) reports on some earlier studies that are worth mentioning. Regarding the aspects of agreement and congruency, Kirkpatrick and Hobart, in a 1954-study, found a 'lack of realism in all stages of courtship and early marriage' indicated by a much higher disagreement than the couples estimated their disagreement to be. Hobart found a narrowing of this gap or a 'disillusion' effect in the transition from engagement to marriage, and Pineo discovered a substantial decrease in marital adjustment scores over a period of 20 years. Also, later divorcing men showed a higher adjustment score at engagement than still married men. These results also regard consensus, since a respondent with a high adjustment score testifies to an idealization of the couple's harmony due to the social desirability component of the test. This means that with a high adjustment score (s)he reports higher congruency than the level of agreement that can reasonably be supposed to exist. Turning to accuracy as another aspect of consensus, Schulman (1974, p. 141) mentions a study of Dymond that showed that happiness of married couples was significantly associated with accuracy of prediction scores.

Quite relevant to the still ongoing debate on emancipation is a study by Stuckert (1963). He studied marital satisfaction in terms of similarity of role expectations and accuracy of perception of other's role expectations. He found role similarity to be most important in the satisfaction of husbands and accuracy of the wife's perception to be most important in the satisfaction of wives. Where role similarity is low and his perception of his wife is accurate, the husband is often dissatisfied. However, *husbands are satisfied when their role expectations are different from their wives, but they do not perceive this.*

We call to mind our earlier statement on the likeness of the more colloquial words 'feelings' or 'emotions' and the more technical terms 'definition of the relationship' and thus 'role expectations'. Stated bluntly, these results therefore can be interpreted that men are feeling satisfied in their marriage if they perceive congruency in feelings, even when incorrect. Women, on the other hand, are capable of feeling happy while being aware of their partners' feelings, even if deviant from their own. Taking into account studies of Calonico and Thomas (1973) and Miller (1977), to be reviewed later, this investigation would be worth replicating with current 'old' and 'new'-style couples. Emancipated men can be supposed to fare better in tolerating imbalance - i.e. incongruency - than their more conventional predecessors.

Schulman's (1974) own research focused on idealization and romanticism among engaged couples. One of her hypotheses on the relationship between potential conflict and perceptual distortion stated that in areas of potential conflict - i.e. disagreement - there will be distortion of each other's views and feelings - i.e. inaccuracy. This hypothesis was confirmed with a highly significant correlation coefficient of .70. A problem with the use of correlation

coefficients is the inherent interdependence of the variables correlated with each other. A possible solution to this problem will be presented later-on. Still, not being aware of disagreements certainly appears to be a way of avoiding conflict that unfortunately also blocks communication. This is consistent with the findings of Laing et al. (1966) that couples misunderstood more often on issues on which they disagreed, which led these authors to wonder if ... *people feel they agree even when they do not because they never probe or make explicit issues when disagreement may exist?*' (p. 85).

A related study with married instead of engaged couples is Bloom's (1984) study on consensus during the transition to parenthood. She investigated the concerns and relationship dynamics during the prenatal-to-postpartum transition of 22 couples expecting their first-born child. It was found that, prenatally, both husbands and wives perceived significantly more agreement than actually existed; post-partum this remained true for the husbands only. The 'disillusion-effect', mentioned earlier, apparently worked more for the wives than their husbands. However, both prenatally and postpartum, husbands were significantly more accurate in their estimation of their wives' views on the importance of concerns. She also found that issues ranked as most important were also those discussed most. It should be noted that the consensus variables, agreement, perceived agreement (i.e. congruency) and accuracy, were operationalized with Spearman correlation coefficients with a rank-order of importance. This may have entailed undue emphasis on the *importance*, in contrast to the *number*, of the issues on which consensus existed. Consensus, if conceived as correspondence or similarity, should not be measured with correlations, as has been argued by Gregson (1975, p. 52) for similarity measures in general.

The same flaw in operationalization of consensus is present in the large and interesting study of Neal and Groat (1976), who measured various aspects of consensus in the marital dyad as correlation coefficients over the total number ($n = 365$) of couples. Apart from methodological objections, this procedure only enables one to identify the issues on which consensus exists. It impedes the identification of the specific couples in consensus or dissensus over several issues.

Jaco and Shepard (1975) remark that in survey research on family matters, married wives are predominantly interviewed. This might indicate that researchers assume consensus within the marital pair. Interest in the conditions under which this assumption might be valid and the idea that consensus on survey questions might depend on type of information elicited prompted their research into the relationship between consensus and *demographic homogeneity* within the marital dyad and secondly into the impact of knowledge vs. attitudinal questions on this potential relationship. They were unable to find a general relation as predicted. Higher levels of education tended to be asso-

ciated with higher levels of consensus, both on knowledge (facts of sex and fertility) and attitudinal items (on family planning). Only religious homogeneity - e.g. both spouses Catholic - showed a consistently higher level of consensus than mixed affiliations. Place of origin and age had no clear effects on consensus. The authors themselves have mentioned several methodological limitations to their study that might account for the lack of support for their hypotheses. Additionally, I note that consensus has only been operationalized as *agreement*. In view of other research, results might have been different if understanding (accuracy) and congruency had been included.

A related study by MacKinnon and Summers (1976) investigated the extent to which *homogeneity in social background* can account for role consensus among teachers and adult students. Consensus was defined as agreement in the sense of similarity of response and measured indirectly by testing the generalized multivariate null-hypothesis of similarity in role expectations between groups. The research was based on solid theoretical and research evidence that will not be reviewed here. The strength of this study is its clear formulation of a null-hypothesis and the thoughtful testing with advanced techniques (principle components factor analysis and multiple discriminant analysis). The findings supported the hypothesized relation between homogeneity and role consensus. There was generally greater role dissensus for student than for teacher behavior. Age, sex, socio-economic and marital status, apart and in interaction, were all aspects of possible homogeneity that were found to be related to role consensus (agreement).

White (1985), in a thoughtful study on consensus in married couples, first relates common assumptions on marriage that are fundamental to consensus research in this area. Phenomenologically, marriage is seen as a process of constructing and legitimating a shared reality; the individual perspectives are thought to be gradually reconstructed. The resulting reciprocity of perspectives, being an intentional activity of both spouses, would imply a high correlation between their perceptions of similarity. In the same vein, understanding - the ability to accurately predict the stance of the spouse - would be high where couples have achieved this 'biographical fusion'. Consequently, perception of similarity - i.e. congruence - and understanding or accuracy should be correlated. A communication and co-orientation perspective like that forwarded by Scheff (1967) and the consistency theory of Heider lead to the same prediction that understanding and perceived similarity are correlated. It is questionable, however, whether such global predictions are valid. An important contribution of White's research - derived from a probability sample and consisting of 87 married couples - is his probing of the *salience* of issues. At different stages of marriage certain issues will probably be more important than others: for a couple without children it may be leisure time pursuits, later on preferred family size or economic issues.

Indeed, on a global level his results do not show the predicted correlation of congruence and understanding. Still, the husband's perceived similarity is closer to the actual dyadic agreement than that of his wife, who perceives a too high agreement. For highly salient - i.e. important - issues the wife's perception is close to that of the husband. Nevertheless, husband's congruence cannot predict understanding by himself or his wife. In understanding the wife is superior, especially for high salience issues. Overall, there was clear agreement on which issues were important.

These confusing results, which contradict theory, may at least partly be attributed to the method employed: For all consensus measures difference scores were used. Objections to difference scores are well-known (Cronbach & Furby, 1970; Fisher et al., 1985). More serious however are the implications of the consensus-paradigm used. The fact that the consensus-scores in White's research, and actually in most cases cited, were difference-scores means that the persons in the dyad - like husband and wife - are correlated by means of subtraction of their scores per item, the value of agree-agree being zero and of agree-disagree one for each item. This procedure presumes the items to be a kind of 'test-items', i.e. to be homogeneous. No reliability coefficient has been supplied and if calculated may be assumed to be low since the items - although all in the domain of 'family concerns' - cover divergent subjects like family-size, leisure time pursuits and women's rights. The statistics used - ANOVA, t-tests and correlation coefficients - are also inappropriate. They lump together the possibly divergent patterns of all dyads. The use of means and variances per level of consensus does not allow to discern which - perhaps significant - number of dyads show the predicted relationship of congruence and understanding.

It should therefore be preferred to use a similarity paradigm for consensus, which is more of a Q-methodology and rather alien to the R-modes commonly employed. This implies that the items should be regarded as Q-sort items with a dichotomous sorting instruction (agree-disagree). The Q-sort may be structured with a partitioning into knowledge items and attitudinal items, in the case of Jaco and Shepard (1975) referred to above, or in salient and non-salient in the case of White (1985). The correlation, or preferably 'association', between the two persons of the dyad is consequently made by a similarity coefficient, since consensus signifies similarity of response.

Calonico and Thomas (1973) have attempted to account for individual role-taking ability through an analysis of differing dyadic characteristics in the nuclear family. Their hypotheses on the relationships between affect, value similarity and role-taking ability have been based on balance- or consistency theory. Randomly selected families (n=222) with four participating members took part in the research. Role-taking ability or accuracy was operationalized as the correct prediction of another family member's choice in a 'behavioral

dilemma' from a 10-item scale. Value similarity can be viewed as agreement. In this line the study is investigating the relationship of agreement and accuracy under differing degrees of affect. Unfortunately, accuracy scores as well as role-taking and value similarity score were again difference scores. The two hypotheses on the influence of high positive affect were supported: role-taking accuracy was high with strong positive affect toward the other family member together with high similarity in social and personal values. Likewise, strong positive affect and low similarity were accompanied by low role-taking accuracy. It was also found that 'in certain conditions value similarity, regardless of the level of affect, is strongly related to role-taking accuracy' (p. 664). There were also indications that especially children with a strong positive affect toward their parents, combined with a low value similarity between parent and child, exhibit the poorest role-taking ability. This confirms the saying that 'love blinds'. Parents did not appear to be bothered so much by this.

Apart from the methodological point on difference scores mentioned above, it is a flaw in the design of this study that 'the affect-condition was taken one-sidedly, i.e. without accounting for the mutuality in affect. High-high affect is very different from e.g. high-low affect, and consistency theory only predicts for the symmetrical affect conditions like 'high-high' a low accuracy in both. Allowing for congruence, the hypothesis on high affect-low similarity would have been: in the case of high congruence and high affect with both persons in the dyad, but low value similarity (agreement), then role-taking accuracy will be low in both. When one takes into account imbalanced states on the individual level, the number of relationships predicted from balance-theory between affect, agreement and accuracy is still more complicated (see White, 1982); for this reason, before analysis individuals who show imbalance should have been removed.

Our final conclusion on this study and others is, that in a study of consensus, the three interrelated levels - congruence (*intrapersonal* consensus), agreement and accuracy (*interpersonal* consensus) - should all be considered, and secondly that the method of analysis must be suited to the *dyadic* unit. In a two-way analysis of variance, as employed here, the dyadic character of the interpersonal measures does not remain intact. A preferable way to test the relations pertaining to the last study is a two-way analysis of variance with similarity coefficients for similarity of values and for accuracy, together with three levels of affect (high-high, low-low and low-high).

A related study is the one by Levine and Arluke (1983) on the relationship between attitude similarity and parent-child attractiveness. The study was done with 85 male and female adolescents, mean age 16, and their mothers. Two-way analyses of variance were done for the child's evaluation of the attractiveness of his mother and mother's evaluation of her child. For the child neither congruence, nor agreement, nor the interaction had any effect on his

level of feeling attracted to his mother. For the mother only an effect of congruence on the level of attraction to her child was found. Although the authors report a correlation of .43 between the evaluations of mother and child, no attempt was made to analyze the data in terms of mutually high and mutually low level of attraction.

3.9 Psychological research on consensus

The question of salience addressed by White (1982) also figured in the study of Zuckerman et al. (1982). This study focussed especially on factors influencing a special kind of congruence, the so-called 'consensus bias', i.e. 'a person's tendency to generalize from a single behavior, whether his or her own or someone else's, to the behavior of the group' (p. 840).

In ingenious experiments the authors manipulated the influence of the kind of attribution - situational or dispositional - that the observers probably made and the salience - here operationalized as the visibility, for the observer, of the target or the experimenter on a videofilm. Finally, representativeness was manipulated, i.e. the observed person, the target, was introduced either as a typical member of his students group or as someone whose personality and behavior did not represent his group.

In a typical experimental situation an actor-experimenter tries to persuade another actor (the target) to walk around campus for 30 minutes wearing a sandwich-board that simply reads 'repent' and to keep record of how people react. Depending on the kind of experiment, the actors themselves and/or the observers of the videofilm with the experimenter-actor dialogue had to judge what percentage of their peers they estimated would agree to carry the sandwich-board. Most hypotheses on the influence of the three factors governing the bias were confirmed. This means that if an observer believes:

- 1 That a person - that may be himself - acts the way he does, because it is consistent with that person's personality and not dependent on the situation;
 - 2 That the acting person is unlike the other persons of the group about which his prediction is requested and;
 - 3 If his attention is not directed to that person (i.e. is not salient);
- then consensus-bias - congruence between oneself and others or between the target-person and others - will be low.

This research is but an example of the many studies reviewed by Marks and Miller (1987), referred to earlier. It has been presented in some detail to convey the flavour of this kind of investigations and to highlight that quite a lot of ingenuity has gone into the detecting of factors that might contribute to consensus-estimates. In our opinion, however, this research has not only failed to produce a consistent theory, as put on record by Marks and Miller (1987),

but has also deterred research into individuals and dyads that demonstrate special patterns of consensus. The review of the literature so far obviously has not bolstered our knowledge in this respect. Although we agree that the study of individual differences in person perception - prominent in the late 1940's and early 1950's - 'came to a crashing halt' (Kenny and Albright, 1987; p. 390) due to the faulty measurement techniques used, we deplore the mainly personal-psychological and inductive way-out this kind of research has taken. As argued before, a different consensus-paradigm and appropriate statistical techniques can inspire interactionally oriented research based on coherent theory. We expect this to produce results that lead to practical knowledge, especially in the field of psychiatric treatment.

Another study on factors influencing person perception and consensus, is Miller's (1977) research on the relation of the way a person communicates, and his perception of communication behavior. In a test situation a communicator-style measure (CSM) that contained a validated measure of dominance was taken from the sets of two persons before and after the dyadic encounter, in which they had to agree on the five objects that best represented American values and therefore earned a place in a time capsule that was not to be opened for 500 years. Results showed that only persons describing themselves as dominant in the CSM were able to perceive correctly the non-dominance of their partners correctly. Degree of dominance - measured with items like being talkative and taking charge - was confirmed by trained raters. The non-dominant person reported after the encounter no difference in dominance between him/herself and the dominant partner. Two interpretations were forwarded for the lack of accuracy of the non-dominant person:

- 1 social desirability
- 2 Perceptual deficiency

As to the first: If dominance is viewed as a social asset, it is adaptive to report the other as just as dominant as oneself. But persons were classified as non-dominant according to their self-reports before the encounter. If this interpretation is correct, then why would they start with reporting themselves as non-dominant? According to the second interpretation, non-dominant persons do not only lack the skill, but also are unable to recognize communicative cues relevant for dominance. The authors cite work of Tagiuri and Goffman that lends support to this view. It should be noted that Miller's study is quite relevant to issues in the treatment of combined behavioral and perceptual deficits.

An interesting application of consensus research was done by Braito et al, (1981), who determined whether male and female college students were accurate in their attribution of behavior to the other sex or reflected a stereotypic view with respect to 'playing dumb' on dates (p. 66). In summary, their results show that men view women as acting more inferior than women report

on their own dating behavior, and women view men as less inferior than men report on themselves about dating. The degree of inaccuracy was different for both sexes however: 29.0% for men and 12.5% for women. The authors conclude that this form of pluralistic ignorance in superficial aspects of dating behavior reflects the persistence of sex-roles stereotypes.

A somewhat related study is the one by Sanders and Mullen (1983), in which they investigated the accuracy of 'minority' and 'majority' (student-) respondents in estimating the percentage of other college students who would indicate possession of a range of attributes. For each attribute - like being optimistic, supporting women's lib - subjects were assigned a majority or minority status, depending on whether the majority of the population did or did not share their attribute. The results are interpreted in terms of the amount of comparison formation a person possesses to found his judgement. The principle finding was that the majority subjects generated more accurate comparison information than did the minority. The majority subjects were also better than chance and better than minority subjects in distinguishing between attributes for which there was high versus moderate consensus; the minority subjects tended to overestimate the consensus for their attributes, while the majority subjects tended to err in the direction of underestimation of their consensus. Several interpretations for these results are forwarded, including avoidance of stigma - being a minority member - and lack of confidence in self-generated comparison due to limited interaction with peers.

3.10 Alternatives in the measurement of consensus

3.10.1 Probabilistic approach

A quite different approach to the measurement of consensus is Markoff's (1982) application of Lieberman's (1969) method for establishing the degree of agreement within a population or between two populations. Markoff notes that Lieberman's index is the probability that a random pair has a certain characteristic and that indices of this kind have been used to assess social and biological diversity, organizational complexity, political fractionalization and the like. This application is illustrated with content-analysis data from the 'cahiers de doléances' of 1789, expressions of public opinion (grievances and recommended actions) on the eve of the French revolution and focused on questions like the internal consensus of groups (like nobility, third estate and clergy), consensus between groups and the variation of recommended actions. The relevance of the research to historians is - among others - to further the interpretation of the French revolution as the work of a class or a whole nation.

The author defines within-group agreement as a probability that is computed by randomly selecting two documents of a group and, within one of those two documents, a random grievance concerning the government institution in question. Agreement is then defined as the probability that this particular grievance is also discussed in the second document. This index of agreement is thus a summation of the products of two probabilities. From analysis of the different kinds of agreement he concludes that the more frequently the Third Estate discussed any aspect of French society, the greater the consensus, and that the Third Estate participated in a nation-wide consensus-forming process, whereas the parishes (rural France) did not (p. 297). The author concludes with recommended uses of this kind of consensus measure.

Although the probabilistic approach is appealing, especially for consensus within and between groups, and the computations are rather simple, the method does not provide a rule to assess significance of levels or between levels. For example, is the difference between agreement with the Third Estate on Lord's recreational privileges and the same agreement within nobility (0.64 versus 0.71) a significant difference? Furthermore, the method does not allow the assessment of other kinds of consensus (like understanding and congruency).

3.10.2 Dynamic approach

To conclude this presentation of measurement models and research examples we will briefly mention the instrument of Spillman et al. (1979) for the dynamic measurement of consensus. The authors aim at establishing consensus as a process - i.e. as changing over time - within a small group, defining the degree of consensus as the extent to which group members have reached unanimous agreement on one of a number of alternatives (actions, preferences). Their approach is based on fuzzy-set mathematics, which offer the advantage of allowing for partial preferences between two alternatives. First, fuzzy sets are made for all possible pairs of alternatives. These sets are used for constructing individual and group preference matrices. From these, consensus and its complement fuzziness (dissensus) are computed by formulas using matrix multiplication. The resulting measures are both between 0 and 1.

The attractiveness of the method lies in the possibility of calculating the level of agreement over a number of alternatives within a group and of distinguishing the individuals who deviate from the other members. Respondents need to have the sophistication, however, to fill out a so-called 'preference matrix scale', and the method is only valid for cases in which agreement can be equated with reaching a unanimous decision in a group over alternatives. Also, since their model has not been rooted in co-orientation theory, there is no question of measuring other levels of consensus like understanding and realization. Finally, the question of levels of significance for the measure has not been worked out.

Comment

The foregoing overview of models and research can be summarized as follows:

- 1 The pervasive applicability of the concept of consensus has been discussed.
- 2 I have endorsed mainly the conception of consensus in the sense of interdependent levels of consensus based on co-orientation theory and discussed some other theories and models to measure consensus, especially in a group setting.
- 3 Although the idea of consensus has stimulated intriguing questions that seem vital for understanding human relationships, defects in operationalization, measurement and statistical analysis have hampered a true development of the research in this field. As a solution we have defended the operationalization of consensus in the form of similarity (see also chapter 8).
- 4 Ancient philosophers posed the problem of 'knowing thyself', and since Freud's elaboration on defense mechanisms, psychiatry has also been thoroughly aware of the difficulties in achieving self-knowledge. How about knowing the other person, are we perhaps doing better in this respect? Consensus in the form of understanding or accuracy is a measure of our knowing the beliefs, opinions, and attitudes of the other. Regarding the issue of 'knowing the other', the overview confronts us with two paradoxical observations. The first is, that consensus is considered to be an indispensable condition for human interaction, while at the same time research evidence shows that we seldom achieve to accurately know the other person, a human shortcoming put on record as early as St. Paul (1 Cor. 13: 12). The second paradox is, that co-orientation theory may lead to the pessimistic view that we are better able to discern differences between our own views and those of the other person when we despise him than when we love him. But at the same time, taking into account the abundance of topics on which men can agree and diverge, it can be more optimistically stated that a positive relationship is doing well according to co-orientation theory of consensus if it follows the folk-wisdom that love aims more for what unites than divides.

Elaborating on the relevance of consensus in psychiatry, I first call attention to Scheff's premise that intersubjectivity - a term quite related to consensus - is primarily based on the expression and recognition of emotions. clinical practice in psychiatry, especially psychotherapeutic interaction, certainly needs successful 'mind-reading'. Psychotherapy is also a field in which affect plays a major role, both in the development of the emotional climate between patient and therapist and in the patient's emotional problems which the therapist should help to overcome. In doing so, the therapist clearly has to be sensitive to the expression and perception of his patient's and his own affects as amplifiers and

qualifiers of all verbal statements. In co-orientational terms, he also has the job of 'unbalancing' or destabilizing his patient out of his unproductive, restricted and affect-laden ideas and perceptions, while at the same time preserving a positive affective relationship. To achieve this, a fruitful psychotherapeutic encounter is not built on a patient-therapist 'match' in terms of stable characteristics (see Goldstein, 1973) but on a co-evolutionary process of emotional development that consists of 'fitting' the actions, ideas, feelings, explanations and descriptions of the therapeutic dyad. Such a therapeutic process extends the consensual domain for both of them by offering new options for behaving, feeling and relating.

The first illustration of such a process is in the following case-study. The idea of 'fitting' implies something similar and dissimilar, but in what sense? Before this will be elaborated, we will first present the case description.

4 Exposure and leave-taking rituals*

As therapeutic methods and techniques increase in number, it becomes more and more important to know on what grounds one is to choose a particular mode of intervention. In strategic or directive therapy, the therapist's interest focuses on the answer to the question: 'What will be effective for this problem, with this client, in this situation?' (Van der Velden & Van Dijk, 1977).

The purpose of this article is to compare two different approaches to emotional problems encountered in working through. This will be illustrated by a case description. The concept of working through is here taken to refer to a process in which vehement emotions attending a traumatic event abate to such a degree that other experiences and courses of conduct can be resumed without impediment (Rachman, 1980). If the process of working through is somehow improper or inadequate, emotions from the past will continue to exert an adverse influence on the individual's present functioning. This chapter is not concerned with treatment of ordinary emotional working-through processes, for these need not be treated as a rule. Rather, it has to do with the distress of clients who were unable to cope with their emotions adequately.

Many publications on difficulties with working through specifically treat *mourning*. Van der Hart (1983a) calls mourning a problem that frequently occurs in important transitions from one stage of life to another. Ramsay and Happée (1977) likewise relate grief and bereavement to all kinds of changes in life; they speak of mourning behavior and grief when the loss involves the death of a parent, spouse, or child, for instance. But 'what we have to say about the mourning process is equally true of other types of loss' (Ramsay, 1979).

Van der Hart (1983a) considers it the task of the therapist to ease the client's passage through a transition into the next phase of life. To this end he employs therapeutic rituals. A significant loss, such as the death of a partner or relative, is regarded as a 'radical transition' from the phase preceding the loss to that following it. As in other transitions or important changes in life, this may generate tensions in those affected. The approach advocated by Ramsay (1979) involves exposure to the stimuli and emotions relating to the deceased until extinction of the grief-linked emotional responses has been achieved. Van der Hart describes transitional rituals as symbolic actions which, in a more or less defined manner, facilitate abandonment of the previous phase and acceptance of the new one. In the following case, the method of exposure, using a

* This chapter has been published in: O. van der Hart (1988), *Coping with Loss. The therapeutic use of leave-taking rituals* (p. 125-134). Irvington, New York, 1988.

technique derived from Gestalt therapy, was initially chosen for treatment. Subsequently a ritual was prescribed for the client. Both approaches contributed to a successful outcome. The discussion will compare and contrast the two methods of treatment.

Case description

Mrs. Jaarsma, the client, is a widow. She is 61 and has a neat, well-groomed appearance. Her behavior is very tense and formal. She has lived alone since the last of her four children left the house some six years before. Her relationship with all four is excellent; she speaks of them with warmth and admiration and says there is no friction in the relationship. All but one of the children live far from her, and she only sees them on such occasions as birthdays and holidays; however, they do keep in touch by telephone. Her husband died suddenly ten years ago. She thinks she coped quite well during the first few years, and until a few months ago her memories of him did not make her cry. Now they do. Her mother, well into her eighties, lives in a home for the elderly in a small neighboring town. The client is an only child and feels highly ambivalent towards her mother; she loves and appreciates her mother's spunk, but she feels blackmailed by all her complaining about the physical infirmities and the loneliness of old age. Mrs. Jaarsma thinks that her mother takes up too much of her time. Mother phones every day between five and six to hear how her daughter is doing and to drop subtle hints that the daughter's love for her old and sick mother leaves much to be desired. Her mother also feels that her daughter's visits are not frequent and long enough; her acquaintances in the home are loved, respected, and admired much more by their children and grandchildren. This pattern of complaining and hinting has been in existence since the death of the client's husband, who was very good at protecting his wife from her mother's claims. But her reactions are never adequate, and she alternately feels angry and guilty.

Before getting married, she had an interesting job; later, she worked as a volunteer for a long time. She quit her volunteer work two years after her husband's death because she felt too unstable. She had several somatic complaints which are still bothering her now. She also felt anxious and depressed. Her youngest son was the last one to get married, a few months before her first contact with the therapist. Since then things have grown even more difficult, for her relationship with her mother, the only relative she sees very often, has become more oppressive. Complaints that had been troubling her to some extent have become worse: headaches, retching, dizziness, panicking if she has to wait somewhere, in a store or when traveling, for instance. She has become gloomy: 'As far as I am concerned, my life may come to an end, only

I am too much of a coward to get out. I feel left behind in an empty nest. I don't know myself any more, I used to be very active and cheerful'.

Her mood is subject to daily fluctuations; her appetite is poor and she suffers from sleeping disorders. There are other depressive complaints and symptoms: she is irresolute, bad-tempered, continually fatigued.

Assessment

The client's primary complaint is the frustrating relationship with her mother. The acuteness of the problem very likely has to do with the recent marriage of her youngest son: the marriage definitely put an end to her role as a mother, and she does not feel capable of dealing with the one remaining relationship, that between her mother and herself, of which her mother is in such firm control.

It looks like her primary problem - lack of freedom from her mother - is linked to a more central one: inability to cope with the transition from the mother-with-children status to that of a woman on her own (the 'empty nest'). Her social isolation and her subassertive attitude towards her mother aggravate the depressive feelings provoked by this inability; she once again starts to miss her husband because it was he who shielded her from her mother's presumptuousness. The mental state of the client reflects an agitated depressive syndrome with vital symptoms.

Treatment

The client's primary objective is to increase her freedom from her mother, but a broader goal would be for her to feel free and happy as a widow living alone. As her motivation largely relates to the goal she has set herself, this is made the focal point of treatment; the therapist expects that the second objective will come up along the way. To accomplish her purpose, the client may have to make it clear to her mother that she is not the self-sacrificing daughter the mother wants and that she does not want to receive so many phone calls from her. Seeing that Mrs. Jaarsma does not want to run the risk of a breach in the relationship with her mother, it is likely to prove very difficult to get her to adopt this assertive attitude. A more indirect manner of helping her free herself from the pressures exerted by her mother will have to be considered (cf. Van der Velden & Habekotté, 1980). The problem is further complicated by her mother's deafness: the client cannot call her mother because she does not hear the telephone ringing. For the calls she dials herself, the mother places a special amplifier on the receiver.

Owing to a tight schedule, the therapist is obliged to postpone an appointment for the commencement of treatment with the client by a few months. Mrs. Jaarsma is to consult her family physician for an antidepressant during the waiting period. Three months later, one of the children phones the therapist to say that his mother is not doing very well. Shortly after that a second talk takes place in which Mrs. Jaarsma mentions that her mother died suddenly four weeks ago. This event has brought about a great change in her life, in the problems she is facing, and, of course, in the therapy. On the one hand her mother's death brings relief, on the other it rekindles her feelings of guilt, now because of the fact that she feels relieved. There is a lot of unvented rage at her mother as well as self-reproach. She is highly perturbed by the fact that she allowed herself to be dominated by her mother for so many years. She believes she would never subject her own children to anything like this, and that she has let go of them as she should. She is proud that it is usually the children who ask her to drop by or to contact them. The therapist indicates he thinks they should wait and see how she handles the situation. But Mrs. Jaarsma feels quite at loose ends, even though her vital depressive symptoms have decreased as a result of the medication. She wants the treatment to go ahead as planned to aid her in working through the confused emotions about her mother. She is afraid that she cannot cope on her own. The therapist decides to go along with her, and the client agrees to the proposed strategy: exposure to the painful emotions, so that they can abate.

At the second session an empty chair dialogue with the mother takes place as part of the exposure technique. Despite her initial trouble in talking to her mother like this, the client ultimately manages very reasonably to express her feelings, which are mainly aggressive. The third session two weeks later she is feeling much better. Her mood has improved suddenly and dramatically; on occasion she has even noticed herself whistling. For the first time, a checkup by her physician has shown a substantial drop in her blood pressure, which means that her medication can be reduced. In consultation with her doctor, she is also going to reduce the dosage of the antidepressant and then give it up entirely. Moreover, she has found the energy to make the necessary arrangements to divide her mother's estate.

This initial success of the treatment has a stimulating effect on both client and therapist. Her complaints have improved and this is proof to both of them that the strategy chosen is the right one. But she found it rather embarrassing and not at all like her to have to express her emotions about her relationship with her mother in front of the much younger therapist. The therapist feels he should respect her sentiments, and he suggests a different procedure with which the same goal can be reached. Now that the exposure session has shown her how her complaints are determined by her feelings towards her mother, the therapist thinks she will be able to accomplish a great deal on her own. They decide that she is to write a 'continuous letter'

(Rubinstein in: Van der Hart, 1983a, p. 121). The daily telephone hour (from five to six) - which, understandably, still reminds her of her former relationship with her mother - is to be her writing time. She will write while sitting in her mother's chair, which she took over from the estate. At this stage it is not yet decided what will be done with the letter eventually, although they agree that something is to happen once it is finished.

The next two sessions are utilized to bring to the forefront significant experiences and episodes in her past contacts with her mother and her feelings about them. This will give more substance to her letter and will help her to put the 'complete message' (Kempler, 1974) down on paper. Old grievances about her mother's reaction to her daughter's deceased husband must be brought up and discussed as well. She never managed to talk about this when her mother was still alive. Now she can also write to her about what a burden it is to go on without her husband, she can tell her mother how much she reproached her for never having wanted to listen to this.

During the next session, the sixth, positive feelings for her mother are also discussed. Client and therapist agree that the letter will now be completed. They consider what should be done with it once it is finished and jointly conceive the following leave-taking ritual: Mrs. Jaarsma will burn the letter and then fix a very fine dinner and treat herself to a bottle of good wine. During this session the client mentions a reunion with former colleagues of hers which she would love to attend. Both client and therapist think it would be nice if she could manage to take leave of her past before the reunion. After the leave-taking ritual she is to phone the therapist for a final appointment.

The seventh and last session takes place six weeks later. She carried out the leave-taking ritual as agreed. The letter, which had become quite a tome, she burned; she felt very solemn while she watched the flames consuming the paper and was surprised that she did not think: 'What a waste of all that work of writing'. The dinner was quiet and pleasant. She put some candles on the table to make it more festive. The next day she went to the reunion by train. She felt free and cheerful and immensely enjoyed seeing her one-time colleagues again. She was surprised but proud that she did not panic in the train, not for a single moment. After the reunion she went to see a newly-born grandchild and then went home, contented and quite at ease.

She truly feels she has mastered the situation and is ready for new experiences and for contacts with friends and with her children. The therapist knows that Mrs. Jaarsma is fond of gardening and, by way of a parting gift, he presents her with a cutting of a plant from his own garden which she had greatly admired. During a follow-up phone talk four months later, she turns out to be doing 'well to very well'. She says: 'I am now much more aware than before I came to you that my own personal life is significant. I used to think that I should always help anybody who asked for it. Now it is much easier for me to turn down my children's requests if they disrupt my own plans'.

Discussion

1 Treatment or not?

It would have been a reason in itself to defer further treatment when, at the second session, the client's mother had died. This would have been a good point to see how Mrs. Jaarsma would cope with this loss on her own. But she urged that the treatment be continued because her complaints persisted and because she felt at loose ends. I feel that in the event of mourning and other working-through problems, it is not so much the period of time that has elapsed but the symptomatology which is decisive in the question of whether or not treatment should follow. Another important argument in favor of treatment was that it was her inability to cope with her status as an older, independent woman rather than her frustrating relationship with her mother that was at the nucleus of her difficulties. There was no reason to expect that the client would handle this problem better now that her mother had died.

2 Treatment

This treatment of working-through problems consisted of a one-session exposure treatment and a subsequent, ritualized writing of a leave-taking letter combined with a leave-taking and a reunion ritual (Van der Hart, 1983a). The burning of the letter, the fine dinner as a woman on her own, and the resumption of social contacts (her reunion and her visit to her grandchild) in this new role are their expression.

We might ask ourselves if we should speak of a *grudge* rather than *mourning*. In my view, this can justly be called mourning: the client is assisted in coping with a significant loss. The death of the client's mother, with whom she had such an intensive relationship for so many years, may certainly be regarded as a tremendous loss, however negative the relationship may have been to the client especially during the last few years. Once the mourning problem had been dealt with, the client turned out to be quite capable of completing the transition to the stage of a woman on her own.

3 Similarities and differences

Mrs. Jaarsma used the letter to take leave of her mother while incorporating certain memories of her husband in the process. The ritualized writing can be regarded as a form of exposure implemented by the client herself. The intention of the letter was to elicit all emotions thus far avoided by inducing provocative stimuli before the client would decide that it was complete. In this sense, the therapeutic ritual shows an affinity with the procedure advocated by Ramsay.

On the other hand, De Tempe (in: Van der Hart, 1988, Ch. 13) demonstrates that the manner in which Ramsay effects a mourning therapy resembles a therapeutic ritual. This is not so surprising, since every psychotherapy can be described as a ritual event (Van der Hart, 1983a, p. xvi).

But alongside these analogies there are differences. A ritual calls for greater creativity on the part of the therapist, but he will find it to be less strenuous and time-consuming than the exposure technique (which, I feel, also encompasses the imaginary confrontation (Burger, 1978) and the primal scream therapy (Janov, 1970)). A Ramsay therapy, as outlined by him and confirmed by my own experience, is often highly exacting and tiring: the therapist must cut off gently but firmly any avoidance maneuver on the part of the client. The duration of a session can hardly be predicted, for it should be ended only after some degree of extinction of the emotions evoked during the session has been achieved. This can create great difficulties in scheduling appointments with clients.

In my opinion, the most eloquent element of a ritual is that the client *independently* performs a series of actions that have far-reaching emotional and cognitive consequences. Going through a mourning process or another working-through problem often involves eliminating a relationship of *dependence*, and a ritual gives fitting expression to this.

However, there are some clients for whom a ritual is not appropriate. There are those who are unaffected by rituals. To them, rituals are empty, childish, merely idle show. It is also difficult to introduce rituals to clients in the very midst of the negation phase of the working through process. An example is a client who, when the question comes up of her difficulty in mourning her husband's death (which took place four years ago), turns out never to have come to terms with the death of her father, who has been dead for forty years. It is most likely that this demonstrates her strong inclination to avoid performing the real 'work' of mourning. By such denials, clients avoid experiencing and expressing their emotions attending traumatic events and they often generate phobic and somatic complaints (see also Ramsay, 1979, p. 12). Clients who persist in their denial are not ready for transition (Van der Hart, 1983a, p. 133), nor can they be motivated to perform the ritual properly. They skip parts of their emotional experience or leave it out entirely, they do not address the deceased or the person they miss directly, they 'forget' to write or to be ritually engaged in other ways. This is understandable: the ritual confronts them with strong emotions, which is exactly what they are trying to escape or against which they are putting up such fierce resistance - there lies the reason for their phobic and somatic complaints. Their complaints seem to be aggravated by the writing and this often induces them to stop, so they never feel the relief of having said everything that was on their mind. An exposure technique properly monitored by the therapist provides a better opportunity to

check the tendency to avoidance and to see to it that during the session the client is continually confronted with the stimuli and emotions he or she so fears. For those who are quite unaffected by ritual, something can be done to overcome this, but with strong avoidance and, hence, denial, the use of therapeutic rituals would seem to be counterindicated.

4 The choice between the two approaches

When equally favorable therapeutic results can be achieved using either technique, I think the therapeutic ritual is to be preferred. Here it is the client who performs the work, while the therapist attends to the context, to the transfer of the 'complete message', and aids the client in appropriately terminating one phase of life, thus facilitating the transition to the next.

In the case of a recent loss, the following should also be considered: as a rule, traditional mourning rituals are performed shortly after a death. This means such rituals resemble what Averill (1968) terms 'mourning': the mourning behavior defined by social conventions and customs.

Seen in this light, a therapeutic mourning ritual might be regarded as a substitute for incomplete or inadequate conventional mourning. Ramsay (1979, p. 14) states that only in exceptional circumstances is 'flooding' advisable if one year has not yet passed since the loss. Clients with more open and accepted mourning difficulties are more easily motivated for the use of a ritual. Processes which have somehow gotten bogged down and are thus usually characterized by avoidance and denial can better be approached using an exposure method guided by the therapist. If a ritual approach is nevertheless favored, the preparatory phase (see p. 16) and, more in particular, motivating the client for this approach, will take a lot of time and energy.

The client whose treatment was reported in this article was sensitive to the meaning of rituals. This was evident from the description she gave of certain habits and activities that she, her husband, and their children enjoyed (celebrating certain events, giving presents) and from her favorite writers (Tagore, for instance). Furthermore, she had not yet had time to disengage herself from her mother's death and from her confusing feelings about it. If the confusion had become prolonged, the client might have come to avoid the working-through process altogether.

A combination of the two approaches does have the advantage that the direct confrontation inherent in the exposure method sets in motion the process of working through. The way is paved for a smooth start to a ritual. The two approaches were similarly combined in the case described here, although we can say in retrospect that it was not strictly necessary since the empty chair dialogue went quite effortlessly.

Further experience with the combined approach has shown that a leave-taking ritual is generally performed with great conviction after a few exposure

sessions the directness of which serves to loosen the process. These sessions are quite strenuous for both client and therapist and it is rewarding for them if they can continue treatment after this and conclude it with a leave-taking ritual. This does the most justice to the client's pace and his or her manner of self-expression.

Comment

To begin with, the foregoing article, that was originally written in 1980, illustrates how ... *all impersonal questions arise from personal reasons and are best understood from their histories* (McCulloch, 1968, p. 1). The article testifies indeed to the coherence or consistency of the author's ideas, undoubtedly based on more private experiences. The more general conceptualization of psychotherapy as a consensual process will be worked out and tested in subsequent chapters, but is already present here 'in status nascendi'.

First, the author 'assesses' his patient (p. 127-128) without a formal diagnostic framework. The last sentence of the assessment has been added at the editor's request. In his assessment he tries to formulate how he understands his patient 'from within', while at the same time reaching for a 'skeleton of truth' (Bateson & Bateson, 1986, p. 6) with the selection of essential relationships - between history, present situation, meanings and emotions - that may be amenable to change. Evidently, the proper understanding of ideas and feelings is important for further consensus and cooperation. In the treatment section the negotiating of agreement is apparent when the therapist 'decides to go along with her' (i.e. the patient) and the patient subsequently 'agrees to the proposed strategy' (p. 129). This consensual quality of a treatment process returns in the statement that 'client and therapist agree' (p. 130). Finally, the author was made aware by a reviewer, that he actively partakes in the consented ritual of leave-taking by offering his patient a farewell-present from his own garden.

In the discussion the author has been rather absorbed by the question of 'similarities and differences' (p. 132) that will return in a more formalized way in chapter 8. He also briefly mentions the ritual as giving a '*fitting* expression' (p. 132, italics added) to overcome a dependency relationship. The case description shows in a rather veiled way, how 'fitting' (i.e. both similar and dissimilar) explanations and descriptions give rise to consensus enabling the development of 'fitting' psychotherapeutic techniques. How patient and therapist both participate actively in the process of consensus and 'fitting' will be illustrated in the following chapter.

5 February man revisited^{*}

Abstract

In working with hypnotherapy we frequently borrow from the work of M.H. Erickson, particularly such approaches as indirect suggestion, metaphors, specific inductions and therapeutic tasks. Because each patient-therapist relationship is unique, it is difficult to duplicate entire Ericksonian treatment procedures. As Erickson suggests, therapy should be adapted to this uniqueness. This article describes a version of the 'February Man' treatment adapted to the specific qualities of the patient-therapist relationship in the context of Dutch culture. A practical adaptation of an Ericksonian treatment procedure is discussed for use in day-to-day psychotherapy.

The first International Congress on Ericksonian Approaches held in Phoenix (1980) was organized around Erickson's motto: 'Each person is a unique individual. Hence psychotherapy should be formulated to meet the uniqueness of the individual's need, rather than tailoring the person to fit the procrustean bed of a hypothetical theory of human behavior'. He also stated that every therapist ought to develop the methods and skills that are suited to his or her personality and experience (Haley, 1969, p. 534), and while offering himself as a model he somewhat paradoxically proclaimed that imitating his method of treatment does not make a person a good therapist.

These ideas are basic to the well-known book *hypnotherapy, An Exploratory Casebook* (Erickson & Rossi, 1979). In the book Erickson and Rossi give eloquent examples of the use of indirect suggestions and of rousing the patient's strengths and experiences. This method leaves much to the inventiveness of the reader. The format of the book thus corresponds to Erickson's basic ideas by appealing to the reader to use his own strengths and experiences in putting these techniques into practice. Erickson and Rossi label these techniques the 'utilization approach'. The final chapter of the book, 'Creating Identity: Beyond Utilization Therapy?', raises an interesting question about the utilization approach. What if the patient has been so severely deprived in some life experience that he or she lacks the unconscious ability to solve his or her problems? In this chapter, Erickson deals with this problem and illustrates how he goes about this. In such cases he may use hypnosis to form a personal

^{*} This chapter has been published as: J.M.M. Dormaar (1987), *February Man Revisited*, *American Journal of Clinical Hypnosis*, 30; 1: 28-35.

relationship with these patients that 'anchors (them) in a more secure inner reality around which (they) can create a new identity for themselves' (p. 460). Erickson is not attempting to negate the influence of the past, but he is 'adding reality to a non-existent thing' (p. 468).

The case of the 'February Man' which he describes in the last chapter is 'a synthesis of several original manuscripts' (or case descriptions) (p. 460) in which Erickson applied the procedure. It is quite complex. Put simply, Erickson introduces himself into the experience of a female patient who had achieved age regression under hypnosis. He does not insert himself as he is, for that would not be appropriate to the patient's childhood world, but he assumes the identity of an old friend of the patient's father who chances to pay a visit in the month of February - the month in which her birthday falls. She reexperiences important moments from her childhood, while Erickson in his role as February Man adds positive experiences to her cold memories, which are permeated by the emotional absence of two busy and preoccupied parents.

The following clinical vignette is a further example of how this Ericksonian approach can be adapted to a treatment procedure tailored to the needs and abilities of an individual patient and an individual therapist. The purpose of this article is to take a practical look at the influence and viability of Erickson's approach. It therefore seems important to know what other therapists in their day-to-day work can glean from the abundance of Erickson's experiences, techniques and ideas. This case is presented in order to share a stimulating learning experience of the author and to outline some principles possibly underlying the successful adoption as well as the practical application of his techniques.

Referral and assessment

Mr. E., 56 years old, came into treatment in a state of panic. He felt 'completely rigid' and complained of great fatigue, pain in his neck, headaches, and gloominess. Mr. E. held a technical position in a large company which puts a lot of pressure on its employees to get the work done promptly. In addition, he had the remarkable hobby of working for over 20 years on a treatise about the development of astronomy, human thinking, and society. This involved a very extensive literature study requiring row upon row of files - the whole attic full - to be forged into a short and lucid treatise. The work had been in its final stages for years and was a lonely task.

Mr. E. was afraid of losing his job. His work and his scientific treatise left him no room for anything else. He had no time for his family, for his wife, for social contacts, or for other 'trifles'. He could rightly be called a 'tormented workaholic': *'Work, work, work is what I have learned to live by. There was no*

time for my wife and children and now that there is, I can no longer change. I feel like a machine, standing on railroad tracks'. His behavior bore out his self-description; even in the waiting room he worked on his treatise. He was a tense and lonely man with no friends. The relationship with his wife was a vast gulf bounded by her silent reproach about his shortcomings as a husband and father. What he wanted most from therapy was to learn to relax.

Diagnostically, he was not clinically depressed, but rather exhibited a state of decompensation with depressive and functional features, perhaps a psychasthenic man in the climacterium virile.

Previous treatment and first phase of therapy

Four years ago he had been in intensive treatment by another psychotherapist, who attempted to decrease his compulsive working by going outdoors with him to draw and paint. This was something Mr. E. did as a child with great dedication. He actually wanted to be a painter. Before this treatment he was under long-term medical care for several complaints: headaches, stomach ulcer and bad feet. Recently he had switched to a chiroprapist.

The first five sessions were devoted to 'relaxation', as he had requested. Not classic relaxation, but to feel free and in control was what he really seemed to need. Relaxation is not a necessary condition for this and may even be a disadvantage (Thompson, 1979). This led to an active, alert form of trance induction and utilization. The imagination took hold of what he would really like to do but never got around to: to go somewhere out-of-doors and not have to think any more. Many indirect suggestions were made about letting go and being personally involved. Although he was a good hypnotic subject during the sessions, the procedure did not work as the effect did not generalize to daily life.

The second phase of therapy

The therapist realized that he needed more information from the patient's life history. He had been the youngest son of a large Catholic family living in Leiden. His father managed to remain on the job during the Depression, working just as hard as Mr. E. did nowadays. The family had to move to a poor lower-class district when the patient was about 10, a place where he never felt at home. He was a slight child and felt quite intimidated by the youths on the street; he did not dare to fight, nor to run away. His mother was a nervous woman, always at work, and unmindful of his plight. When he was 14, he went to work to contribute to the family income.

On the basis of detailed information it became clear that the message Mr. E. got was to sustain himself by hard work. If not, all the dangers in the world hover ominously over him: poverty, attack, subjugation, abuse. Secondly, work has helped him to disguise shortcomings in his personal functioning.

As Mr. E. was inclined to turn more direct suggestions into tasks and orders which blocked his own spontaneity, a metaphoric strategy which merged the legitimate and metaphoric domains (Van der Hart, 1982) would be the most likely to let him discover new content and meaning in life. The strategy should aim at extending and improving his primary functions in a socio-cultural sense by increasing playful, competitive, artistic, and imaginative capacities (Huizinga, 1955). Moreover, the choice of the therapeutic strategy should take into account that Mr. E. ascribe his tense and mechanical functioning primarily to his childhood experiences. Because of Mr. E.'s emphasis on hard work combined with the emotional and affective impoverishment of his early and current life, the therapist chose to use the technique described by Erickson as the February Man (Erickson & Rossi, 1979).

With these considerations in mind the therapist now suggested that he continue working on the relaxation technique Mr. E. had asked for. He easily entered into a trance which was used to introduce time distortion and age-regression. Mr. E. regressed to his first visit to this therapist, to the start of the previous treatment, to when he had just gotten new orthopedic shoes, and finally to the time when his feet were not yet finished growing because he was still a boy of 12. In this state a bridge was built to a pleasant period when the family was still living in a spacious and airy house in a good neighborhood in Leiden: 'You remember it so well that it is as if you were sitting out in the sun parlour once again. And wouldn't it be nice to sit down at that table and draw?'. There follows extensive discussion of the drawing and the pencils, thus strengthening the reality of the regressive experience. The therapist then introduced himself gradually into the experience of the patient, speaking in a thoughtful, evocative manner: 'It's a beautiful day, isn't it? How quiet it is outside ... Are you going to color that part blue?'. Then a visitor was introduced: 'Someone is coming to your front door, do you see him?'. When the client followed these suggestions, the therapist shifted his reactions to the first person as a friend of father's: 'I think it is very nice what you are making there. My, you draw well. I love those bright, warm colors. Did you make that up yourself or did you see it somewhere? I don't have anything special to do. I'm just going to sit down and make myself at home with you, and wait until your father comes back'. At this point the client reexperienced an event with his parents. He had made a painting and wanted to show it to his mother. She did not react. He went upstairs crying and laid down on the floor in his room. Father peered inside, saw him lying there crying and closed the door again without saying a word. He felt very lonely. The friend came upstairs for a

minute and soothed him. After admiring the painting, he put the parents' reactions into perspective: It is terrible if you are 6 and your parents act like that. But when he would be bigger he would understand his parents better and no longer take their incomprehension so hard. The boy calmed down and went outside to play. In the same way, the following sessions were devoted to re-experiencing various events together with the therapist as father's friend: taking trips all over town; basking in the sun in parks and public gardens; walking the quays on the Rhine and Schie Rivers; strolling the canals and to the seaside town of Noordwijk; sometimes alone, sometimes together with 'father's friend'. When out to play football the friend waited for him, cheering him up when other children called him names. At home, lying sick on the sofa, they played games together. In the small house, he sat on the flat roof of the shed drawing the houses behind theirs with a setting sun. By then he had grown bigger and told about school, about the big boys, and about his father who was never there. In this way a great many angry, sad and anxious memories were relived, and all the additions were aimed at support, understanding, playing, having fun, being out-of-doors.

Throughout the session the therapist used a grammatical variation possible in the Dutch language for addressing in the second person (formal 'u' versus personal-informal 'je' or 'jij' in Dutch) to indicate in what roles they were dealing with each other. Thus every session was closed as it had started: a short formal trance induction in the 'u' form, from which he awakened 'refreshed and wide awake'. Any commentary on the session related only to the final moments; the regressive experience was never discussed, and the therapist did not investigate amnesia.

After five of such sessions Mr. E. related that things have been going well recently; he hardly had pain in his neck and he had helped his wife with a drawing stating that 'she doesn't push me away anymore'. At his work he set clearly his boundaries by stating that he will be finished when he is finished. In addition he no longer worked through lunch.

During the next session numerous episodes were brought up such as his bashfulness with women and his contempt for his boss to which he rapidly associated experiences that the therapist could not follow. Without talking very much, he laughed, cried, and sometimes trembled in a way that looked like repulsion. His reactions seemed to point to a catharsis (Scheff, 1979).

Soon hereafter, Mr. E. felt the time has come to conclude the treatment. His complaints had nearly disappeared. He again mentioned the improved relationship with his wife, but nothing about his masterwork, the treatise. It did not seem opportune to the therapist to bring it up at this point either: the impression could easily be aroused that the therapist really thought it is time Mr. E. completed his manuscript.

Follow-up after three months

Mr. E. felt that things were 'really going unbelievably well' for him. Both at home and at work, he took his time for self-hypnosis. He went into a trance almost automatically if his neck or his fatigue told him it was necessary. The therapist praised this development, but was surprised by it: the possibility of self-hypnosis was never explicitly mentioned.

To improve his condition, he went for fitness training with his daughter, a mutual activity which gave him great pleasure. Concerning the relationship with his wife, he remarked with a nonchalant grin that, to her surprise, he recently did not comply with what she asked. This made her mad, but he remained calm. It is telling how he spontaneously summed up another memory about 'feeling left in the lurch by father'. 'He probably just didn't know how to reach us - the children'.

Discussion

The deep, somnambulistic trance, the positive hallucinations, and the covering of trance-work with amnesia described by Erickson show the artistry of the master-therapist, whereas the present case is a more straightforward use of essentially the same technique: adding new, positive experiences to the revivification of important childhood episodes, which enable the patient to gain a different perspective on his past and less restricted ways of responding in the present.

Some differences, however, in the *practical execution* of the therapy can be highlighted. First, Erickson trained his patient in hypnosis in a session lasting for 5 hours, whereas in this case five sessions of one hour were used. One wonders about the extraordinary impact on both patient and therapist of merely being together for such a long time in an intimate, personal interaction; such intensive encounters with one other person, personally or professionally, are indeed rare. Sessions of 5 hours and subsequently of several hours (p. 470), however, are also inconsistent with the social and practical constraints of therapeutic interaction. A second difference lies in the kind of cue used to define transitions in the hypnotic state: Erickson gently introduced a wrist cue, using his watch; in the present case a verbal cue was used: in the Dutch language the *u-jij/je* alteration is an accepted way of structuring a relationship from respectful and formal to addressing a friend or child. Since this cue is no foreign element, it makes use of the patient's social and cultural context defining the nature of the relationship. As mentioned before, the essence of the procedure lies in supplementing the reality of the past; in discussing the principles behind this, Erickson and Rossi remark that it is an adding of new

experiences to the past. When further deposits to the memory bank are made, which the patient may use to change his own history and with it an important part of his own perception, traumatic experiences can be compensated. This is, however, a construction which may be termed a therapeutic myth. It rests on the unproven theorem that by using hypnotic revivification it is possible to add current hypnotic experiences which the person and his memory will accept, after the therapy, as being just as true and real as what the person reported about his childhood before therapy.

Milton Erickson is already quite a mythical figure, and though a myth can be very exciting, it does not further real understanding of his therapeutic technique as one can easily lose sight of the practical applications of his approaches in daily psychotherapy. The question mark in the subtitle of the chapter - *Beyond Utilization Therapy?* - suggests that the authors themselves doubt the mythical claim of being able to add something new to a person's memory. Clarifying the doubt Rossi '... points out the self-corrective aspect of the psyche; it is in a continual process of reformulating and resynthesizing itself to achieve a more integrated pattern of functioning' (p. 472). But how does Erickson's approach in the 'February Man' actually further this process? Not by adding something completely new, as a closer look at both cases will reveal.

Erickson's patient had a rather clear idea about what she missed during her youth; she was convinced 'that her rejection by her mother ... would in some way adversely affect the handling of her own baby' (p. 462). Mr. E. on the other hand, experienced a lack of joy and creativity and now felt 'like a machine standing on railroad tracks'. Moreover, both patients had explicit experience with what they were presently missing: in Erickson's case the father 'did have genuine affection for his daughter' (p. 461), while the patient described here had enjoyed drawing and painting when a child. Consequently both therapists geared themselves precisely on these 'old learnings'. Erickson did not introduce himself as a mother-figure, as one might expect if he really were to substitute for her lack of affection, but acted like a loving and supportive father-figure. Likewise, the present therapist explicitly joined with the patient's early playful and artistic capacities.

Thus, in both cases the therapists actually reactivated what already existed, which is clearly within the utilization framework. Instead of 'adding reality to a nonexistent thing' (p. 468), they are enlarging and freeing the existing realities of emotion and experience for the patient. But the 'February Man' is more than just another example of the utilization approach. It also illuminates an important treatment factor that in order for a psychotherapeutic technique to be effective there must be a fit between the explanations of patient and therapist regarding problem-formation and cure.

That this fit exists in both cases is readily apparent. The similarity between Mr. E. and the patient described by Erickson lies primarily in their explanation about the origin of their problems: an unhappy, restrictive, and cold childhood. When people account for their attitudes by means of their childhood experiences, it is probably advisable to go along with this explanatory model during treatment. In general there is remarkable agreement between both professionals and lay persons about the origins of neurotic problems, their persistence, and required treatment. The source of the difficulties lies in the past, primarily in traumatic experiences or childhood events, and the influence of this source must be removed via insight. As long as the person has no insight - especially to the powers of the unconscious - the problems will persist, and insight can be brought about by treatment. According to classic Freudian theory, repression lies at the basis of psychopathology, and the elimination of repression by means of conscious insight is the primary goal of therapy. This is, concisely, the common view. Erickson, however, did not share this view of human behavior and experience (Zeig, 1980, p. 103), but he certainly used parts of it as can be seen by his treatment technique. Erickson actually attempts to *complete* the implicit logic of the patient's explanation in a new way: if something went wrong during childhood, you do not need insight but you have to relearn as a child, by the revivification of childhood experiences.

In general the concept of fitting the therapist's explanation to the patient's explanation may be understood when we characterize the activities of the therapist as acting or being one side of an interaction. This is to say that psychotherapeutic activities create the reciprocal responses in the patient, thereby setting the context for a certain class of responses to occur (Bateson, 1972). The patient's ability to respond in the expected way has to preexist. One builds on existing presuppositions of thought and social behavior; for example, a context exists for showing affection or creativity as in the previous cases. The regressive hypnotic state is used by the therapist to free the person from more recent limitations in responding. Within this framework, the therapist's explanation of problem-formation and cure can be summarized as: spontaneous emotions and ways of relating are thwarted by lack of stimulation in the patient's youth. A patient with such problems needs to reintegrate these feelings to function more freely.

The patient's view and the therapist's view are explanations in the sense of being constructions of reality that are not alike but somehow fit together. This idea of fitting is constructivistic (von Glaserfeld, 1984). The therapist must take care that his patient's constructions and his own - instead of being identical - fit together so that a psychotherapeutic technique can accomplish its goal. Explanations that 'fit' in this way provide 'multiple versions of the world' (Bateson, 1980). This creates a 'bonus' and adds an extra dimension to reality that offers new options for behavior and experience.

I surmise that creating this level of fit is essential to the adaptation of treatment techniques to the abilities of both patient and therapist. The uniqueness of this fitting together may also enable us to comprehend more fully why Erickson said '... that you cannot imitate someone else' (Haley, 1967, p. 535) to which we might add 'and even not yourself'.

Author Notes

I wish to thank Sandra Escher, Rob Rotteveel, Kees van der Velden and Marten de Vries for their comments and suggestions.

Comment

The last case-report again shows the search for a consensual framework that was only found after the therapist reconsidered the patient's history that led him to a better understanding and a fitting treatment strategy (p. 30). The idea of 'fitting' takes different meanings. First, a fitting strategy seems to imply to add what is missing, or at least to reactivate traits and behaviors that already existed in a suppressed form. More importantly, 'fitting' is elaborated as the fitting of explanations and the word 'explanation' is here clearly used to denote the perspective in which a person puts his experiences. This idea of fitting is also characterized as 'acting or being one side of an interaction' and thereby creating reciprocal responses in the patient by 'setting the context for a certain class of responses to occur' (p. 34). The reference to Bateson (1972) was to his articles 'the cybernetics of self: a theory of alcoholism' (p. 309-327), especially the section on 'pride and symmetry' and to 'the logical categories of learning and communication' (p. 279-308). The quotations from the present article may seem rather cryptic and are therefore expounded.

It is not only the patient who is bound to his neurotic and thereby restricted definitions of himself and his relationships, he also binds the partners he is interacting with. Limitation being his weakness and his strength, the patient thus 'sets the stage'. In therapeutic parlance: the patient wants unilateral control of the relationship. The patient of M.H. Erickson that has been referred to in the article, essentially says: 'I cannot take care of children, because I have not been taken care of' and probably implies that 'people still do not care about me'. In common human relationships the partner plays a part in this script, either by confirming the patient's outlook - 'indeed, you are a nuisance' - or by refuting it - 'But I *do* care about you' - thereby creating a kind of 'game without end' (Watzlawick et al., 1974) as the person does not feel accepted and understood in his private experience. A therapeutic relationship characteristically starts with a thorough understanding that implies both confirmation and refutation by redefining the relationship: 'indeed that was true' - while confirming the private experience the therapist includes a 'distancing device' (Scheff, 1979), namely a time-lag. Detailed attention to history-taking often works that way, albeit often more to the benefit of the therapist's aloofness than to promote the patient's handling of his present difficulties. At the same time he contradicts subtly, in Erickson's case by implying 'now you will learn what it takes to care for a child by acting like one during therapy'. If the patient accepts the therapist's 'double-talk', he also allows a slight altering of the stage because the therapist shifts the role of the protagonist in such a way that the play can be brought to a happy end. This change of the stage-

setting is foremost a question of 'fitting' in the constructive sense, as made clear by Watzlawick (1984). He contrasts 'fit' with 'match'. In his terms matching signifies 'being alike', while 'fitting' stresses the point that the therapist's construction of the patient's reality just works well and that we will never know for sure what the patient's construction of his private world really looks like.

This kind of deliberations on consensus made me wonder whether I could recognize the influence of consensus in the work of colleagues. This is the basis, in a nutshell, for my research into consensus as a vital aspect of successful psychotherapy, whatever the theoretical stance of the therapist or the nature of the patient's difficulties. The following field study has thus been guided by personal experience in clinical practice which I deem indispensable for doing psychiatric research, since practical and personal knowledge - as opposed to textbook knowledge - elicits research questions in a natural way. Consecutive theoretical understanding often needs far more general and abstract ideas than psychiatry itself can offer.

When investigating whether consensus might promote successful outcomes, we first have to know how to define success. Therefore the first part of the field study will deal with the question how improvement can best be operationalized.

6 Psychotherapeutic outcome: Measuring reliable clinical change

Abstract

Psychotherapeutic research is generally carried out to investigate the effectiveness of specific techniques with selected, often homogeneous samples of patients. This leads to problems when applying research findings to the clinical reality where the patient population is heterogeneous. In this study, we examine psychotherapy outcome in a naturally occurring treatment setting using criteria established by Jacobson and his colleagues for measuring improvement in psychotherapy, reliable change and clinical significance. Outcome was evaluated in a sample of 117 Dutch ambulatory mental health center patients with the SCL-90 Symptom Checklist and the Target Complaints List. On both instruments improvement rates varied greatly and were only moderately correlated. Our findings suggest that a more limited application of the Jacobson criteria derived from selected samples is warranted when applied in a general patient population.

Most psychotherapy outcome research examines specific techniques and patient samples. This research trend, which may be viewed as the 'flight into specificity', gains in precision and theoretical understanding, but loses in clinical relevance. clinical practice does not provide us with homogeneous samples of patients and problems. We deal with heterogeneous populations with diverse and often multiple problems. Under these circumstances common features that can account for effectiveness in psychotherapy are at least as relevant as specific ones (Frank, 1969). Common or nonspecific factors are by definition active under divergent conditions; therefore research into their effect requires heterogeneity in patients, therapists, treatments and problems.

Another reason why classic psychotherapeutic outcome studies often have little bearing on clinical practice is that by comparing experimental and control groups the central question is whether patients improve significantly more in the one group than the other. clinically, however, a different question is of interest: what proportion of patients is clinically improved under certain treatment conditions to a level that may be considered a cure?

* This chapter has been published as: Maarten Dornaar, Chantal I.M. Dijkman-Caes and Marten W. de Vries (1988), *Psychotherapeutic outcome. Measuring reliable clinical change. Psychotherapy and Psychosomatics*, 50: 95-101.

Addressing this question, Jacobson et al. (1984a) have recommended a twofold criterion to provide 'information about the effects of psychotherapy which is both clinically meaningful and statistically reliable'. They operationalized the *clinical significance* of attained changes and proposed a *reliable change index* which is 'able to determine that the change is real' that is 'of a sufficient magnitude to rule out chance as a plausible competing explanation' [p. 344]. Jacobson's approach has, however, rarely been used up to now. The original authors, notably Jacobson et al. (Jacobson et al., 1984b; Jacobson et al., 1985) and a few others like Cline et al. (1987), have reported results of marital therapy in this way. Mavissakalian (1986) and Arrindell et al. (1986) have used it to examine treatment outcome in agoraphobia.

We extended this line of research by examining the effects of common factors on the level of psychotherapeutic improvement in a general patient population. We did this within the context of an ongoing study into the influence of selected characteristics of the patient-therapist relationship on the success of psychotherapy. Employing general outcome measures in a typical outpatient population, we evaluated the usefulness of Jacobson's criteria to categorize patients into improved, unimproved and worsened groups. We also report the correspondence of this grouping to one based on a second outcome instrument and another criterium for improvement.

Methods

Research setting

The study was carried out in a Dutch Center for Ambulatory Mental Health Care. Two short-term treatment teams for adults participated, where an eclectic range of therapies is provided by a total of 22 therapists using behavioral, experiential, interactional and psychodynamic approaches.

Subjects

Subjects met the following criteria: age between 18 and 60 years, no overt signs of psychosis or mental retardation and at least 4 contacts with the same therapist. All subjects were patients who gave their informed consent during the 12-month research period. In the normal routine many patients had only a few contacts for advice or were referred to a different team or therapist. Consequently, these patients were included in the total number of dropouts. No control group was used since data were collected with the explicit aim of estimating the progress during therapy in naturally occurring therapy settings. The total group consisted of 162 patients: 94 females and 68 males. Their mean age was 34 years (SD 11 years). 64% were married, 28% single and

11% widowed or divorced. 60% of the patients belonged to the low and low-middle income groups, 84% were in individual treatment with a mean number of sessions of 10 (range: 6-20). These demographic characteristics were compared to those of patients who refused or were omitted ($n = 254$), or who dropped out ($n = 174$). No significant differences between the three groups were found. Although diagnoses were not always assigned, most patients suffered from neurotic and adjustment disorders.

Procedure and instruments

Therapy outcome was assessed by means of two patient self-report questionnaires and one report by the therapist. Before intake the patient filled out the Dutch version of the SCL-90 symptom checklist (Arrindell and Ettema, 1981), range: 90-450, originally developed by Derogatis (1977). After 2 sessions both the patient and the therapist filled out the Target Complaints List (TCL) devised by Battle et al. (1966), which evaluates change in individual complaints (range: 0-10). Both instruments are well known for measuring outcome in short-term treatment (Emmelkamp et al., 1985; Koss et al., 1983). The posttest for both SCL-90 and TCL was after 6 months, representative of an average treatment period. In order not to lose cases due to difficulties in tracing patients and not to confound posttests with follow-up data, this interval seemed best.

Analysis

Improvement reaches clinical significance when 'the level of functioning at posttest is statistically more likely to be (present) in the functional than in the dysfunctional population' (Jacobson et al., 1984a). With the following formula a cutoff point 'C' is computed between the (presumably normal) distributions of functional and dysfunctional populations with unequal variances:

$$C = \frac{s_0 X_1 + s_1 X_0}{s_0 + s_1}$$

The suffix '0' refers to the well-functioning, normal population, '1' to the distressed, dysfunctional population; s = standard deviation, X = mean.

clinical significant change is a *movement* from the dysfunctional to the functional distribution (Jacobson et al., 1984a; p. 340). In our results we will show this movement clearly by a division into four groups.

Of the instruments used in this study, distribution norms of a well-functioning population were available only for the SCL-90. Arrindell and Ettema (1981) report a mean of 125 and a standard deviation of 30.3 for a

normal representative Dutch sample ($n = 907$). Our patient sample ($n = 162$) had a mean of 203 and an SD of 66 at pretest, which is very close to the mean = 203 and SD = 60.2 reported by Arrindell and Ettema in a sample ($n = 703$) of phobic patients. For the current sample the cutoff point between functional and dysfunctional scores is 150.

The reliable change index is computed by dividing the difference between pre- and postscores by the standard error of measurement. According to Christensen and Mendoza (1985) it is more appropriate to use the standard error of the difference between two testscores (s_{diff}). The reliable change index RC' is therefore:

$$RC' = \frac{(X_2 - X_1)}{s_{diff}}, \quad s_{diff} = s_E \sqrt{2} = 18.6$$

A value over +1.96 means reliably changed toward improvement and likewise a value below -1.96 can be called reliably worsened.

To ascertain which of the patients who had attained a clinically significant level of change could also be considered *reliably* changed according to their reliable change index, data on both criteria were combined into *reliable clinical change*.

For the TCL, improvement was defined in two ways:

- a for comparison with clinical significance of the SCL-90, a patient was considered improved with a postscore under the cutoff point of the mean (prescores) minus 1 SD (prescores). For the SCL-90 the cutoff point is close to 1 SD below the mean (prescores), and the cutoff of the TCL was similarly defined;
- b for comparison with reliable change of the SCL-90, a patient was considered improved with a change of more than 1 SD of the change distribution. Relative to the SCL-90 this criterion for improvement is more stringent. For the SCL-90 a change of 1 SD of the change distribution corresponds to an RC value of 3.3.

Results

The pre-/posttherapy differences at the group level were highly significant, $p < 0.001$ (t-tests, two-tailed), for all three outcome measures SCL-90, TCL reported by patients (TCL-P) and TCL reported by therapists (TCL-T).

Table 1 shows the distribution of the patient-sample regarding the level of clinical significance.

Table 1: Distribution of clinically significant change

SCL-90 postscores	SCL-90 prescores	
	< 150	> 150
> 150	group 1 (worsened) n = 1 (0.8%)	group 2 (unimproved, distressed) n = 52 (44.4%)
< 150	group 3 (unimproved, well) n = 19 (16.2%)	group 4 (improved) n = 45 (38.5%)

Of the 64 patients who had a postscore of positive clinical significance (i.e. lower than 150), quite a number (group 3, n = 19) had a prescore that was already in the well-functioning range. For reliable clinical change, the results are presented in *table 2*:

Table 2: Number of patients with reliable clinical change (SCL-90)

RC groups	clinically significant change groups				Total
	1	2	3	4	
1 Worsened (RC < -1.96)	1	6	0	0	7
2 Undecided (-1.96 < RC < +1.96)	0	27	18	5	50
3 Improved (RC > +1.96)	0	19	1	40	60
Total	1	52	19	45	117

Patients (29%) who worsened or were 'unimproved, distressed' at posttest and who demonstrated no reliable change toward improvement (n = 34; 29%) really did poorly. On the other hand, those who changed from distressed to improved with a reliable change (n = 40; 34%) are the patients who made real progress during therapy. A remaining group of patients (n = 43, 37%) demonstrated ambiguous results.

For the TCL-P improvement, in the sense of clinical significance, is reached at a postscore severity rating of 4.8 and for the TCL-T at 6.5. The condensed crossbreaks of TCL-P and TCL-T postscores with clinical significance of the SCL-90 are shown in *table 3*.

Table 3: Number of patients with functional and dysfunctional scores on the TCL and the SCL.

		SCL-90		total
		functional (postscore < 150)	dysfunctional (postscore > 150)	
TCL-P	Functional (postscore < 4.8)	44	16	60
	Dysfunctional (postscore > 4.8)	15	36	51
	Total	59	52	111
	Chi-square 19.6, $p < 0.001$			
TCL-T	Functional (postscore < 6.5)	44	33	77
	Dysfunctional (postscore < 6.5)	15	16	31
	Total	53	49	108
	Chi square 0.4, NS			

As table 3 shows, 80 patients (73%) are either functional or dysfunctional as established with the SCL-90 and the TCL-P ($r = .59$, $p < 0.001$). Only 60 patients and their therapists (56%) agreed on the (dys-)functionality as assessed with TCL-T and SCL-90 ($r = .20$, $p < 0.05$).

Patients were improved, in the sense of reliable change, with a raw difference score over 2.7 for the TCL-P and over 2.5 for the TCL-T. *Table 4* shows the condensed crossbreaks of improved versus other patients for TCL-P with SCL-90 and TCL-T with SCL-90.

Table 4 shows that 38 + 30 patients (65%) had reliably changed (improved and not improved) according to both the TCL-P and the SCL-90 ($r = .37$, $p < 0.001$). For the TCL-T and SCL-90 this concordance rate was 63% ($r = .26$, $p < 0.05$).

Table 4: Number of patients showing reliable change, measured with TCL-P/TCL-T

SCL-90		a	n	d
		SCL-90		
		improved (RCI > + 1.96)	undecided and worsened (RCI < + 1.96)	total
TCL-P	Improved (changescore > 2.7)	38	22	60
	Undecided and worsened (changescore < 2.7)	15	30	45
	Total	53	52	105
	Chi square 8.1, $p < 0.01$			
TCL-T	Improved (changescore > 2.5)	30	15	45
	Undecided and worsened (changescore < 2.5)	23	34	57
	Total	53	49	102
	Chi square 6.0, $p < 0.05$			

Discussion

The division of clinically significant change into 4 levels provides a means for controlling for patients who were already in the well-functioning range before therapy. The combination of clinically significant change with the reliable change index into *reliable clinical change* gives a viable definition of 'cured', since it implies a significant change toward health. These criteria are clinically meaningful and allow the comparison of treatment effectiveness across studies. Another advantage is that the criterium of *clinical significance* offers a comparison with a sample of the population and with a sample of untreated patients. These comparisons can substitute for those with traditional control groups, particularly if the stability of the measure has been established.

It is improbable that the high variance in the SCL-90 scores of the dysfunctional population is accidental, since the means and standard deviations of our patient sample and that of a Dutch phobic sample (Koss et al., 1983) are nearly equal. This great variance gives rise to a high standard error of the

difference between two scores (s_{diff}) and therefore to a substantial range of difference scores that do not qualify as reliable change: a high variance in the patient population makes it more difficult to ascertain improvement. Likewise, a high variance in the dysfunctional population in proportion to that of the functional population results in a cutoff point for clinical significance moved in the direction of the functional population. The high variance makes the criterion for clinical significance more stringent than in the case of lower and equal variances for the functional and dysfunctional populations. With overlapping distributions the cutoff point as calculated here is the best choice according to Jacobson et al. (1984a) and is to be preferred over a criterium for functionality based on 2 standard deviations from the mean of the functional or dysfunctional population. This approach does mean, however, that quite a number of patients have functional scores and that many persons of the normal sample score in the dysfunctional range.

As for the clinical relevance of the two criteria, clinical significance has some drawbacks. On the question of well- or dysfunctioning, patients and mental health professionals have different vantage points (Strupp and Hadley, 1977). clinical significance is based here on patient reports, but also reflects the values of the professional, who tends to view a patient as having regained health if he does not distinguish himself by his level of psychopathology from the population at large. The substantial numbers of both normals and patients that score on the 'wrong' side of the cutoff point illustrate that many non-professionals do not share this view. It may also be pointed out that it is mainly the level of demoralization that determines whether one 'feels' like a patient (Frank, 1975), and highly demoralized persons express their despair by many and severe complaints (Van der Velden, 1987). So, if his self-report scores show a reliable improvement, the patient thereby indicates that he feels less demoralized and less a patient. If his posttherapy score is still in the dysfunctional range, he now possibly belongs to the vast number of people who can be given a psychiatric diagnosis in epidemiological surveys (Myers et al., 1984), but who do not seek help in the mental health care system. These remarks lead to the conclusion that when establishing clinical improvement, greater weight should be given to pre-/posttherapy differences as assessed with the reliable change index than to the level of functioning as reflected in the clinical significance criterion.

Comparing the improvement according to the three measures it is striking that patients show a fair agreement between measures regarding *clinical significance*, whereas the correspondence between patients and therapists is equal to chance. This difference between patients-correspondence and patients-therapists correspondence is not marked for the *reliable change* criterium. This result strengthens our preference for reliable change over clinical significance in judging improvement.

In view of the drawbacks of the great variance of a general outcome measure like the SCL-90, one might argue in favor of a restricted patient sample, or of a more specific group of patients to be measured with a specific instrument. This procedure would certainly lower the s_{diff} and separate the functional and dysfunctional distributions more clearly. The proportion of patients who change reliably and reach the level of clinical significance at posttest might then increase. These strategies, while methodologically sound, eliminate an important characteristic of the patient population in a general practice, namely its great variability. As stated earlier, we hold that retaining natural variability is desirable in research on nonspecific or common factors.

If the combined criteria are not feasible, we recommend the criterion of reliable change for reporting improvement, as it is the easiest to obtain and also seems preferable to clinical significance. If no reliability coefficient is available or cannot be computed - as in the case of the TCL - the criterion of one standard deviation of the change distribution may be a reasonable alternative to determine clinically meaningful change.

Acknowledgement

The research was supported by the NFGV (Dutch National Foundation of Mental Health) (grant No 1.816) and by the University of Limburg. The authors are grateful to participating clients and therapists of the 'RIAGG Maastricht e.o.' (director: H.B. Pomerantz) as well as to M.A.J. Romme, chairman, Department of Social Psychiatry and J.J.C.B. Bremer, chairman, Department of Medical Psychology, Medical Faculty University of Limburg, for their support.

Comment

The main conclusions on the measurement of clinical change have already been discussed in the article. In this brief interlude I would like to offer some remarks on the general model and presuppositions of psychological measurement. First, we draw the reader's attention to the rather intricate manoeuvres with the 'raw data' that have been made to allocate patients as improved, undecided or worsened. This kind of manipulation is quite common and accepted, but it makes clear that we apparently benefit little from 'raw data'. One of the merits of the scientific methods of recording, recoding and analyzing data is that the process is manifest, can be scrutinized by others who may adopt the same procedure or reach a well-founded conclusion that this - in our case the assessment of improvement of psychiatric patients - is not the proper way to do it. This is in sharp contrast to the methods employed by the patients in filling out their answers to the tests we provide them. What we know about their sets of minds, procedures, biases and the like, is largely by indirect evidence.

Continuing this line of basic observations, it is also obvious that a social scientist basically extracts 'data' via some sort of communication with others. We should distinguish this from the practices in the physical sciences, where data are extracted from non-living, or at least not linguistically communicating, entities via so-called *instruments*. We proudly call our checklists, tests, questionnaires, etc., 'instruments' as well. Risking the danger of a controversial debate with all outstanding psychometrists and other methodologists, I venture to doubt whether some essential distinctions between physical and social instruments have been properly taken into account to warrant the use of the instrument-metaphor in social sciences in the ways it is now done.

To substantiate my doubt I propose the following arguments. In primary or basic measurement, instruments are some sort of *extensions* or aids to our sense-organs. The basic act of perception is making distinctions by comparison (Spencer-Brown, 1969) or reacting to differences (Bateson, 1972). In measuring length we directly compare one object (the standard) with another on the same attribute. Basically this is the same as when comparing one's thumb or arm to the other object, a practice presumably abolished on the grounds that people noticed the variability in lengths of limbs between them and thus decided to use for measurement purposes the same 'limb' for all. Measuring weight (see Bateson & Bateson, 1987; p. 63), like with a simple balance, also does not really involve the *direct* measurement of a certain quantity of matter, but the comparison of one weight to the other.

In both cases of measurement the comparison - assessment of difference or equality - is made with our eyes, i.e. with our sense organs. We also owe to Bateson (1979, p. 99-146) the 'criteria of mental process', quite illuminating reading that is taken into account too little in the practice of social science. His second criterium is 'mental process requires collateral energy', i.e. living entities are always energized 'from within'. The organism's own metabolism procures the energy that enables it to perceive, i.e. to react to differences. Likewise, more advanced technical instruments - especially for the measurement of phenomena not directly accessible to our sense organs or only in a dangerous way - are quite autonomous 'extensions', in that the instrument basically operates in the same way as our sense-organs. These instruments also possess their own source of energy. A volt-meter's needle will not react to electric potential when its own batteries are dead unless it has been designed to use some of the electric energy it should measure, to do its work. To my knowledge the same basic characteristics - of being either a direct or autonomous extension of sense organs - applies to all technical instruments.

Now, what are the correspondences to this basic structure of instruments in our tests and questionnaires? Clearly, a test is no 'direct extension'. An intelligence test does not contain itself a fixed quantity of intelligence. It contains items of agreed-upon knowledge. How to measure intelligence with such a test is an intricate problem of construct validity that is never conclusively solved (see Flynn, 1987). The test also does not have its own metabolism or other source of energy and thus cannot react, by itself, to differences. So, actually, the person filling out the test is himself the 'instrument' and the test is the thing being measured by the so-called respondent. After this, the completed test, together with many like it from other persons, goes to another 'instrument' - i.e. the researcher. Where the person being tested reacted to difference in the sense of comparing some sort of inner symbolic, verbal 'reference-values' (assuming that all kinds of cognitions require a form of symbolization) to the presented external verbalizations or symbols, the researcher compares copies of the same test from different persons. The researcher does his comparing in a scrupulous and intersubjective way by reporting variabilities, reliability and so on. As it is assumed in the whole process of measurement, that similar test results reflect similar inner states in different respondents, it is also assumed that the first form of comparison is done in an albeit unknown, but essentially similar way by all respondents. There is little evidence, however for this assumption. In some cases we have some evidence that it is partly false (e.g. Guppy & Goyder, 1984; Chapman & Chapman, 1967), and we certainly can recognize factors like social desirability, performance anxiety and other factors as influencing the response to a test.

Specifically with symptom-checklists, like those used in our research, we essentially do not know in what - possibly divergent - ways people react. On what selection of personal experiences and evaluations do they base their response, what kind of motivations play a role in the answers and what effect does the respondent suppose or want his responses to have on himself, the researcher or other individuals his test-results are discussed with? This topic is especially interesting when considering the numbers of non-patients that score on the pathological side of the SCL-90, as reported by us, and the number of patients that score on the normal side.

These notes on measurement and instrumentation have been brought forward in spite of the fact that the same methodology that is being criticized has also been used in our research. In the future the author hopes to have the opportunity to either work out his doubts on instruments and psychological measurement to achieve an alternative solution or solve it in a different way for himself. In the meantime, if some of the doubt has been contagious, the reader and author alike can feel consoled with the well-known observation that research raises more questions than it answers.

Continuing the reporting of our research on consensus, I remind the reader that the whole business of assessing reliable clinical change was done in order to make a grouping of patients. This was needed for the primary interest of our study, i.e. to see if measures of consensus could predict the outcome of the therapeutic process. These main findings will now be presented in the following article.

7 Consensus in patient-therapist interactions*

7.1 A measure of the therapeutic relationship related to outcome

Abstract

This paper discusses consensus between patient and therapist as a measure of the quality of the psychotherapeutic relationship. Consensus is defined by two interpersonal concepts, agreement and understanding, and two intrapersonal concepts, congruency and experienced consensus. 162 Patients and their therapists in a Dutch Center for Comprehensive Mental Health Care were rated on the degree of consensus. Two patient self-report outcome measures and one reported by the therapist were related to consensus. The three consensus measures that predicted outcome after 6 months were understanding by the therapist as well as agreement and experienced consensus by the patient. This result suggests that a complementary patient-therapist relationship favors a good psychotherapy result. The instruments in this paper provide a means of measuring consensus during the process of psychotherapy. This is a feasible way to both monitor progress and improve the chances of a successful outcome.

The therapeutic relationship is an evolving interpersonal process that is difficult to research. While generally acknowledged as the key ingredient of psychotherapy the interpersonal aspect seldom guides psychotherapy research (Garfield and Bergin, 1978). Most psychotherapeutic research investigates the outcome of specific techniques in relation to patient characteristics and overlooks the therapeutic relationship that is indispensable to rendering any technique effective. This shortcoming is in part due to the difficulty of pinning down the crucial characteristics of a therapeutic relationship. In this paper we define interactional characteristics such as consensus and cooperation between patient and therapist and show them to be an essential feature of a fruitful psychotherapeutic relationship. Since consensus is a vague concept, we first operationalize it for practical use in psychotherapy research. Secondly, we present the results of a field study on the relationships between measures of consensus and outcome and propose a method of evaluating and improving the quality of the therapeutic relationship.

* This chapter has been accepted for publication (1989): Maarten Dormaar, Chantal I.M. Dijkman-Caes and Marten W. de Vries, *Consensus in patient-therapist interactions. A measure of the therapeutic relationship related to outcome. Psychotherapy and Psychosomatics.*

Consensus

Psychotherapy requires that the patient and therapist cooperate. This includes finding common names or definitions for the patient's problems and dividing the work that will lead to mutually agreed upon goals. A complete presentation of models and research on consensus is beyond the scope of this paper. We limit ourselves to a short outline of the sociological model proposed by Scheff (1967) as we have used this in our research. Scheff's model points out the necessity for social coordination - defined as cooperation or competitiveness - which gives rise to the seeking of consensus. Consensus, in turn, derives from interpersonal communication that creates intersubjectivity or co-orientation. In Scheff's model consensus is then equated with co-orientation, which includes two further concepts, agreement and understanding.

Agreement is defined as the zero level of co-orientation, since people may or may not be aware that they agree or disagree. *Understanding*, which is also named accuracy (McLeod and Chaffee, 1973; Kenny and Albright, 1987), is the state of 'knowing' the point of view of the other person(s) and is therefore called the *first* level of co-orientation. Further levels of co-orientation - such as knowing that the other person knows your point of view - have not been included in our research. *Agreement* and *understanding* are truly interpersonal and relational characteristics since they include the perspectives of two persons. In addition to this co-orientational definition of consensus, we also further define consensus as an *intrapersonal* concept. When two persons disagree and misunderstand each other, one or both may still have the subjective belief that agreement and/or understanding are present, and this may very well influence his behavior. By questioning his private experience of being in consensus we measure what we have called a person's *experienced consensus*. Another intrapersonal aspect of consensus is called *congruency* (McLeod and Chaffee, 1973), which is the match between one's own perspective on an issue and the presumed perspective of the other person. The different aspects of consensus-agreement, understanding, congruency and experienced consensus - are illustrated in *figure 1*.

Consensus between patient and therapist on problems, tasks and goals are prerequisites for cooperation. In this study we limit the research to consensus about problem definitions on two questions.

The first is a question of predictive validity and asks:

- 1 What aspects of consensus are related to outcome and can outcome be predicted? According to Scheff's model, consensus should be directly related to cooperation and consequently to outcome as the criterion variable.
- 2 Are maximal values of co-orientational consensus also optimal? Here we investigate whether the relationships between consensus and outcome variables are linear or curvilinear.

Figure 1: Aspects of consensus

7.3 Methods

Subjects and setting

The research was carried out in a Dutch center for comprehensive ambulatory mental health care serving a population of 200,000 and registering over 2,500 new patient systems per year. The centre has 8 different treatment teams of which two short-term treatment teams participated. Subjects included (a) all new patients of the two participating teams during the research period of 12 months who satisfied criteria of informed consent, age (between 18 and 60 years), no contra-indications (psychosis or mental retardation) and at least 4 contacts with the same therapist and (b) all therapists ($n = 22$) of the same teams.

No significant differences on demographic variables were found between the research sample ($n = 162$), the patients who refused to participate or were contraindicated ($n = 254$), or those who dropped out within 3 sessions ($n = 174$). The majority (58%) of our sample was female and the mean age was 35 years. In most cases (86%) the patient system was an individual. Most patients were married (65%) and lived with their family (68%). Approximately one-

third of the patients reported that they had suffered less than a year from their complaints, another third between 1 and 5 years and less than one-third more than 5 years. The wide range of presenting problems is depicted in *table 1*. Although diagnoses were not always assigned, most patients suffered from neurotic and adjustment disorders.

Table 1: Main problem areas according to patients and therapists* in percentages relative to the total number of problem areas (n = 324) mentioned by the patients themselves and the total number (n = 399) mentioned by their therapists.

Problem areas	Therapists	Patients
Interpersonal relations (outside family)	12.7	14.5
Fear & anxiety	17.6	12.3
Education, leisure & household	11.1	9.0
Family & marriage	19.7	19.0
Somatic complaints	15.7	10.3
Decision-making and self concept	7.1	13.3
Coping with trauma & problems of development	6.5	11.3
Total of 3 remaining areas (addiction, depression, practical problems)	9.6	10.3

* Chi square 20.0, $p < 0.01$.

Instruments and analysis

experienced consensus was measured with a 5-point checklist of 8 items regarding the subjective experience of patients and therapists of understanding the other party and of feeling understood.

As shown in *figure 1*, the consensus variables agreement, understanding and congruency are relationships between measures. The measures themselves are from the interpersonal perceived consensus (IPC) list. This list requests the

opinions of patient and therapist (agree/disagree) regarding 8 problem definitions and also their presumed opinions of the other on these same problem definitions.

Agreement, or the degree of correspondence between the opinions of the patient and his/her therapist, is computed with the similarity coefficient of Jaccard (Jaccard, 1908; Snijders et al., 1988), ranging from 0 (no agreement) to 1 (maximal agreement). Understanding and congruency are likewise computed. The variability of the consensus measures* is shown in *table 2*.

Table 2: Variability of consensus measures

Consensus measure	Range Mean SD		
Experienced consensus			
experienced by patients	8-24	19.1	3.8
experienced by therapist	8-24	17.5	3.8
Interpersonal perceived consensus			
Agreement	0-1	0.49	0.27
Understanding by patients	0-1	0.49	0.28
Understanding by therapists	0-1	0.48	0.27
Patient-congruency	0-1	0.75	0.27
Therapist-congruency	0-1	0.59	0.32

To measure outcome the Target Complaints List (TCL) (Battle et al., 1966) and the SCL-90 (Arrindell and Ettema, 1981; Derogatis, 1977) were used. The TCL discomfort ratings as reported by the patient himself (TCL-C) and as judged by his therapist (TCL-T) range from 0 (no discomfort) to 10 (maximal discomfort). Two or more ratings were averaged to one score per patient and per therapist. The TCL permits the evaluation of progress on individualized complaints whereas the scoring system gives results comparable across subjects (Kiesler, 1977). The Symptom Checklist SCL-90 is a 90-item multidimensional indicator of psychopathology sensitive to changes in the level of subjective distress that has been used earlier in short-term psychotherapy outcome research (Boelens and Debats, 1983; Emmelkamp et al., 1985). The scoring ranges from 90 (no distress) to 450 (maximally distressed).

* A copy of the consensus instruments, instruction manual and computation of measures can be obtained from the first author.

Procedure

Before the first contact with his therapist each patient had to fill in the symptom checklist SCL-90. At this time the demographic data and the duration of his problems or complaints were also assessed. After the second contact with the same therapist both patient and therapist were separately requested to fill out the consensus instruments and the TCL.

Finally, the posttest, composed of a second discomfort rating of the TCL (for patient and therapist separately on the problems or complaints as reported previously) and of a second administration of the SCL-90 (for the patient), was taken 6 months after the beginning of treatment or earlier if the treatment had been terminated already. The mean number of sessions at posttest was 10 (range: 6-20).

Results

Firstly, none of the linear (product-moment) correlations between consensus measures and outcome defined as raw difference scores of the SCL-90 and TCL-P were significant. For the TCL-T, only experienced consensus by patient was significantly correlated ($r = 0.18$, $p < 0.05$, $n = 143^*$). No significant correlations were found between consensus measures and residual change scores of the outcome measures.

Curvilinear relationships between consensus and outcome were investigated to test the hypothesis of *optimal values* for the measures of interpersonal perceived consensus, especially agreement. For this purpose the scaling of consensus was transformed following the method of Green et al. (1985): the mid-value of 0.50 was arbitrarily taken as an optimal end point and higher scores were transformed to scores less than this mid-value but at equivalent arithmetic distance from it. Thereafter each IPC measure was recoded into a variable with 3 values: high (between 0.40 and 0.50), middle (between 0.20 and 0.39) and low (between 0.00 and 0.19). Coded in this way significant correlations were found for agreement with outcome ($r = 0.23$, $p < 0.01$, $n = 141$) as assessed with the SCL-90 (raw difference scores) and for therapist-congruency with the same outcome measure ($r = 0.25$, $p < 0.01$, $n = 141$). This means that mid-level agreement and mid-level therapist-congruency are associated with a better outcome as assessed by the Symptom Checklist.

* A case is represented by the scores of patient and therapist on their *individual* measures of consensus and by the *combined* scores of the patient-therapist dyad on their *relational* measures of consensus.

The relationships between consensus and outcome were further explored with a one-way analysis of variance (ANOVA): postscores of the outcome measure were used as the dependent variable and prescores as covariate. The independent variables consisted of 5 levels of interpersonal perceived consensus measures and 4 levels of experienced consensus. SCL-90 postscores varied significantly with the understanding by the therapist ($F_{4,130} = 2.626$, $p < 0.05$; $\eta^2 = 0.095^*$) and with the experienced consensus of the patient ($F_{3,130} = 2.835$, $p < 0.05$; $\eta^2 = 0.078$). The adjusted marginal means as presented in the multiple classification index of the ANOVA reveal that the highest level '5' of understanding and to a lesser degree level '2' were associated with a better outcome. For experienced consensus high levels only were associated with a better outcome. The postscores of the TCL discomfort rating by therapist again varied significantly with the degree of understanding by the therapist ($F_{4,130} = 3.358$, $p < 0.05$), as well as with patient-congruency ($F_{4,130} = 2.613$, $p < 0.05$). In the multiple classification index both effects showed a positive relationship with outcome: again the lowest level of understanding associated with a slightly better outcome than the grand mean and the highest level with the best outcome. For the patient congruency the higher levels were associated with a better outcome. Analysis of variance of the postscores of the TCL discomfort rating by patient showed no significant main effects of the independent variables. Because of the great number of levels in the independent variables no interaction effects were explored in the analysis of variance. Finally, discriminant function analysis was used to predict outcome on the basis of predictors, i.e. the consensus measures.

The cases were grouped according to outcome level as follows: group 1: improved, i.e. significant positive change on at least two outcome measures; group 2: not improved, i.e. no significant positive change on any outcome measure. Cases with only one significant positive change were not classified ($n = 48$).

The criterion for significant positive change for the SCL-90 was *reliable clinical change*. As reported elsewhere (Dormaar et al., 1988) this two-fold criterion derived from Jacobson et al. (1984) determines improvement both as clinically significant and reliable. An improved patient shows a level of functioning at posttest that is statistically more likely to be in the functional than

* η^2 is a measure for the proportion of the variance in the dependent variable that is shared with or predictable from the independent variable. It is therefore a measure for the strength of association.

in the dysfunctional range, and the probability that his change from pre- to posttest can be attained by chance is less than 5%. The SCL-90 satisfies the psychometric demands of this criterion.

For the TCL outcome measures, raw change scores above one standard deviation of the change distribution were classified as significant positive change and the remainder as no significant change. After elimination of multivariate outliers and cases with missing predictors or outcome scores 81 grouped cases remained. The results of the discriminant analysis are shown in *table 3*.

Table 3: Classification results of prediction of outcome by consensus measures

Actual group	n	Predicted group membership	
		not improved	improved
Not improved	45	26 (57.8%)	19 (42.2%)
Improved	36	10 (27.8%)	26 (72.2%)

Percent of 'grouped' cases correctly classified: 64.2%.

Three out of the eight predictors were accepted with Wilk's stepwise method: experienced consensus by patient, understanding by the therapist and agreement. These are the same consensus measures that showed most of the correlations with outcome in the previous results. The canonical discriminant function had a Wilk's Lambda of 0.85, chi square 11.7, significance (d.f. = 3): $p < 0.01$. This means that the function has a highly significant discriminating power. *Table 3* shows the classification results without prior probabilities. The overall percent of correct classification (64.2%) is rather low, but quite satisfactory for this small number of cases.

Discussion

The association between mid-level *agreement* and successful psychotherapy may be attributed to the fact that a moderate level of agreement between patient and therapist regarding the problem areas to be dealt with gives evidence of a creative tension between their views. Maximal agreement indicates that their

views are identical and that the therapist has no new options to offer, and minimal agreement means that patient and therapist are living in completely different worlds.

Likewise, mid-level *therapist congruency* associated with the best outcome indicates that his patient stands a better chance if the therapist can tolerate a moderate discrepancy between his own views on the patient's problems and the views he supposes to be held by his patient. Most therapists have a rather high score on congruency. So they tend to perceive their patients too much like themselves to reach this optimal level of cognitive dissonance.

Of the measures of consensus reported by the patient, consensus experienced by the patient and patient congruency are positively associated with outcome. High consensus experienced by the patient reflects his subjective feeling of mutual understanding between himself and his therapist, and high congruency reflects his idea that his own views on his problems and those of his therapist are well-matched.

Although the level of accurate prediction in the discriminant analysis is moderate, it is quite remarkable that outcome assessed after 6 months of psychotherapy can be predicted reliably on the basis of the consensus measures that were taken after the first two therapy sessions. Two out of the three predictive consensus measures, understanding by therapist and consensus experienced by the patient, characterize a favourable patient-therapist relationship that can be typified as a complementary one: the patient *feels* understood and assumes that he understands his therapist, the therapist *understands* his patient correctly. The discriminant analysis shows that the complementary relationship favors a positive outcome, although a minority can do without it. 72% of the improved patients have this type of relationship with their therapist.

Some limitations caution against a more definite interpretation of our results. First, a plethora of intervening variables could not be controlled for in this field study, specifically demographics, diagnosis, duration of complaints, expertise and technique of the therapist. The number of cases precludes a full investigation in a multivariate analysis. Further research will focus on manipulation of consensus. This will enable us to answer questions like how different therapists and patients achieve consensus and how levels of consensus develop during therapy. Second, a limitation of the consensus model is that it presupposes motivation to coordinate actions (Scheff, 1967). Therapists know by experience that their patients do not always show that motivation. Finally, therapist and patient may encounter problems that they are unable to solve, notwithstanding their full cooperation.

The results of this study represent a step towards a very desirable goal in psychotherapy and research: to have at one's disposal practical instruments that monitor the process and progress of psychotherapy and that take into account the perspectives of both parties concerned. Contrary to the desired

state of affairs - well defined methods and procedures for circumscribed groups of patients and/or disorders - daily practice shows a rather unmethodical mixture of therapeutic attempts to deal with patients with a rather idiosyncratic pattern of reactions. To bridge the gaps between clinical practice and research, we need instruments that tap factors that are really present, that are suited for the psychotherapy situation 'in reality' and not only for a laboratory-like 'ideal' situation. Consensus reflects a clinical reality of working together by patient and therapist. The easy administration and scoring of the instruments as developed make them suited for further research, for evaluation of clinical practice and for feed-back of the results to patient and therapist in order to improve their efforts in cooperating toward their common goal of a successful outcome.

Acknowledgement

The research was supported by the NFGV (Dutch National Foundation of Mental Health), grant No. 1.816 and by the University of Limburg. The authors are grateful to participating patients and therapists of the 'RIAGG Maastricht e.o.' (Director: H.B. Pomerantz) as well as to M.A.J. Romme, Chairman, Department of Social Psychiatry, J.J.C.B. Bremer, Chairman, Department of Medical Psychology and M.A. van den Hout, Chairman, Department of Mental Health, University of Limburg, for their support.

Comment

In retrospect, it is not difficult to have some methodological reservations about the study as reported here. For instance, the number of patients with a complete data set was but a subset of all subjects originally available. This may cast doubt on the validity and generalizability of the results, although we could report some superficial checks on possible sources of selection. At least part of the incompleteness of the data-set can be explained by the fact that at the time the study was executed (1983-1984), research was still a rather 'foreign body' to the mental health care institution where the investigation has been done. We have tried our best to achieve some control over procedures and to get all therapists involved. Certainly most of them were faithfully co-operating, although research remained a minor interest and sometimes a nuisance to them. Despite the recently more institutionalized forms of cooperation between university and mental health care centers, the questions of how to introduce research and how to divide responsibilities and control over the procedures to be followed remain important. The fact that both patients and therapists are questioned on sensitive subjects they are very personally involved with puts a special responsibility on the researcher to create a 'consensual domain' in which all parties concerned can cooperate wholeheartedly in mutual trust and respect.

Secondly, we have skipped over the implications of the circumstance that some therapists participated up to twenty times, whereas others were included only once or twice. We have treated every patient-therapist dyad as a unique specimen. It is obvious, however, that a therapist responding to our questionnaires a couple of times is likely to be influenced in his opinions not only by the information of the patient of the particular moment, but also by his own preconceptions. This means that he will not behave independently from previous occasions. It is therefore unanswered how much 'constraint' exists in the data that may have influenced the results.

Finally, a tricky point will be mentioned briefly, namely to what extent the results have been 'fabricated' or reflect the real state of affairs. I think that all who have been practically involved in data analysis are aware of manipulating their raw data - which are not that 'raw', objective or virginal in the first place - in many different ways that cannot all be reported. Also, in using advanced techniques of statistical analysis, most researchers, being largely unacquainted with the presuppositions and implications of the mathematical tools of calculus and matrix calculations, somehow remain sourcerer's apprentices, basically unsure about creating a miraculous kind of order or actually havoc. This being the case, I consider the problem of 'fabrication or reality' as

basically unsolvable. In previous sections I have already embraced the constructivistic position, which holds that we cannot but create our reality, that furthermore the objectivity of this creation depends on the degree to which that claim is endorsed by others and that, in turn, this endorsement depends on the consensual domain the person claiming that objectivity has been able to establish (see also Maturana, 1988).

Nevertheless, and also notwithstanding the satisfying and promising results, the self-imposed criticism cited above amounts to the intention to do it next time differently and hopefully better. As to the actual procedures followed, for instance, it is a pity that consensus could be assessed only once. It would have been very interesting to see how the different aspects of consensus developed, both in successful and unsuccessful therapies. Finally, the relationships of consensus to outcome have been called a question of *predictive* validity. Actually, it is more appropriate to speak here of *construct* validity, since the theory predicts the relationships. The next chapter also concerns an aspect of the construct validation of consensus. If consensus rests on the reciprocal influencing of patient and therapist, then the absence of consensus means no reciprocal influence, that is a responding to the 'interpersonal consensus checklist' (IPC) by patient and therapist uninfluenced by, or independent from, the other. To test the hypothesis of reciprocal influence leading to consensus, we have to establish:

- 1 Whether the distribution of consensus-values of patient-therapist dyads - who supposedly went through a process of mutual influencing during their first two contacts - differs significantly from the distribution of dyads without such a process, and
- 2 How many patient-therapist dyads have reached a level of consensus after their first two contacts that is improbable under conditions of independence.

As the null-hypothesis of independence could be most easily operationalized for agreement, we decided to tackle this problem of construct validation for this level of consensus.

Although a *high* degree of agreement does not have to be the most typical characteristic of the kind of consensus that favours successful outcomes in psychotherapy (see p 103), it can certainly be maintained that the distribution of agreement of patient-therapist dyads should show a higher mean than to be expected under independence, and that the number of dyads showing a high degree of agreement should be significantly higher than to be expected from the independent distribution.

8 Consensus in terms of similarity*

When two persons are claimed to be in consensus - by themselves or others - this means that these two persons are considered similar in certain respects. In the case of our research, when asserting that a therapist and patient are in consensus we mean that therapist and patient resemble each other with respect to their definitions of the problems for which the patient is asking help. When focused on agreement, we mean that they both use the same definitions and, when focused on understanding, we mean that the one's estimation of the other's problem definitions is the same as that other's real definitions. When speaking of resemblance or sameness, we speak of similarity. Therefore, establishing consensus is a matter of establishing similarity, which is an underlying principle for classifying objects and making generalizations (Tversky, 1977). Many different forms of similarity and dissimilarity data are used in social research, as in the concepts of communality, correlation and diversity. In our consideration of the concept of similarity, two topics will have to be discussed:

- 1 the models of similarity and
- 2 the assessment of levels of significance.

The presentation of the first topic will be based on Tversky (1977). He first points out that similarity is most commonly based on a geometric model. In such a model, similarity and dissimilarity (difference) correspond to the *metric* distances between the two points that represent the objects to be compared in some coordinate space. The coordinate space implies that objects can be represented through a few quantitative *dimensions*. From the metric and dimensional assumptions follow the axioms of minimality, symmetry and triangle inequality. These axioms, however, are not always valid when we look at the ways we actually make comparisons between objects or judgements of similarity and difference. The first axiom implies that the similarity between an object and itself is the same for all objects. But it appears that the probability that two identical stimuli will be judged as 'same' is not constant for all objects. Secondly, symmetry also does not hold in all cases. Similarity judgements in the form of 'a is like b' make this clear. We are more likely to say 'an ellipse is like a circle', than 'a circle is like an ellipse', or 'the son looks like his father', instead of 'the father looks like his son'. Sometimes the reversed direction of comparison even carries a different meaning, as in the metaphors 'a man is like a tree' and 'a tree is like a man'. Likewise, there are cases in

* This chapter is partly based on: T.A.B. Snijders, M. Dornaar, W.H. van Schuur, C. Dijkman-Caes and G. Driessen (1989), *Journal of Classification*, accepted for publication.

which the third axiom is violated. This axiom holds that if 'a' is similar to 'b' and 'b' is similar to 'c', then 'a' cannot be very dissimilar from 'c'. The violation of the axiom appears, for instance, when comparing countries. We might say 'Jamaica is like Cuba', referring to their geographical proximity, and also 'Cuba is like Russia', meaning their political affinity. But Jamaica and Russia are not similar at all.

The foregoing reasoning leads to a description of similarity based not on a few quantitative dimensions, but on many qualitative features. Features denote values of binary or nominal variables, but they can easily include qualitative and quantitative dimensions - i.e. higher order variables - by means of a set-theoretical representation. In this way similarity of 'a' to 'b' is expressed as a function of three arguments: the features that are *common* to both 'a' and 'b', the features that belong to 'a' but not to 'b' and the features that belong to 'b' but not to 'a'. These last two kinds are called the *distinctive* features. The idea of 'feature matching' has been worked out in a number of assumptions and theorems, whose validity has been illustrated and mathematically justified.

Feature matching leads to two models of similarity: the contrast model and the ratio model, in which the assumptions and axioms of the geometric model are no longer needed. The contrast model gives a scale S that expresses similarity as a linear combination, or a contrast, of the measures of the common and distinctive features. The ratio model presents a scale S that is a matching function expressed in set-theoretical terms as follows:

$$S(a,b) = \frac{f(A \cap B)}{f(A \cap B) + \alpha f(A - B) + \beta f(B - A)} \quad \alpha, \beta \geq 0.$$

'f' Reflects the salience or prominence of the various features. If all features are regarded as equally prominent and if $\alpha = \beta = 1$, this matching function is equal to the Jaccard index of similarity (Jaccard, 1908) that was used in our research. Most of the remainder of Tversky's article is a discussion of how the scale 'f' and the parameters α and β depend on the context and the task of making comparisons.

In most research of which we are aware, consensus is assessed by means of correlation coefficients or analysis of variance techniques. The main objection to such methods is that they are measures of the association between *variables*, not between *units*, e.g. persons. Furthermore, two variables can be highly correlated without similarity in response. If consensus is conceived as similarity between persons over a set of attributes, then these methods are basically unsuited. The problems in assessing levels of significance for similarity coefficients may be one of the main reasons why such coefficients are not yet

widely used in the social sciences, for measuring consensus or other aspects of association or relationship between individuals.

The second topic, that of assessing levels of significance, will therefore be elaborated on the basis of the article by Snijders et al. (1989) that originated out of our endeavor to confront the problem of how to evaluate levels of consensus, especially agreement.

Using the Jaccard index for the measurement of consensus

Agreement may have cultural or interpersonal origins. In the first case, two persons hold the same opinions quite accidentally: they share some beliefs or attributes without having been in contact with each other. In the second case, they agree because they have exchanged views on the subject-matter and thus influenced each other's previously held opinions. The latter situation is called 'co-orientation through communication' in Scheff's consensus model, where this process is postulated as the basis of interpersonal consensus. If psychotherapist and client have discussed the nature of the latter's problem, as is generally the case in the first therapeutic contacts, communication has explicitly taken place.

We have used the Jaccard index to measure consensus between a psychotherapist and his client with respect to the definition of the client's problem(s). attributes here are specific formulations of possible problems. If both therapist and client 'agree' that certain attributes (problem formulations) do apply to the client, then they exhibit agreement or 'zero-level co-orientation' (Scheff, 1967). If, however, both therapist and client agree that certain attributes do not apply to the client, then such agreement should not contribute to consensus between therapist and client. Therefore, the J index is appropriate here; it avoids the 'common deficiencies' problem present in many other coefficients (Gregson, 1975; p. 48).

In our research on consensus (Dormaar, Dijkman-Caes, and De Vries, 1989), we used data based on the responses of psychiatric clients and their therapists to a list with eight problem formulations. Clients and therapists were asked to state for each problem formulation whether they considered that formulation valid for the client. The 162 clients filled out the list after the second therapy session, as did their therapists. There were 22 different therapists who participated between 2 and 12 times in this research and who completed the list of problem formulations for a total of 115 of the 162 clients. For this sample of 115 client-therapist dyads the mean J-value was 0.49 and the standard deviation 0.21.

We hypothesize that a distribution of consensus values in a population of dyads, each consisting of the dyad 'client and client's *own* therapist', will have a higher mean value than the distribution of these values in a population of

'client - *arbitrary* therapist' dyads. We expect that within the first type of dyads co-orientation has occurred, which will have led to more agreement and higher levels of consensus than within dyads of the second type. The latter distribution corresponds to the null hypothesis formulated above, that clients and therapists respond to the problem formulations independently of the responses of their partner.

Empirical rejection of this null hypothesis provides some support for the model of interpersonal consensus outlined above. Moreover, the availability of the probability distribution under the null hypothesis of statistical independence will enable the researcher to select those dyads which exhibit evidence of having gone successfully through the process of co-orientation. These dyads could be defined as those reaching a degree of consensus that is highly improbable under the null hypothesis of mutual independence, operationalized for instance as J-values that have p-values under statistical independence lower than 5%.

As a first approach to the question of which values can be interpreted as a significant level of consensus, a simulation study was carried out, analogous to Johnson and Millies's (1982) procedure for finding confidence intervals for Stander's similarity index. A dataset of 10,000 independent dyads was randomly simulated on the basis of the marginal probabilities of the eight problem formulations for clients and therapists separately. The simulation model assumed no association among the attributes (problem formulations). In this case the mean J-value was 0.38 and the standard deviation 0.19.

In a second simulation approach all 115 response vectors of clients to the eight problem formulations were combined with all 115 therapist response vectors. This results in a total of $115 \times 115 = 13,225$ simulated J-values, among which 115 are genuinely observed and the other 13,110 are hypothetical. The entire set of 13,225 simulated dyads had a mean J-value of 0.39 with a standard deviation of 0.22. The difference with the genuine dyads is considerable, but cannot be legitimately tested with the t-test, because the assumption of independent observations is not valid.

Some coefficients for binary data

In general, similarity between pairs of operational taxonomic units (or 'units', for short) can be defined in terms of the common presence and absence of attributes. coefficients of similarity based on multivariate binary data are widely used in the field of taxonomy (Sneath and Sokal, 1973; Everitt, 1980; Anderberg, 1976), in biology (Washington, 1984), and in other disciplines (Austin and Colwell, 1977; Hubalek, 1982). These coefficients are often used to convert a two-mode matrix (e.g., units by attributes) to a one-mode (e.g., units by units) matrix of correspondence between pairs of units. Here we restrict ourselves to three of the better known similarity coefficients: the Matching coefficient, M

(Sokal and Michener, 1958), the Jaccard index, J (Jaccard, 1908), and the Dice coefficient, D (Dice, 1945; see also Dice, 1952; also called Czekanowski-Dice coefficient by Wishart, 1978). These are defined by

$$M = \frac{a+d}{a+b+c+d} \quad J = \frac{a}{a+b+c} \quad D = \frac{2a}{2a+b+c}$$

where a , b , c , and d refer to the contingency table of attributes present (1) or absent (0) among n attributes, for two units:

Table 1: Contingency table of dichotomous attributes for two units.

		Unit B		
		1	0	
Unit A	1	a	b	$a+b = N_A$
	0	c	d	$c+d$
		$a+c = N_B$	$b+d$	$a+b+c+d = n$

The similarity coefficients J and D are increasing functions of each other:

$$J = D/(2-D), D = 2J/(1+J)$$

Thus J and D express the same information in numerically different ways. Our discussion of similarity coefficients is concerned with testing their values under a null model of absence of association between units A and B .

According to each coefficient M , J , and D , two units are more similar to the extent that a larger number of attributes is jointly present in both. Each coefficient ranges from 0 to 1. The difference between J and D on the one hand and M on the other lies in the treatment of joint absence of attributes in both units. These 'negative matches' are ostensibly ignored in J and D , but used in M . The Jaccard index was first proposed by Jaccard (1908) in order to remove the artifact that two units - ecological sites - would be considered similar because *neither* contains many of the attributes (e.g., the plant species under consideration).

Comparing observed values of similarity coefficients with expected values under statistical independence

In applications of similarity coefficients, substantive theory often predicts similarity between specific units (individuals, species, sites) to be higher than between arbitrary units. This situation will cause the researcher to test the data against a null hypothesis of independence. However, knowing that a similarity coefficient is higher than expected under the null hypothesis is not enough; we want to know whether it is *significantly* higher. In the absence of criteria for significantly high values of similarity, some researchers (e.g., Quesada, et al., 1983) resort to using arbitrary levels (e.g., 0.65) as a lower boundary for an interpretable level of similarity. We will here present the procedure for a significance test.

Coefficients such as M, D, or J are not only sensitive to similarity between the two units, but also to other characteristics of the data. Since the aim of using these similarity coefficients is to study similarity within dyads, it is necessary to control suitably for factors which could influence these coefficients, and thereby may cause problems in evaluating their observed values. Two of these factors which should explicitly be taken into account are discussed first.

The first is that some attributes may be present more frequently than others. If so, average values of similarity coefficients will be higher than if all attributes have equal relative frequencies. When units A and B in (2) are drawn from two different populations (e.g. from populations of therapists and clients, as presented above) a further complication arises. If the frequency orders of attributes are the same in both populations, the average values of similarity coefficients will be higher than if the frequency orders are different. Differential frequency of attributes, and different populations of units A and B, do not constitute serious problems, but they do need to be taken into account.

A second factor results from the association *within* units among the given attributes. The attributes used in applications of M, D, and J may be related in a meaningful way; in some applications they can be regarded as different indicators of one or more underlying latent concepts. When the attributes are associated from a conceptual point of view, they will be mutually statistically dependent in the population(s) of units. This dependence affects the probability distributions of M, D, and J. We may expect that association among attributes will affect the standard deviations of the similarity coefficients, since it affects the standard deviations of the numbers N_A and N_B of attributes present. This problem is more serious than the first one and has hardly been addressed in the literature we know, which commonly assumes, explicitly or - more often - implicitly, that under the null hypothesis the attributes are mutually independent within units. We propose to take these problems into account, first for the Simple Matching coefficient, and subsequently for the Jaccard index, using the Dice index as an intermediate step.

These considerations lead to the following formulation of the null hypothesis, relevant in our opinion for testing an observed value of a similarity coefficient. A population of dyads of operational taxonomic units is postulated; either the members of each dyad are *a priori* indistinguishable, or the members of each dyad can be distinguished as one being of type A and the other of type B. For a given set of n attributes, the presence (denoted by 1) or absence (denoted by 0) of each attribute can be observed for both members of the dyads investigated. There may be association (i.e., statistical dependence) among attributes within units (see Coleman, Mares, Willig, and Hsieh, 1982, for a formulation of independence in which association is assumed to be absent). If units of types A and B are distinguished, then the joint probability distribution of attributes present may differ between population A and population B. The null hypothesis is *independence among units* in every single dyad under investigation.

If sufficiently large samples of units A and of units B are available, the joint probability distribution of presence of the attributes within either population of units is known. Parameters of the probability distribution of the similarity coefficients M , D , and J will be derived under the null hypothesis. The null distribution of many other similarity coefficients can be investigated using the same approach. The null distribution of similarity coefficients in the situation where no distinction between units A and B is made is identical to the null distribution in the situation where such a distinction is made, but the joint probability distribution for type A units is the same as for type B units. Therefore, no generality is lost by formulating all results for the case where the distinction between both types is made.

Snijders et al. (1989; section 4) first discuss earlier approaches for calculating the mean and variance of the null distribution (the distribution under the null hypothesis of independence between units within dyads) of several similarity coefficients. In this chapter we limit the presentation of results to the approximations for the mean and variance of the null distribution of J :

$$EJ = ED / (2 - D) \approx ED / (2 - ED) + 2(2 - ED)^{-3} \text{ var } D .$$

For this approximation of EJ the second order Taylor expansion has been used. For the approximation of $\text{var } J$ the first order Taylor expansion of ED yields:

$$\text{var } J \approx 4(2 - ED)^{-4} \text{ var } D .$$

The mean and variance of the null distribution of the Dice coefficient D (ED and $\text{var } D$) have been calculated in a mathematically exact way that allows for different types of units within the dyads and for association of attributes within

units. For the mathematical expressions of ED and var D, I refer to Snijders et al. (1989), section 6.

In order to make a probability statement for assessing the significance of a given observed value of J, it is necessary to evaluate the cumulative distribution function (cdf), or at least to approximate it. Exact expressions have been derived for the mean and variance of D, but only approximate ones for the mean and variance of J. Therefore, the exact expression of ED and var D are used for the approximation of the cdf of J. Taking account of the boundedness of the interval [0,1] of possible values of J and D we favour the approximation by means of a Beta distribution instead of a normal distribution. The distribution of D can be approximated by a mixed distribution with a probability π for the discrete outcome $D=0$ and a probability $(1-\pi)$ for the Beta distribution of the remaining outcomes. From this distribution approximate cumulative probabilities for J can be calculated with a formula for the correspondence between the cdf's of J and D (see Snijders et al., 1989; section 7).

Elaboration of the example of consensus assessment

The values found for J in the study of consensus between psychotherapist and client, discussed briefly in Section 3 of Snijders et al. (1989) and more fully in Dormaar et al. (1989), were investigated under the null model defined above. The parameters of the mathematical expressions for var D and ED were estimated from the sample of 115 psychotherapist-client dyads, for which complete data were available. This implies that the estimated null distribution of J coincides with the distribution of the 13,225 simulated J-values obtained by taking all pairs of any psychotherapist and any client, which was referred to above. This distribution thus can be regarded as a permutation distribution. Thus, given that estimated probabilities are used, the probability distribution of J, for which the approximations have been computed as outlined above, is simply the permutation distribution. The availability of the permutation distribution can be used to study the values of the various approximations in this particular case.

Under the permutation distribution, results of the Dice coefficient were $ED = 0.521$ and $\text{var } D = 0.0555$ and for the Jaccard coefficient $EJ = 0.386$ and $\text{var } J = 0.0476$. Approximations for EJ and var J that use the exact mean and variance or the Dice coefficient as an intermediate step, yielded 0.386 and 0.0463, respectively. The good performance of this approximation is striking.

When particular J-values are to be tested under the null model of absence of consensus between psychotherapist and client, formulated statistically as independence between the attributes chosen by the client and his/her psychotherapist while allowing for association of the attributes chosen by each single person, the upper 5% level of the null distribution would be a reasonable threshold value. In the permutation distribution of the J index, this is beyond the value $J = 0.75$. Relevant null probabilities are:

$P \{J \geq 4/5\} = 0.046$, $P \{J \geq 3/4\} = 0.063$, and $P \{J \geq 5/7\} = 0.084$; there are no possible outcomes for J between $5/7 \approx 0.71$, $3/4 = 0.75$, and $4/5 = 0.8$. From the approximation of the cdf of D by a mixed discrete - Beta distribution the following probabilities were derived:

$$P \{J \geq 4/5\} \approx 0.040, \quad P \{J \geq 3/4\} = 0.064.$$

The correspondence with the values obtained from the permutation distribution is quite satisfactory.

Of the 115 true client-psychotherapist dyads, 16 had a J -value at least equal to .75, while 27 had a J -value at least equal to .71. From a statistical point of view, these frequencies of occurrence are considerably more than the 6.3% and 8.4% which were to be expected under independence. From the clinical point of view, however, the number of dyads where the obtained value of J suggests that consensus-after-communication occurred is rather disappointing. Tentatively, we conclude that the Jaccard index based on 8 attributes is not a powerful method for establishing consensus.

One of the authors of the original publication (G. Driessen, see Snijders et al., 1989) has made available a Fortran computer program for calculating the means and variances of the null distributions of the similarity coefficients J and D . The use of this program can help other researchers to test the significance of their observed similarity data.

Discussion

First, I have argued for conceptualizing consensus as a form of similarity, according to Tversky's (1977) ratio model. Second, a method for testing the significance of observed values of similarity coefficients, notably the Jaccard index, has been presented by means of an outline of the original publication (Snijders et al., 1989). With sufficiently large samples of units available an estimated mean and variance of the similarity coefficient can be computed, and the null distribution of the similarity coefficients D and J can be quite well approximated by Beta distributions that should include an extra discrete probability for the outcome 0. The method used here for determining the mean and variance of J and D can easily be extended to similar coefficients.

Methodologically, the message of this chapter is twofold. First, the null mean and standard deviation of similarity coefficients will in practice often be large, unless the number of attributes is great. As a consequence, observed values of the coefficients can be interpreted as contradicting the null model only if they are extremely high. In our example with 8 attributes, the 5% threshold value for the Jaccard index was as high as 0.75. This result suggests

that it is difficult to draw statistically significant conclusions if one wants to establish 'consensus' on the basis of only 8 attributes. Increasing the number of attributes with a given average probability of presence will decrease the variances of M and J. Users of these measures of association, which are so prevalent in methods of classification, ought to be aware of these limitations. Second, one should take account of the possible association of attributes within taxonomic units. Such an association may strongly affect the null distribution of similarity coefficients.

I have operationalized agreement - as an aspect or level of consensus - by means of a similarity coefficient. The question may be raised, however why an agreement index like Cohen's kappa has not been used. Considering the vast literature on both similarity coefficients and agreement indices, this question will be summarily dealt with.

I take notice of the following points on 1) basic procedures and assumptions and 2) methodology in relation to types of research questions, that explain the preference for a similarity coefficient in our research on consensus.

- 1 Cohen's (1960) kappa is basically a measure of agreement between observers, in most cases between *pairs* of observers (Popping, 1983). Typically, two independent observers judge a set of observational data according to a nominal classification scheme of both exhaustive and mutually exclusive categories. Also, judgements of the observations from the set are considered to be made *independently* from each other. Cohen's kappa accounts for chance agreement, defined as the sum over all categories of the product of the (probability) fractions per category of the two judges. To the extent that both - or all - observers assign the same categories to the same observations, the assignment can be considered to be independent from the particular observers. A significant kappa - when assessing inter-observer agreement according to the usual a posteriori method (Popping, 1983, p. 3) - therefore is a necessary condition for reliability and interrater validity of the classification scheme.

A similarity coefficient is basically a measure by which an observer - e.g. the researcher - assesses the similarity between two observational units with a fixed classification scheme of dichotomous categories (presence or absence of attributes) chosen by the observer of the units. The categories are neither exhaustive nor mutually exclusive and there even may exist association among attributes. A method to correct for chance similarity under these assumptions has been presented in this chapter.

- 2 In addition to the differences in basic assumptions between interrater agreement and similarity the following methodological considerations are in order. One might use Cohen's kappa for assessing agreement on problem formulations between a specific patient and therapist, by applying kappa to the 2 x 2 contingency table of presence/absence data (see table 1). However, the correction for chance agreement would then be incorrect because it would be based

on the unwarranted assumptions of equal probability of each problem formulation and of considering problem formulations as independent replications of each other.

Alternatively, if we take one specific problem formulation and calculate its kappa coefficient for all pairs of observers (i.e. patient-therapist dyads), no statement on the significance of the level of agreement over *all* problem formulations of any *specific* patient-therapist dyad would be possible (as is needed in our research), since the results would merely show a distribution of kappa coefficients over all pairs of observers per individual problem formulation.

In our use of a similarity coefficient, notably the Jaccard index, the question is not one of testing interrater reliability or agreement. Basically, the researcher offers a set of possible problem formulations - neither exhaustive, mutually exclusive nor independent - and requests his research subjects (i.e. patients and therapists) to use this 'classification scheme' to formulate the problems that exist in their respective 'reality domains-in-language' (Maturana, 1988). If a specific patient and therapist use the same problem formulations to denote a reality that *depends* on descriptions in language (problems only exist in the eye of the beholder), then they have a shared reality (in language) regarding these problems, at least for the researcher - who in turn is a dependent observer through his presentation of the list of problem formulations and calculation of their similarity coefficient. Reducing the several problem formulations to a count of the number of problems present would be quite alien to the kappa coefficient. Moreover, the basic philosophy behind kappa is that it is a viable task to categorize the reality of observations *independent* from the observers. kappa coefficient testifies to the measure in which this task has been carried out successfully. In practice, the training of observers that usually precedes the assessment of their interrater agreement is an exercise in extending the consensual domain in language (verbal descriptions of observations) of the observers that is needed to create or reinforce their illusion of observer-independence, which is consecutively confirmed by significant kappa coefficients.

Comment

After having been introduced in chapter 3, the idea that consensus should be defined as a form of similarity has finally been worked out in the current chapter. The utilization of similarity for the assessment of consensus is but one application of the concept, however. Turning to clinical practice for the last time, we will illustrate how similarity operates in the use of metaphors in psychotherapy.

Generally, the degree of similarity has been defined as a function of common and distinctive features. In research on similarity, the domain of features deemed relevant will be delineated beforehand, unless the research has been designated in order to retrieve the so-called 'feature space' (Tversky, 1977) employed by the subjects making the similarity judgements. But in the therapeutic use of metaphors, we explicitly take advantage of the freedom to choose the features the therapist and patient may attend to. This is made possible by the fact that the number of features that can be distinguished in an object, topic or experience is in principle unlimited. Regarding the act of distinguishing or 'making differences', Bateson observes: *'... there is an infinite number of differences around and within the piece of chalk. There are differences between the chalk and the rest of the universe ... And within the piece of chalk, there is for every molecule an infinite number of differences between its location and the locations in which it might have been. Of this infinitude, we select a very limited number, which becomes information. In fact, what we mean by information - the elementary unit of information - is a difference which makes a difference ...'* (1972; p. 453, emphasis in original).

Turning to the complex issue of metaphor, we limit the discussion of this subject to a few points. In connection with similarity, we may first observe that we speak of a metaphor when we succeed in linking two topics -called the subject and the referent - by means of language in such a way that speaker and listener feel that the two have something in common, i.e. are similar to some extent. The interactional view on metaphor (Black, 1962) holds that a metaphor creates or suggests a similarity that did not exist before. The linking of subject and referent lures us into paying more attention to the relatively few common features than to the abundance of distinctive features. To the extent that we perceive the real and metaphoric topic as meaningfully associated, we have somehow selected the 'feature space' in such a way that we feel right in our intuitive judgement of similarity. It can therefore be proposed that the art of selecting a proper metaphor in psychotherapy largely consists of being able to construe a 'feature space', induced by the metaphor, that elicits therapeutically useful and beneficial responses in the patient. An example may make this clear.

When treating a fearful patient we might first use a metaphor like 'fear is a tiger'. The patient may accept this, while attending to the dangerous, unpredictable, uncontrollable, attacking and swallowing features of both tiger and fear. These features, however, will not help him to master his fears. Alternatively we might try with 'fear is a fire-alarm'. This may be somewhat better. The common features of fear and fire-alarm alert to the fact that something is wrong and that action is needed, although you do not yet know where and what precisely. Finally, we could propose that 'fear is a mountain', focusing on the fact that a mountain may be difficult to conquer, that the climbing requires a lot of effort and that one often needs a guide. But also, that you feel very satisfied when on top of it, enjoying the scenery of all the other peaks you might now dare to climb. In passing it may be noted that the correspondence between mountain climbing and overcoming problems may be an important reason why M.H. Erickson so often urged his patients to climb Squaw Peak, a famous mountain near his home-town Phoenix.

This example points to an important aspect of therapeutic metaphors: they are not presented in bare form, but as extended stories - commonly called anecdotes - with a central theme, motive or message. Moreover, in practice successive anecdotes are used, guided by the - nonverbal - feedback of the patient, which shows which metaphors best fit the unique experiences and skills of the patient. In discussing M.H. Erickson's use of anecdotes, Zeig (1980) enumerates the following therapeutic uses and characteristics: Anecdotes are non-threatening, engaging and foster independence. The individual himself needs to make sense out of the message. Anecdotes can bypass natural resistance to change and present directives and suggestions in such a way as to maximize the possibility that they will be accepted. Anecdotes can be used to control the relationship. The listener needs to work to make sense of the anecdote, and he is kept off balance. Finally, anecdotes model flexibility, create confusion and promote hypnotic responsiveness. They also tag the memory; they make the presented idea more memorable (p. 25-26).

To let the patient take full advantage of these characteristics, we may add that a therapist often tells his metaphors in an implicit and personalized way, he also keeps them often unfinished and rather vague or evocative. Implicit, because it is often advisable to hide the aspects of the problem behavior that are addressed and the implied suggestions. If the shoe fits, the patient may wear it. Personalized, because this enlivens the vicarious experience for the patient. The therapist tells his metaphors in the first person or he stages 'a friend of mine' or 'one of my patients'. Unfinished, as this keeps the patient in suspense and the therapist can embroider the metaphor the next occasion. By letting the patient find the most fitting moral, the therapist avoids being pedantic. Alternatively, endings with shock, relief or mystery can be employed (Gordon, 1984). Vague and evocative, because such use of language,

especially when seemingly precise and concrete, encompasses many different feelings and events.

The case-histories in the following article show some of these elements of extended metaphors. The treatment in the first case has been built on the unusual similarity between sex and music, the second on the still more improbable similarity between sugar and love.

9 Individualized metaphors in hypnotherapy*

Abstract

Therapeutic metaphors in naturalistic hypnotherapy can be derived from the patient's own positive assets in the areas of work, leisure, physical and mental abilities. This extension of the utilization approach of Erickson and Rossi (1979) is illustrated with two case descriptions. A framework of how to tailor such metaphors to the individual patient is offered and contrasted to the less specific metaphoric language that is implicit in psychotherapy models. The article concludes with some theoretical considerations toward understanding the influence of metaphors on human behavior and experience.

Introduction

In employing naturalistic techniques, the hypnotherapist does not resort to formalized techniques of induction and treatment, but he uses everyday language and behavior that befits the individual patient. Normal language abounds with metaphors that represent a way of understanding and perceiving something in terms of something else (Schwartzman, 1982). When a therapist wants to extend his patient's patterns of perceiving and understanding, the use of metaphors - as a form of indirect suggestion - better suits his aim than emphasizing that more rational mode of language which is contiguity (Jacobson & Halle, 1980). Accordingly, metaphor is the central subject of a number of books (e.g. Lankton & Lankton, 1983; Barker, 1983; Gordon, 1978) that teach the complex ways in which metaphors can be constructed to fit the patient's problems and present possible solutions. An interesting alternative way to find a suitable metaphor is presented by Van der Hart (1985), who uses *metaphoric 'Kernel' statements* from the patient about his problems as the basis for hypnotherapy.

A characteristic of naturalistic hypnotherapy is to have a patient achieve his own goals by employing the abilities he already possesses but that are underutilized. This procedure, called the utilization approach (Erickson, 1967; Erickson and Rossi, 1979), is, however, often difficult to put into practice in individual cases. In this paper we demonstrate the practical application of this

* This chapter has been submitted for publication as: M. Dormaar and J. Olthof, *Individualized metaphors in hypnotherapy*.

approach, in which the patient's assets are used instead of his problems to serve as the core of a therapeutic metaphor. This provides a different perspective, enabling the patient to make positive emotional and behavioral adjustments. Two cases of this approach are presented and discussed.

The essence of this treatment technique consists of:

- Tracing an activity that is well-known and cherished by the patient and in which he is consequently absorbed with positive affect and a feeling of competence. Past and present hobbies, sports, work and other specific motor, perceptive, relational and cognitive abilities have to be reviewed in detail.
- The therapist has to identify relationships that hold both for the activity and the problem area. This may be rather general, such as involvement and passion in both music and sex, or more unique for the particular patient. The second case illustrates how overriding or wearing out, being exploited and lacking in affection are rather singular similarities between horse-riding and the patient's suffering from diabetes.
- Assess the skills, feelings and attitudes that the patient is lacking in the problem area.
- Expand the activity that is used as a metaphor for solving the problem by stressing skills, positive feelings and attitudes that the patient may be supposed to experience in the metaphoric domain. These may be attitudes like tenacity and confidence, feelings like enjoyment and caring, or skills like dexterity and concentration.
- The therapist should make himself well acquainted with the activity: its challenges, pleasures and rewards.

It is obvious that this list points at the same time to severe limitations to the application of the technique. Patients may not be able to report any positive activity, the therapist cannot find commonalities with the problem area or he cannot associate with the activity of the patient.

The first time that this metaphoric technique was reported concerns the training in hypno-analgesia by Erickson of the florist Joe with an extended tale about the growing of tomato-plants (Haley, p. 300-306, 1972). The more general idea of using a positive quality of the patient as a metaphor for overcoming the problem was not recognized, however.

Case 1

A middle-aged (51 years old) man, Mr. H., is 'living apart together' with Mrs. I (a 56 year old woman), for 8 years. They are referred to a private practice (J.O.) by their family doctor and request hypnosis to overcome their sexual problem. Mr. H. progressively did not feel like having intercourse. His partner had always taken the initiative up to now, but that made him more and more tense and so she stopped it. Mr. H. explains that he feels very angry when

sexually aroused and also senses vague images and memories that ruin everything for him. To avoid this state of mind he goes to sleep some hours after his partner or in his own house. When drunk he dares to have intercourse, but Mrs. I. rejects this as it reminds her of bad experiences with her alcoholic former husband. Mr. H. acknowledges that he is a heavy drinker; he regularly stays at home all night to drink beer while listening to classical music, his 'great love'. Although dispirited by their present sexual relationship, they are still fond of each other and both want to continue their affair. We will not go further here into the separate treatment of Mrs. I., who wants to get over her previous marriage and other misfortunes, or into the couple-therapy sessions, but focus on the treatment of Mr. H.

He presents a history of an emotionally deprived and conflicted childhood. Of his parents he mainly remembers their violent quarrels, ending in a divorce when he was 13. He lived with his father until he was 19 and then moved in with his mother, who he describes as a cold woman. His life was little more than just labouring as a construction worker until he was declared disabled 10 years ago because of chronic backpain. He has had no previous relationship with a woman, and he still lives in the house of his mother, who died 13 years ago. Further inquiry reveals that he harbours a negative outlook on the future and suffers from a very low self-esteem. He shows no other signs of a mood disorder.

Diagnostically, the presenting problem is categorized as an atypical sexual dysfunction characterized by unintelligible and unacceptable aggressive feelings and reminiscences that interfere with an otherwise normally developed sexual desire and sexual response cycle. His drinking pattern, although problematic, does not yet qualify as abuse or dependency and apparently functions to avoid situations of sexual arousal. The treatment-strategy is mainly built around his passion for music and proceeds as follows: each session the induction of hypnosis simply involves a request to close his eyes and to fantasize that he puts one of his favourite discs on the record-player and starts listening. Judged by his immobility and successful hand-levitation he goes into trance immediately. During most of the trance-work in this and seven successive sessions he listens to popular classics like piano concertos by Chopin and symphonies by Schubert and Beethoven. He relates the trance only to the music he has heard 'for hours' and shows spontaneous amnesia for all suggestions. The therapist starts with the conductor-symphony metaphor for conscious-unconscious (Beahrs, 1982), and he indirectly suggests positive self-directed feelings and attitudes connected with sexuality in the following way: While you're listening you know that the conductor takes care that all instruments have their proper turn, old pieces of long-forgotten music can be found under the dust and be allowed to sound again. All music is inspired by the unconscious. You are the conductor and you are firmly in charge, the orchestra and the conductor trust

each other and cooperate to make beautiful music. All moods have their own music. There is wild, tender, romantic and even erotic music.

When regressed to the age of six he is told: You do not have to keep quiet anymore or to hide the music that you like in order to live in peace. The grown-up conductor will accompany you to retrieve the music and together you can argue with those who have forbidden you to listen and enjoy. The conductor helps you to be angry and permits you to play the music as you like. During trance Mr. H. relates a recurrent dream of being a boy who has to labour in anguish beyond his strength and with a lot of anger and shame he also tells a secret: when he was twelve years old he was standing in the vicinity of a swimming pool, looking around. A policeman came and sent him away because he was 'peeping'. He was not allowed to stand there. The therapist considers for himself that this incident and perhaps others had created the patient's mixed feelings of sexual curiosity, aggression and shame. 'Not being allowed to stand' is a kind of 'Kernel' statement of the patient (Van der Hart, 1985), and suggestions of being allowed to stand upright are woven into the suggestions about music. While listening to Schubert's Unfinished Symphony he is told: The conductor has a firm grip on his baton and he proudly stands in front of the orchestra to let the music come out of the confinements of the unconscious. You can complete what was left unfinished and play life's music with your own hands.

In successive dreams outside the therapy sessions Mr. H. has seen himself working as a craftsman now, and his boss told him that he, the boss, had a difficult time. He enjoyed that remark and says that he rebels over his lost chances: When I was 14 years old I stopped living and was suddenly a grown up. In a deep trance he again returns to the age of a small boy, and the adult Mr. H. shows the boy how he can free himself of prohibitions. To the rhythm of music you can fantasize, explore and enjoy growing up slowly. You can make new arrangements and even compose new music; you can enjoy the music of your body, hear sounds that have long been silenced. The conductor is proud of his orchestra. Mr. H. hears the 'Ode an die Freude' of Beethoven's 9th Symphony and feels strong and confident when he returns from trance.

In the remaining part of the therapy his aggression against his mother, the policeman and others who have banned his enjoyment of sex and life in general is ventilated with silent abreaction (Edelstien, 1984) and by blowing all his anger in a balloon that floats away. He is also coached to develop his childlike sexual fantasies into mature and realistic desires.

The therapy comprised 23 sessions, 16 of which in individual therapy with Mr. H. In all his sessions, hypnosis was the only treatment technique. In addition, there were 4 sessions with Mrs. I and 3 joint sessions. During the last follow-up session the couple reports that they now have a tender and quite satisfying sexual relationship. They are also planning to live together in

Mrs. I's house. Mr. H. has completely abstained from alcohol the last four months. He also bought a piano and practices at length with great delight. He finally notices that he is a lot more confident and enjoys a positive outlook toward the future.

Case II

A 35 year old male suffers from diabetes. He claims to be certain that his disease has a psychogenic origin. Diabetes was diagnosed at the age of 21 after years of different psychosomatic complaints and depressive states. He has been under constant medical care for his ailment since then. The therapist accepts the patient's statement on the origin of his diabetes as a personal construction of meanings, which can be utilized. He is aware, however, that there is no accepted medical evidence for such a belief. From childhood on the patient had a great affection for horses. As a child he worked for a farmer in his spare time and was only occasionally rewarded by riding the farmer's horse. As an adult he pursued a career as manager of a riding school. He owned some horses and has won his spurs in breaking horses.

The horse-world serves as a context which provides metaphors for the understanding of the development of his diabetes and for a treatment perspective. The rider-horse relationship denotes the mind-body relationship and the horse also symbolizes the patient as a hard-working boy who had little time left to ride and often did not get his 'sugar-candy'. Finally, the horse is like the boy who was in want of his parents' love, and love is the 'sweet thing of life'. The patient describes himself as a lame horse, not well in his hairs and burned-out. He is, however, still positive about the horse's future possibilities.

After a classical induction procedure of eye-fixation and suggestions for relaxation, therapeutic metaphors were developed in the following way. The horse is trained to achieve suppleness and variation without too high hurdles. The therapist acts as a trainer, who treats the horse lovingly, respectfully and with much encouragement and rewards. Sadness with respect to his unloving parents, anger at the farmer and other 'riders with the whip' are experienced in metaphors of a horse that was run-out and had to be self-supporting, but could not retain anymore the sugar within himself.

After 10 biweekly sessions his need for insulin had dropped from 74 to 28 units per day, and, even more important, his glucose blood-level remains stable. Formal hypnosis has been used in 4 sessions, but many kinds of horse-world metaphors were the central theme of all therapeutic work. He is also more receptive to signs of affection from his wife and son. Love now means to him a natural and lasting way to store his energy. He is more able to code somatic sensations into signs of need for food or affection instead of signs of

disease. Furthermore, he has developed a quiet rhythm in his life. At the time of writing this article, therapy is still continuing with sessions every 4 to 6 weeks.

Discussion

In order to evaluate this approach the following issues need to be clarified:

- 1 Is the improvement reported by the patients due to the metaphoric technique?
 - 2 How can we understand the effectiveness of the technique in these specific cases? and
 - 3 How can the alleged effectiveness of therapeutic metaphors be explained on a more general, theoretical level?
-
- 1 The claim that the reported improvements are due to the employment of metaphors in hypnosis is hard to substantiate. The case-reports can by nature only supply clinical evidence. In both cases the induction of a trance-state served to foster the receptivity of the patient and to by-pass the scrutiny of the conscious mind to the suggested implications. Techniques as described here remain largely a question of clinical acumen. Witzlum et al. (1988) have provided excellent guidelines on strategies and tactical approaches. Nevertheless, lack of research and formal theory hamper our understanding of relative effectiveness, operating procedures and relationships to trance-phenomena.
 - 2 The elaboration of the music metaphor in a trance-state is central to the first case. Corresponding elements of the metaphoric domain and sexuality as the reality domain are enjoyment, being active and passive, being competent and 'standing up for it'. The metaphor has been extended in a metonymic way. Both metonymy and metaphor are the basic ways in language to construe a coherence of experience through linkages (Schwartzmann, 1982). In metonymic language the linking of experiences and perceptions is based on their contiguity in time and/or place. In metaphor connections are made between two disparate fields of experience on the basis of similar elements or relationships in both. This interactional view is exemplified by Black (1962) with the common metaphor of 'a man is a wolf'. The listener knows the 'associated common-places' about the wolf, and he will be led to construe a corresponding system of implications about the second subject. Thus the wolf metaphor organizes our view on man. Also, from now on we will tend to perceive a wolf as more human-like. Accordingly, we believe that in this case and many others, metaphors owe their effectiveness largely to letting the patient make corresponding implications in the problem area to those suggested in the metaphoric domain.

The second case proceeds essentially along the same lines. For a true understanding it is noteworthy that a stable glucose blood-level and lower

insuline dependency were no conscious purposes (Bateson, 1972) of either therapist or patient. Therapy aimed, rather, at the development of the patient's true nature, especially his giving and receiving of affection.

- 3 In addition to the earlier remarks, our present theoretical understanding of therapeutic metaphors is as follows: By examining the relationships within and between concepts - like id and ego, self-structure and experience, individual and family, stimulus and response, or conscious and unconscious - therapeutic models describe in their own metaphors the limited repertoire that characterizes neurotic problems and disorders. A limited repertoire implies deficient forms of connections or links between whatever concepts the observer is able to perceive. In the same vein, all forms of psychotherapy share a common aim of promoting new links between experience, behavior and perception.

Given that metaphor is the ubiquitous way of talking about experience (Gordon, 1978), the *explicit* use of metaphor that organizes and determines our attitudes and perceptions (Lakoff & Johnson, 1980) gives more freedom to expand the patient's potentials to feel, behave and relate than interventions that are derived from the fixed metaphors of a theory of psychopathology. Furthermore, an activity or experience in which the person himself is strongly involved elicits more affective and cognitive associations - i.e. links or connections - than a metaphor that is either alien to the private world of the patient, worn thin by colloquial use, or originating primarily from the therapist. Finally, the evocative language of metaphors is a potent vehicle for putting images into words. The power of imagination or visualizing is well-known in hypnotherapy (e.g. Fowler, 1988) and psychotherapy in general (Singer & Pope, 1978).

These theoretical considerations put ingenuity to the test because of the inherent emphasis on attending to the idiosyncracies of the patient. It might be hypothesized, however, that the use of 'fitting' metaphors, i.e. derived from the patient's own behavior or experience, will produce a better outcome than standard metaphors based solely on the therapist's theoretical stance and the patient's problem descriptions.

A complete presentation of the theoretical model goes beyond the scope of this article, but we also predict that the use of 'organic' metaphors, derived from the living world, will be more effective than 'trivial machine'-like metaphors in promoting imagery. Finally, we believe that effective metaphors are characterized by an extension of the consensual domain (Maturana, 1988) between patient and therapist, based on research evidence (Dormaar et al., 1989) that consensus can partly predict positive outcome.

Authors' note

The authors are grateful to Onno van der Hart, Marten W. de Vries and Thurman Mott Jr. for their comments.

10 Conclusion

10.1 Overview

The preceding chapters discussed consensus between patient and therapist as a means of describing and measuring the degree of collaboration during treatment and the relevance of consensus for a successful outcome. The main conclusions of the clinical research and case descriptions have been presented already in the respective chapters and will now be briefly summarized. The introduction discussed the range and diversity of sociological and psychological studies on the subject, showed how the different components of consensus are interrelated and demonstrated that consensus is a condition for social coordination. We further demonstrated how consensus can be assessed both intrapersonally and interpersonally and that certain aspects of consensus - notably consensus experienced by the patient, understanding by the therapist and agreement - can partly predict outcome. The predictive consensus measures 'understanding by therapist' and 'consensus experienced by the patient' characterize a complementary patient-therapist relationship that favours a successful outcome according to these results. Outcome was defined as encompassing not only the degree of change but the level of functioning attained after treatment as well. The results indicated that a large variance within the patient population in outcome scores makes it difficult to ascertain improvement. Results further revealed that, even before treatment, a sizeable number of patients have scores that belong to the normal range and that many non-patients score in the dysfunctional range. Furthermore, I discussed how consensus is best understood as similarity and presented a method for assessing levels of significance for the similarity coefficient, as used to measure interpersonal consensus, especially agreement.

Whereas consensus regarding problems brought to treatment was dealt with in the research section, the first two case descriptions showed that consensus on treatment goals and procedures is vital to a successful outcome. In comments on these clinical cases I have argued that consensus displays the 'fit' of patient and therapist, which mainly consists of their intersubjectively constructed, symbolic reality, necessary for therapeutic change. The last case report illustrates how metaphors can be an effective means of inducing change. In devising metaphors we make use of the human proficiency at linking two disparate fields of experience by means of similarity.

I have not gone into the issue of how we can benefit from the measurement of consensus in the practical execution of psychiatric treatment/-

psychotherapy. I do believe that an open discussion by patient and therapist about their responses on the consensus questionnaires can clarify misunderstandings and help to strengthen the collaborative basis for treatment. This belief is presently founded only on a few personal experiences, which were sufficiently positive, however, to warrant a more sophisticated clinical trial.

In the meantime, comprehension of the relevance of consensus remains very limited. I have noted that the single assessment of consensus, as in the current research presented here, precludes our understanding of how consensus develops during the treatment process and which aspects are important either in conjunction or separately, over successive phases. Another major shortcoming of the research on consensus reported here is that the *affective* side of the therapeutic relationship has not been investigated. I have conducted my research before becoming acquainted with the balance model as presented in the introduction. I believe that it is unwarranted to assume that the affective relationship between patient and therapist is always mutually positive. By taking into account the implications of the balance model, the associations between consensus and outcome can therefore be ascertained in a more sophisticated way. Finally, a more experimental approach to investigate the importance of consensus to attaining positive results in psychotherapy would need a research design with a manipulation of the feed-back of consensus-data.

10.2 Toward the use of interpersonal variables in psychiatry

Although the work presented in this thesis centers around consensus, its essential message is a broader one. I deem the investigation of consensus in psychotherapeutic practice to be highly interesting in itself. But on the level of theory and practical research methodology, the approach employed in the research and the discussion of the case studies is but one application of a more extensive concern, namely a clear appreciation of the *relational* character of psychiatric data. In the introduction I touched on the importance of this topic for research in the social sciences, and I will conclude now with outlining some examples of an approach to psychiatric diagnosis and treatment that clearly utilizes interpersonal variables.

When looking at the diagnostic process, we observe that diagnosing implies a classification of phenomena that is both general enough, by containing a limited number of categories, and precise enough to give the uniqueness of each case its due. In the *basic* criteria commonly used for classification man reveals some of his pervasive preoccupations, namely 'function' and 'history'. A functional classification is one in which we assign objects to categories on the basis of how they operate, what we can do with them, what use they have to us or what tasks they can fulfill for us. Examples are the classification of all

objects devised with a purpose, like tables and cars, but the functional principle is also evident in categorizations of some animals as domestic, of certain plants as crops and in number categories.

A 'historical' classification is commonly used to categorize biological phenomena like in biological taxonomy, and in medicine, notably human pathology. Such a classification is called historical here to emphasize the common denominator in concerns of reproduction, phylogeny, ontogeny, causality and etiology. The above distinction in functional and historical classifications can be clarified by indicating that functional ones espouse the principle of *analogy*, whereas historical ones focus on ways of reproduction and *homology*. Thus, in biology a whale belongs to the same class as man and elephants but not to that of fishes. Also a whale's tail and the legs of other mammals are considered to be alike. On the other hand, butterflies and birds - although functioning similarly in terms of flying - do not even belong to the same phylum. We also note that from ancient times and transculturally, men have likewise tended to classify themselves primarily in terms of history and reproduction through descriptions of lineage and family relationships. Functional criteria like profession have always been quite secondary to human categorization.

In the early years of psychiatry, pioneers like Kraepelin and Freud attempted to categorize abnormal human behavior with guidelines like history and etiology. Among other factors *equifinality* (Von Bertalanffy, 1979), meaning that different initial conditions may lead to a same result in living systems, doomed this endeavor. For psychiatry this general principle implies that adults with distinct traumatic experiences during infancy may show a similar disorder, whereas similar infantile experiences may lead to different outcomes in adult life. More recently, categories of disorder have included elements of a functional dimension, notably in behavior therapy.

Present psychiatric classification like the Diagnostic and Statistical Manual-Revised (DSM-III-R, APA 1987), has no clear-cut principles for its categories, apart from a probabilistic or statistical one: symptoms that frequently go together are clustered into disorders. In some categories age of onset or duration of symptoms is included as a decisive criterion. Apart from other objections to the DSM-system (e.g. Frances, 1980; Klerman et al., 1984; Williams, 1985), the lack of a general criterion is a serious deficiency, since we need such a criterion to found our making of distinctions (Spencer-Brown, 1969). To illustrate this, the word 'axis' implies a *dimension*. We might wonder, notably for axis I and II, why is that dimension so hidden? Another major inadequacy of the DSM-system - though not of the classification by itself - is that its diagnoses are not systematically tied to specific treatment interventions.

Considering the weighty consequences for both patient and mental health care provider, the making of a psychiatric diagnosis is a severe responsibility. Together with the objections to the DSM-system listed above, it is therefore justified to challenge the robustness of the categories of our diagnostic standard. With the use of a similarity coefficient the reliability of the categories can be tested in the following way.

First I propose, that if a diagnostic system is a reliable one, then two individuals suffering from the same disorder should be more alike than two persons suffering from different disorders, if disorder-symptoms define the feature space. Being alike is operationalized as sharing certain symptoms together with the common absence of certain other symptoms.

Now the DSM-system can be put to an unorthodox test of reliability:

- 1 Take a random sample of psychiatric patients.
- 2 Have them diagnosed, either by means of the usual clinical interview or some standard method.
- 3 Independently from step 2, check all patients with a standard clinical instrument for absence and presence of common psychiatric symptoms. An instrument like the CPRS (Asberg and Schalling, 1979) that is based on both self-reports and clinical judgements, might do, if properly modified.
- 4 Assign independent weights to all symptoms. Dependent weighting - i.e. judging a symptom as important if certain others are present or absent - is to be avoided, since such a procedure would bias the instrument with a structure in the data that still has to be discovered.
- 5 Compute similarity coefficients like the Matching coefficient (Sokal & Michener, 1958) or Jaccard index (Jaccard, 1908) for all possible dyads with the weighted symptoms.
- 6 Select the dyads that consist of persons both belonging to a certain diagnostic category - say schizophrenic disorders - as assessed in step 2.
- 7 Compare the distribution of similarity coefficients of dyads consisting of two persons with that same diagnostic category with the distribution of similarity coefficients of dyads consisting of patients with different diagnoses.

If the DSM-system can group patients reliably, then the mean and standard deviation of the first distribution should be significantly different from those of the second. Alternatively, the number of dyads, in the same category that exceeds the level of similarity observed with 95 percent of all possible dyads as calculated in the permutation distribution, is an indication of the reliability of the diagnosis. These procedures are similar to the ones described in Snijders et al (1989) and in chapter 8.

In addition, starting from the proposition that the distinction between two categories is a valid one if the distinctive features of the objects assigned to the two different categories outweigh the common features, strategies may

be developed to compare the distributions of weighted common and distinctive features of the two diagnostic groups.

A major advantage of such procedures is that they provide independent checks on reliability and validity. Usual research in this area investigates the coherence or consistency of symptoms and relationships of diagnosis to patient variables obtained from other sources. Such validation can never basically challenge our belief in - or use of - the DSM as a 'gold standard'. The procedures outlined above can elucidate whether the system meets the logical requirement of similarity between persons, in addition to test criteria like internal consistency between symptom-variables and inter-rater reliability.

If then we want a psychiatric classification that meets the shortcomings of categorization and specificity regarding treatment, how should we proceed? When taking a closer look at taxonomic classifications like those for the animal and plant kingdom, it becomes clear that taxonomy reveals a hierarchy of complexity of organization from microbes to man.

This is obvious when contemplating the ecological niches of different organisms. Certainly, some microbes have adapted to extraordinary circumstances, such as a highly saline or sulphurous environment. But this niche also defines their range of interactions, i.e. they can live nowhere else. Only man lives both in deserts, at sea and near the poles. Man is so adaptable because he excels in extending his possible range of interactions with the help of 'tools', in its broadest sense, that affect the impact of his environment. He survives because he knows 'how to fit'. Von Glaserfeld (1984) is quite right in pointing out that the meaning of 'survival of the fittest' is often misunderstood to signify that the most rigorous or robust individuals of a species survive. Thanks to the higher rates of reproduction that usually go along with lower levels of complexity, lower species can survive ecological changes thanks to genetic transformations when the adaptation of the individual members fails. Thus, higher complexity entails more possibilities of adaptation (Lilienfeld, 1978), and adaptation in turn can be operationalized as the range of interactions an organism can enter without losing its organizational identity (Maturana and Varela, 1980).

During his development, man becomes involved with an increasing complexity of the interactions he is supposed to maintain in his different roles. In referring to social breakdown and less severe limitations in social involvement, we note that psychiatric patients share the common characteristic of being limited in their range of, notably *social*, interactions. This limitation in range of interactions is therefore the most conspicuous and general feature of patients suffering from a mental disorder. I therefore propose to use the level of adaptation - in the sense of 'range of interactions' - as the basic criterion for creating a continuum of psychiatric problems or disorders, an idea recommended earlier in a somewhat different form by Menninger et al. (1958). To

some extent this proposal is in line with the interactive-developmental model for schizophrenia of Strauss and Carpenter (1981).

We need, then, to measure the range of interactions to arrive at the continuum of psychiatric disorders according to the adaptation criterion stated above. The following plan will do so by taking into account the relational character of person-to-person interactions. Defining the range of interactions in a *dyadic* framework does not mean ignoring the importance of interactions at the group level. In discussing agreement within a group (chapter 3) it was already stated that interactions involving more persons basically consist of multiple dyadic interactions.

The concept of similarity presented in this thesis and especially in chapter 8 can be employed to measure the human relationships that qualify patient-environment interactions, i.e. the interactions of the two individuals that constitute the basic therapeutic system. The plan to be outlined here can ultimately result in a diagnostic scheme built on the adaptation criterion, which offers concrete guidelines for treatment.

Granting that adaptation can be equated with the range of interactions an organism can enter, we should realize that when dealing with adaptation we refer to the total individual in all his possible fields of behavior, affect and experience. So, we first have to meet the possible objection that, as psychiatrists, we cannot pretend to be experts on the entire human being, but only on his disorders. Both on practical and theoretical grounds this objection can be discarded. First, both professionals and educated laymen obviously have their good reasons to psychologize or 'psychiatrize' nearly all human actions, especially emotional (i.e. commonly interactional) behavior. Also, in principle no field of interaction of his patients is devoid of interest for the psychiatrist. More important, however, is the fact that man - like other organisms - can best be considered to behave *as a whole*. This idea of integrated functioning has been elaborated more fully for the central nervous system (Von Foerster, in: Segal, 1986; McCulloch, 1965). We therefore maintain that psychiatric classification is best conceived as a categorization of patients with respect to their degrees, levels and manners of adaptation. Incidentally, it can further be remarked that concepts of mental health generally imply that a healthy individual is well adapted to the different aspects of social life, in particular.

Human interactions take place in the form of diverse relationships. Thus, a measure of 'range of social interactions' will be operationalized as the range of relationships a person can master or cope with vis a vis another person. In chapter 2 it has been stated, that a person's emotions can be viewed as his means of communicating his definition of the relationship. For the measurement of relationships we therefore need first a system to survey the range of all possible definitions of relationship. For this task, Leary's (1957) work on the interpersonal diagnosis of personality is the best starting point. This author is

the originator of a model of human interactions known as 'the rose of Leary'. This model is based on two orthogonal dimensions of human relationship, called 'dominance' and 'affiliation' (hostility-solidarity) (see figure 10.1).

Figure 10.1: The rose of Leary, depicting definitions of relationship on the basis of two dimensions

The first abbreviated notations in the figure refer to the leading dimension, the second to the subordinate dimension. The model allows us to discriminate degrees of intensity, like non-adaptively minimal, adaptively moderate and non-adaptively intense. The two inner circles of figure 1 indicate limits of such degrees.

The circular arrangement of relationship definitions has been independently verified by several investigators, and a number of factor analytic studies of human interactions, reviewed by Beekers (1982), have corroborated the two dimensions. A score of studies confirm the relationships between an effective range of interpersonal behavior and mental health and report significant associations between deficiencies in relationship qualities and psychiatric disorders (Beekers, 1982). Factor analytic studies of the interpersonal checklist (ICL) of Laforge and Suczek (1955), based on this model, consistently show the same two dimensions, sometimes with an additional third factor interpreted as an activity dimension (Beekers, 1982). Vertommen and Rochette (1979) report strong correlations of different MMPI-scales with scale scores on the VIR (Vragenlijst Interpersoonlijke Relaties), a Dutch version of this kind of relationship inventory for assessing relationship definitions regarding others in general and specific others.

In order to measure the level of adaptation of the individuals in the interacting dyad, the two basic modes in human relationships first have to be defined, i.e. symmetry and complementarity. If certain behaviors of the first person are linked to behaviors of the second person that are regarded as

similar by both individuals, the relationship is symmetrical (Bateson, 1972; p. 323). Thus, in the model above, if a person defines himself as dominant in relation to the other person and that other person does so too, they are likely to engage in a symmetrical relationship of the dominant type. Likely, because the link between self-reports on relationships and actual interacting behavior is only a correlational tendency (Mischel, 1968). A complementary relationship signifies that certain behaviors of the first person are linked to different, but fitting behaviors of the second person. So, if the first person defines himself as dominant in relation to the other and the second person defines himself as submissive, they are likely to engage in a complementary relationship of the type 'dominance-submissiveness'. Similar complementary relationships are those typified as dependent-supportive, exhibitionism-spectatorship. It should be noted that the different behaviors in a complementary relationship are not 'dissimilar', like mentioned by Bateson (1972; p. 323), but also fitting because they are more properly defined as both similar *and* dissimilar. Complementary behaviors are, for instance: requesting and giving support, guiding and wanting to be guided.

Second, the two dimensions are not equivalent. The dominance-submission dimension is a complementary dimension, since a relationship definition of a dominant type fits in with a relationship definition of the submissive type. The affiliation dimension is an antagonistic dimension of incompatibilities. Solidarity does not stimulate hostility, nor the other way round. If anything, solidarity elicits reciprocating friendly behaviors, and the same applies to hostility. Interaction based on mutual relationship-definitions around one of the poles of this dimension can therefore be supposed to represent a combination of weak, but stable symmetrical and complementary patterns. On the other hand, symmetrical patterns around one of the poles of the dominance-submission dimension, like mutual dominance, mutual submissiveness, or dependency, will tend to be rather forceful, but unstable. Two persons cannot be pitiful and requiring help together and both at the same time for a longer period. One of the two has to shift its position in order to maintain a relationship. The foregoing leads to a different configuration of the model which represents the diverse types of symmetrical and complementary relationships (see figure 2).

Figure 10.2: Leary model depicting quadrants of symmetrical and complementary interactions.

I-I

unstable symmetrical relationship patterns

III-III

II-II

stable symmetrical patterns with moderate complementarity

IV-IV

I-III stable or rigid complementary patterns

II-IV incompatibility (no stable patterns)

The importance of the relationship definition in all human communication has been stated in chapter 2. The definition of the relationship has three components:

- a how the person sees himself;
- b how he sees the other(s) and
- c how he sees his relationship to the other person.

To illustrate this, imagine a woman who communicates sobbingly to her husband: 'You should take me seriously'. She then essentially conveys:

- a I see myself as pitiful;
- b I see you as responsible for my sorrow;
- c Try to make amends.

The example also serves to make clear that the relationship aspect of human communication - foremost consisting of different kinds and degrees of so-called 'emotional' behavior - implicitly *commands* the communication partner to behave complementarily, that is, to show 'fitting' behaviors. If in this case the husband complies, a submissive-supportive relationship exists. The communication as regards content of the messages exchanged will then be *un-*

problematic, i.e. without 'noise' from the communication about the relationship or metacommunication. If the husband does not comply, the communication as regards content is *problematic*. Metacommunication may take precedence.

I will now turn to the the range of relationship definitions as measured with the Dutch relationship inventory VIR (Vertommen and Rochette, 1979). With this instrument scores are calculated per octant as depicted in figure 1. The VIR (form: direct perspective, 'I versus unspecified others') contains 10 items per octant with a 5-point scale, thus comprising 80 items. They are formulated like: 'I need the support of others' and 'I defend my rights against others'. The intercorrelations reported on the octant-scores lend strong support to the quadrant-division proposed above in figure 2 (see Vertommen and Rochette, 1979; p. 44). The VIR also has an indirect or metaperspective which surveys the views a person holds about the relationship definitions of others toward himself with essentially the same items as used to survey the person's own definitions. This metaperspective is comparable to the level of understanding in measuring consensus.

To measure the range of interactions of a dyad in terms of their symmetrical and complementary relationships as derived from their respective definitions of relationship, the direct perspectives of the two persons and the direct perspectives of the first person and the metaperspectives of the second person (and vice versa) have to be combined. For such combination in the form of a similarity coefficient, item scores have to be dichotomized by creating dummy variables or using a cut-off (high-low). The subsequent steps in the measurement procedure consist of making contingency tables and the computation of similarity coefficients. This procedure is illustrated in table 1 and table 2 for symmetrical and complementary relationships concerning the first quadrant.

The Jaccard-coefficient, or another similarity coefficient depending on the research questions, is used for operationalizing the degree of symmetrical relationship of the dominant type (I-I in fig. 2) between persons A and B. Likewise, a complementary relationship can be measured with a similarity coefficient based on a contingency table of direct and metaperspectives (see table 2). Strictly speaking, the linking of direct perspectives concerning quadrant I and III is the correct way to measure complementarity. Since the items of quadrant I and III are differently formulated, the *direct* perspective concerning quadrant I of the first person and the direct perspective concerning quadrant III of the second person (and vice versa) cannot be compared directly.

Table 10.1: Contingency table for measuring degree of symmetrical relationship in quadrant I of the Leary model of relationship definitions.

		direct perspective of person A, items of quadrant I		
		agree	disagree	total
direct perspective of person B, items of quadrant I	agree	a	b	
	disagree	c	d	
total				20
Jaccard coefficient:		$\frac{a}{a + b + c}$		

Table 10.2: Contingency table for measuring degree of complementary relationship in quadrant I of the Leary model of relationship definitions.

		direct perspective of person A, items of quadrant I		
		agree	disagree	total
indirect perspective of person B, items of quadrant I	agree	a	b	
	disagree	c	d	
total				20
Jaccard coefficient:		$\frac{a}{a + b + c}$		

The four similarity coefficients for symmetrical relationships and the four for complementary relationships, taken together, constitute the relationship pattern which, at least partly, determines the range of interactions the two persons of the dyad can enter into. Although we may assume that all types of relationship

have their adaptive value, depending on the situation and the task at hand, it is obvious that certain relationships are more suited for therapeutic interaction - e.g. a mixed symmetrical-complementary relationship in quadrant II - than others - such as a mixed symmetrical-complementary relationship in quadrant IV. Nevertheless, purely symmetric relationship qualities are needed, both in therapy and daily life, because too much complementarity tends to escalate toward extremes - e.g. growing support and dependency - and to impede those kinds of communicative interactions, which can only occur in a symmetrical relationship.

The procedure proposed above gives an idea of the *structure* of the relationships in a dyad at a certain moment. The *process* of relationship development can be captured in a different way. The structure at the onset can be assessed by comparing direct and indirect perspectives of a patient and therapist prior to their first meeting. After the first contacts the assessment should be repeated. This time the two are requested to answer in terms of how they view their relationships vis-a-vis their interacting partner instead of unspecified 'others', like in the first time. Changes in structure from first to second assessment show the degree to which the dyad can *learn*, i.e. change their relationship definitions through the feed-back of previous experience of interaction. The difference between structure and process is between 'being adapted' and 'learning to adapt', the latter being of a higher order than the first.

A possible objection to the method proposed so far is that it offers only a categorization of levels and patterns of adaptation of patient-therapist dyads, but not of adapted and non-adapted individuals, e.g. our patients. In a sense this is true, since the unit of survival - and hence of adaptation - is not the individual, but 'the flexible organism in its (changing) environment' (Bateson, 1972; p. 451). We are interested in social adaptation, and thus the environment is necessarily of a social nature. We can only assess an individual's level of adaptation in relation to someone else, and this requires a level of description beyond the level at which we describe an individual. The matter of levels of description and the associated idea of operational closure are central to the model of perception and theory of knowledge (Fleischaker, 1984) advocated by Varela (1977) and Maturana (1980, 1988), whose ideas on living systems have guided these conceptualizations. To clarify this point, it is obvious that we need to know how the interacting partner defines the relationship, in order to know if a person's dominant definition of relationship belongs to a symmetrical or complementary relationship.

If one still prefers to limit the description to the level of the individual's most probable range of interactions in terms of the relationship definitions he is generally inclined to, this individual's personal octant or quadrant scores on the VIR can be related to the distributions of the octant scores of reference

groups, as reported in Vertommen and Rochette (1979). With suitable ways to report such relationships of individual scores to distributions, this also constitutes information regarding the patient that clearly shows the relational nature of scientific facts.

Finalizing this first outline of a categorization based on adaptation, some vital questions still remain to be answered, e.g.:

- 1 How to interpret such relationship patterns to achieve a categorization of levels and degrees of adaptation?
- 2 What might be the practical utility of such patterns?

Both questions require a lot of theoretical and empirical work, which can build on what has been accomplished in interpersonal research over the past thirty years. In the context of this thesis I will limit myself to some remarks.

First, the importance of 'range' implies that a dyad which shows a rather uniform distribution of relationships is better adapted than a dyad with only a few well-developed relationships. Second, the equilibrium between complementarity and symmetry in relationships - a standard for excellent communication (Watzlawick et al., 1967) - is a related criterion for good adaptation. Third, it may also be assumed that *moderate* mutual relationship definitions, and hence moderate levels of intensity of relationship, will be more adaptive than extremely intense or very low degrees of relationship. Fourth, the matter of adaptation in terms of process may well be of more importance than adaptation in terms of structure. When speaking of chronic patients, untreatable or difficult patients, schizophrenics and so on, we often convey that there are a number of patients with whom we - and our fellow-men - can interact only in limited ways, that we cannot change or expand these ways and that we suppose (with our patients, mostly) deficits in learning how to adapt. Empirical data on patient-therapist dyads as proposed here, together with data from other sources, can be used to test a model of a hierarchic categorization that is referred to here in a still unspecified form.

On the questions of practical utility, I would like to offer that a diagnostic scheme of a hierarchy of adaptation offers guidelines for treatment. If confronted with complaints of fear, anxiety, depression, unmanageable conflict, mourning or even with thought disorders and hallucinations, the main problem in therapy is not that we do not have at our disposal suitable techniques to help the patient in mastering his problems. The main problem in daily practice is how to relate to our patient in such ways that the application of such techniques avail him. The management of these problems are taught and learned in very diverse and often poorly articulated ways, e.g. in case-reports and professional training. The manner of diagnosing sketched here may guide the art of healing in psychotherapy with clear and transferable knowledge. Essential knowledge for the psychotherapist is to know in what respect he should try to

change himself in order that the other person - i.e. his patient - in the therapeutic system has the opportunity to change as well. It is this kind of situation- and person-dependent self-knowledge that the diagnosis should procure so that the therapist can adequately use his own person as an instrument to promote change. If he is unable to modify his relating sufficiently, the 'interactional diagnosis' of his patient may help in selecting a colleague with the qualities needed for the particular case.

Psychotherapy as a planned and goal-directed influencing of one or more persons (the patient(s)) by another (the therapist) is a process we know both a great deal and very little about. Extensive research literature informs us of the effectiveness and limitations of therapeutic techniques and the impact of patient or therapist characteristics. At the same time this knowledge does not enable us to predict adequately our chance of success in a particular treatment case. Although psychotherapy accomplishes more than might be expected by chance, the degree of success of even the most advanced and thoroughly investigated techniques is quite modest (Smith, Glass and Miller, 1980).

In trying to account for this inconsistency in results, we should not settle for the truism of the unpredictability and diversity of human behavior, but we should scrutinize the model of psychotherapy that is implicitly employed in research and that also guides our therapeutic efforts. The loose definition of psychotherapy at the beginning of this section reflects that model, but is actually quite erroneous. That implicit model of psychotherapy, especially in most research, is one of a therapist operating upon the patient with certain techniques, in order to change the latter's experience and behavior. Perhaps the attractiveness of that model and case histories that exemplify it is that a picture is evoked of powerful magic: a person possessing the rather superhuman quality of altering another person. The research literature strengthens that ideal of unilateral control and one-sided influence. Apparently we like to look upon person-to-person influencing in the same way as we tend to view the manner in which a mechanic fixes a car.

In doing therapy we are, however, quite aware that a passive patient requesting to be 'fixed' is no advantage, but poses a serious therapeutic problem. This paradox of a model of unilateral control and the practical experience of the necessity of active cooperation points to the fact that the model of psychotherapy does not properly accommodate the *reciprocity* of the influencing process of therapy. The model of unilateral control is incorrect, because we never meet a patient passively accepts the therapeutic techniques of his therapist, behaves accordingly and also benefits from the therapy. In daily practice as psychotherapist we know that we are just as much influenced or determined by the actions of our patients in what to say and how to act as

the other way round. In the same vein, both patient and therapist can be looked upon as equally autonomous, i.e. uninfluenced by the other.

A conceptualization of psychotherapy that sufficiently includes the interpersonal and reciprocal nature of the process of influencing may benefit from studies in the field of social psychology (e.g. Newcomb et al., 1966), which present general principles of how people mutually affect each other. In the field of psychotherapy research a focus on general characteristics - in contrast to those that are deemed specific to a particular treatment technique - has been represented most notably by Frank (1968, 1969, 1971, 1974a, 1974b, 1975, 1978, 1979, 1986) in his investigations of factors like placebo-effects, fostering positive expectations, the rousing of affect and giving a rationale. Factors like these are present in all forms of psychotherapy, a type of personal encounter that Frank subsumes under the category 'persuasion and healing' (Frank, 1974c). Frank's common factors are still conceived, however, as primarily residing in one of the two partners concerned and not directly in the interactions that constitute the therapeutic relationship. Both in our research on consensus and in the ideas on psychiatric classification propagated in this concluding chapter, I have tried to further the line of research into 'common factors' by including interpersonal factors.

In a sense, I end where I have started: it is a major challenge for psychiatry to map human relationships in order to describe psychiatric disorders in interactional terms that may guide in turn the therapist's interactions with his patients, that is: the treatment itself. In so doing, we do well to remember that the map is never the territory (Korzybski, 1941), thus preventing the misplaced concreteness that threatens our open-mindedness.

Finally, I want to defend why I have purposefully skipped over details, notably in this chapter. Zeig's (1980) book on Milton H. Erikson begins with an anecdote on Niels Bohr, when lecturing on Heisenberg's principle of uncertainty, also called the principle of complementarity. It says, that when focusing on the location of a particle, the observer sacrifices information about the momentum and, when discovering its exact momentum, sacrifices information about location. Faced with the general validity claimed for this principle, a student asked Bohr: 'What is the complement of clarity?'. After a moment's thought Bohr's reply was: 'precision'. It is clear that, by simplifying complex issues and avoiding detailed descriptions of a vast body of interpersonal research, I have strived for some clarity, leaving the meticulous task of precision for the future. For my inquiries at this moment in time, this choice reflects the heuristic principle of science.

References

- Allport GW (1960): *Personality and social encounter*. Beacon Press, Boston.
- Anderberg MR (1976): *Cluster analysis for applications*. Academic Press, New York.
- Anderson H and Goolishian HA (1988): Human systems as linguistic systems: preliminary and evolving ideas about the implications for clinical theory. *Family Process*, 27: 371-394.
- Arrindell WA and Ettema H (1981): Dimensionele structuur, betrouwbaarheid en validiteit van de Nederlandse bewerking van de Symptom Checklist (SCL-90); gegevens gebaseerd op een fobische en een "normale" populatie. *Nederlands Tijdschrift voor de Psychologie*, 36: 77-108.
- Arrindell WA, Emmelkamp PMG and Sanderman R (1986): Marital quality and general life adjustments in relation to treatment outcome in agoraphobia. *Advances in Behavior Research and Therapy*, 139-185.
- Asberg M and Schalling D (1979): Construction of a new psychiatric rating instrument, the comprehensive psychopathological rating scale (CPRS). *Progress in Psychopharmacology*, 3: 405-412.
- Ashby WR (1964): *An introduction to cybernetics*. Methuen, London.
- Auerswald EH (1985): Thinking about thinking in family therapy. *Family Process*, 24: 1-12.
- Auerswald EH (1986): Family therapy as a movement: epistemological barriers to ontological freedom. *Journal of Strategic and Systemic Therapies*, 5: 14-19.
- Auerswald EH (1987): Epistemological confusion in family therapy and research. *Family Process*, 26: 317-330.
- Austin B and Colwell RR (1977): Evaluation of some coefficients for use in numerical taxonomy of micro-organisms. *International Journal of Systematic Bacteriology*, 27: 204-210.
- Averill JR (1968): Grief: its nature and significance. *Psychological Bulletin*, 70: 721-748.
- Barker P (1983): *Using metaphors in psychotherapy*. Brunner/Mazel, New York.
- Bateson G (1972): Conscious purposes versus nature. In: Bateson G, *Steps to an ecology of mind*. Ballantine, New York.

- Bateson G (1972): Double Bind, 1969. In: Bateson G, *Steps to an ecology of mind*. Ballantine, New York.
- Bateson G (1972): *Steps to an ecology of mind*. Ballantine Books, New York.
- Bateson G, Bateson MC (1987): *Angels Fear. Towards an epistemology of the sacred*. McMillan, New York.
- Battle CC, Imber SD, Hoehn-Saric R, Stone AR, Nash ER, Frank JD (1966): Target Complaints as Criteria for Improvement. *American Journal of Psychotherapy*, 20: 184-192.
- Baucom DA and Lester GW (1986): The usefulness of cognitive restructuring as an adjunct to behavioral marital therapy. *Behavior Therapy*, 17: 385-403.
- Bavelas JB (1975): Systems analysis of dyadic interaction: the role of interpersonal judgment. *Behavioral Science*, 20: 213-222.
- Beahrs J (1982): *Unity and Multiplicity*. Brunner/Mazel, New York.
- Beekers RJM (1982): *Interpersoonlijke vaardigheidstherapieën voor kansarmen*. Swets & Zeitlinger, Lisse.
- Beutler LE and Hamblin DL (1986): Individualized outcome measures of internal change: methodological considerations. *Journal of Consulting and Clinical Psychology*, 54: 48-53.
- Black M (1962): *Models and metaphors. Studies in language and philosophy*. Cornell University Press, Ithaca, New York.
- Boelens W, Debats D (1983): Vergelijking van een puur gedragsmatige met een puur cognitieve behandeling voor depressieve patiënten. In: Beer R, Mulders HSF (eds): *Psychologische benaderingswijzen van depressie*. Swets & Zeitlinger, Lisse (The Netherlands).
- Bogdan JL (1984): Family organization as an ecology of ideas: an alternative to the reification of family systems. *Family Process*, 23: 375-388.
- Braten S (1984): The third position: beyond artificial and autopoietic reduction. In: Geyer F and Zouwen J van der (eds), *Sociocybernetic paradoxes: observation, control and evolution of self-steering systems*. Sage Publ, London.
- Broom BL (1984): Consensus about the marital relationship during transition to parenthood. *Nursing Research*, 33: 223-228.

Consensus in Psychotherapy

Braito R, Dean D, Powers E, Bruton B (1981): The inferiority game: perceptions and behavior. *Sex Roles*, 7: 65-72.

Burger A (1978): Imaginaire confrontatie. In: Orlemans JWG (ed), *Handboek voor gedragstherapie*. Van Loghum Slaterus, Deventer.

Campbell DT and Fiske DW (1959): Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56: 81-105.

Calonico JM, Thomas DL (1973): Role-taking as a function of value similarity and affect in the nuclear family. *Journal of Marriage and the Family*, 35: 655-665.

Chapman LJ and Chapman JP (1967): Genesis of popular but erroneous psychodiagnostic observation. *Journal of Abnormal Psychology*, 72: 193-204.

Christensen L and Mendoza JL (1985): A method of assessing change in a single subject: an alteration of the RC-Index. *Behavior Therapy*, 17: 305-308.

Cline VB, Jackson SL, Klein N, Mejia J and Turner C (1987): Marital therapy outcome measured by therapist, client and behavior change. *Family Process*, 26: 255-268.

Coe WC, Curry AE, Kessler DR (1969): Family interactions of psychiatric inpatients. *Family Process*, 8: 119-30.

Cohen J (1960): A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20: 37-46.

Coleman BD, Manes MA, Willig MR and Hsieh YH (1982): Randomness, area and species richness. *Ecology*, 63: 1121-1133.

Cronbach LJ & Furby L (1970): How we should measure "change" - or should we? *Psychological Bulletin*, 74: 68-80.

Cronbach L (1970): *Essentials of psychological testing (3rd ed)*. Harper & Row, New York.

Dell PF (1982): In search of truth: on the way to clinical epistemology. *Family Process*, 21: 407-414.

Dell PF (1985): Understanding Bateson and Maturana: toward a biological foundation for the social sciences. *Journal of Marital and Family Therapy*, 11: 1-20.

Dember WN (1969): *The psychology of perception*. Holt, Rinehart & Winston, London.

- Derogatis LR (1977): *SCL-90: Administration, Scoring and Procedures Manual-I for the (R)vised Version*. John Hopkins University, Baltimore.
- De Tempe J (1988): Grief therapy from an anthropological point of view. In: Hart O van der (ed), *Coping with loss. The therapeutic use of leave-taking rituals*. Irvington, New York.
- Dewey J (1958): *Experience and nature*. Dover, New York.
- Dice LR (1945): Measures of the amount of ecological association between species. *Ecology*, 26: 297-302.
- Doherty WJ (1986): Quanta, quarks and families: implications of quantum physics for family research. *Family Process*, 25: 249-264.
- Dormaar JMM (1987): February Man Revisited. *American Journal of Clinical Hypnosis*, 30/1: 28-35.
- Dormaar JMM (1988): Exposure and leave-taking rituals. In: Van der Hart O, *Coping with Loss. The therapeutic use of leave-taking rituals* (p. 125-134). Irvington, New York.
- Dormaar M, Dijkman-Caes CIM, deVries MW (1988): Psychotherapeutic Outcome: Measuring Reliable Clinical Change. *Psychotherapy and Psychosomatics*, 50: 95-101.
- Dormaar M, Dijkman-Caes CIM and deVries MW (1989): Consensus in patient-therapist interactions. A measure of the therapeutic relationship related to outcome. *Psychotherapy and Psychosomatics*, accepted for publication.
- Dormaar M and Olthof J (1989): Individualized metaphors in hypnotherapy. Submitted for publication.
- DSM-III-R (1987): Diagnostic and Statistical Manual of mental disorders (3rd ed, revised). *American Psychiatric Association*, Washington DC.
- Edelstien MG (1981): *Trauma, trance and transformation. A clinical guide to hypnotherapy*. Brunner/Mazel, New York.
- Emmelkamp PMG, Mersch PP, Vissia E and Helm M van de (1985): Social Phobia: a comparative evaluation of cognitive and behavioral interventions. *Behaviour Research and Therapy*, 23: 365-369.
- Emmelkamp PMG, Mersch PP, Vissia E et al. (1985): A Comparative Evaluation of Cognitive and Behavioral Intervention. *Behaviour Research and Therapy*, 23: 366-369.

Consensus in Psychotherapy

Erickson MH (1967): Further techniques of hypnosis-utilization techniques. In: Haley J (ed): *Advanced techniques of hypnosis and therapy. Selected papers of MH Erickson*. Grune & Stratton, New York.

Erickson MH and Rossi EL (1979): *Hypnotherapy: An exploratory casebook*. Irvington Publishers, New York.

Everitt BS (1980): *Cluster analysis (2nd ed)*. Gower, London.

Farrell TB (1976): Knowledge, consensus and rhetorical theory. *Quarterly Journal of Speech*, 62: 1-14.

Festinger L (1962): *A theory of cognitive dissonance*. Stanford University Press, Stanford, Calif.

Fields JM, Schuman H (1977): Public beliefs about the beliefs of the public. *The Public Opinion Quarterly*, 40: 427-448.

Fleischaker GR (1984): The traditional model for perception and theory of knowledge: its metaphor and two recent alternatives. *Behavioral Science*, 29: 40-50.

Flynn JR (1987): Massive IQ gains in 14 nations: what IQ tests really measure. *Psychological Bulletin*, 101: 171-191.

Foerster H von (1981): *Observing systems*. Intersystems Publ, Seaside CA.

Fowler K (1988): Hypnotic transformation. Three studies of theatrical role-playing. *The International Journal of Clinical and Experimental Hypnosis*, 36: 249-255.

Frances A (1980): The DSM-III personality disorders section: a commentary. *American Journal of Psychiatry*, 137: 1050-1054.

Frank JD (1968): The role of hope in psychotherapy. *International Journal of Psychiatry*, 5: 383-395.

Frank JD (1969): Common features account for effectiveness. *International Journal of Psychiatry*, 7: 122-127.

Frank JD (1971): Therapeutic factors in psychotherapy. *American Journal of Psychotherapy*, 25: 350-361.

Frank JD (1974): *Persuasion and healing. A comparative study of psychotherapy*. Schocken Books, New York.

- Frank JD (1974a): Therapeutic components of psychotherapy. *Journal of Nervous and Mental Diseases*, 159: 325-342.
- Frank JD (1974b): Common features of psychotherapies and their patients. *Psychotherapy and Psychosomatics*, 24: 368-377.
- Frank JD (1975): General psychotherapy: the restoration of morale. In: Freedman D and Dyrud J (eds): *American Handbook of Psychiatry (2nd ed)*, Vol. 5: 117-132. Basic Books, New York.
- Frank JD (1979): The present status of outcome studies. *Journal of Consulting and Clinical Psychology*, 47: 310-316.
- Frank JD (1986): Failures in psychotherapy. *Readings: A Journal of Reviews*, 13-14.
- Frank JD, Hoehn-Saric R, Imber SD, Liberman BL and Stone AR (1978): *Effective ingredients of succesful psychotherapy*. Brunner Mazel, New York.
- Garfield SL, Bergin AE (eds) (1978): *Handbook of Psychotherapy and Behavior Change*. Wiley, New York.
- Glaserfeld E von (1984): An introduction to radical constructivism In: Watzlawick P (ed): *The invented reality*. Norton & Cy., New York.
- Goldstein AP (1973): *Structured learning therapy: toward a psychotherapy for the poor*. Academic Press, New York.
- Gordon D (1978): *Therapeutic Metaphors*. Meta Publ, Cupertino.
- Gordon D (1984): *Personal communication*. Workshop Therapeutic Metaphors, Amsterdam.
- Green RG, Kolvzon MS, Vosler NR (1985): The Beavers. Timberlawn model of Family Competence and the Circumplex Model of Family Adaptability and Cohesion: Separate but Equal? *Family Process*, 24: 385-398.
- Greer FL (1980): Prognostic expectations and outcome of brief therapy. *Psychological Reports*, 46: 973-974
- Grignon RAM (1975): *Psychometrics of similarity*. Academic Press, New York.
- Guppy N, Goyder JC (1984): Consensus on occupational prestige: a reassessment of the evidence. *Social Forces*, 62: 709-725.
- Habermas J (1975): *Legitimation crisis*. Beacon, Toronto.

Consensus in Psychotherapy

Haley J (ed) (1967): *Advanced techniques of hypnosis and therapy. Selected papers of Milton H. Erickson, M.D.*, Grune & Stratton, New York.

Haley J (1972): *Uncommon therapy. The psychiatric techniques of MH Erickson*. Norton, New York.

Hargens LL (1975): Anomie und Dissens in wissenschaftlichen Gemeinschaften. *Köllner Zeitschrift für Soziologie und Sozialpsychologie*, 375-392.

Hargens LL (1978): Theory and method in the sociology of science. *Sociological Inquiry*, 48: 121-139.

Hargens LL, Hagström WO (1982): Scientific consensus and academic status attainment patterns. *Sociology of Education*, 55: 183-196.

Hart O van der (1982): *Personal communication*.

Hart O van der (1983): *Rituals in psychotherapy; transition and continuity*. Irvington, New York.

Hart, O. van der (1985): Metaphoric and symbolic imagery in the hypnotic treatment of an urge to wander: A case report. *Australian Journal of Clinical and Experimental Hypnosis*, 13: 83-95.

Heider F (1958): *The psychology of interpersonal relations*. Wiley, New York.

Hiller P (1973): Social reality and social stratification. *Sociological Review*, 21: 77-99.

Hoffman L (1981): *Foundations of family therapy: a conceptual framework for systems change*. Basic Books, New York.

Hubalek Z (1982): Coefficients of association and similarity, based on binary (presence-absence) data: an evaluation. *Biological Review*, 57: 669-689.

Huizinga J (1955): *Homo ludens: A study of the play element in culture*. Beacon Press, Boston.

Jaccard P: Nouvelles recherches sur la distribution florale. *Bull Soc Vaud Sc Nat* 1908, XLIV; 163: 223-270.

Jaco DE, Shepard JM (1975): Demographic homogeneity and spousal consensus: a methodological perspective. *Journal of Marriage and the Family*, 37: 161-169.

- Jacobson NS (1984): A component analysis of behavioral marital therapy: the relative effectiveness of behavior exchange and communication/ problem-solving training. *Journal of Consulting and Clinical Psychology*, 52: 295-305.
- Jacobson NS, Follette WC and Revenstorf D (1984a): Psychotherapy outcome research: methods for reporting variability and evaluating clinical significance. *Behavior Therapy*, 15, 365-352.
- Jacobson NS, Follette WC, Revenstorf D, Baucom DH, Hahlweg K and Margolin G (1984b): Variability in outcome and clinical significance of behavioral marital therapy: a reanalysis of outcome data. *Journal of Consulting and Clinical Psychology*, 52: 497-504.
- Jacobson NS and Follette WC (1985): Clinical significance of improvement resulting from two behavioral marital therapy components. *Behavior Therapy*, 16: 249-262.
- Jacobson NS, Follette VM, Folette WC et al. (1985): A component analysis of behavioral marital therapy: 1 year follow-up. *Behaviour Research and Therapy*, 23: 549-555.
- Jacobson R and Halle M (1980): *Fundamentals of language (4th ed.)*. Mouton, The Hague.
- Janov A (1970): *The primal scream therapy; the cure for neurosis*. Sphere Books, London.
- Johnson BE and Millie DF (1982): The estimation and applicability of confidence intervals of Stander's Similarity Index (SIMI) in algal assemblage comparisons. *Hydrobiologica*, 89: 3-8.
- Kaplan HI, Sadock BJ (eds) (1985): *Comprehensive Textbook of Psychiatry (4th ed.)*. Williams & Wilkins, Baltimore.
- Kazdin AE and Wilson GT (1978): Criteria for evaluating psychotherapy. *Archives of General Psychiatry*, 35: 407-416.
- Keeney BP (1979): Ecosystemic epistemology: an alternative paradigm for diagnosis. *Family Process*, 18: 117-129.
- Keeney BP (1983): *Aesthetics of change*. Guilford, New York.
- Kenny DA, Albright L (1987): Accuracy in Interpersonal Perception: A Social Relations Analysis. *Psychological Bulletin*, 102: 390-402.
- Kempler W (1974): *Principles of Gestalt family therapy*. Joh. Nodahls Trykkevi, Oslo.
- Kerlinger FN (1973): *Foundations of behavioral research (2nd ed.)*. Holt, Rinehart and Winston, New York.

Consensus in Psychotherapy

Kiesler DJ (1977): *Use of Individualized Measures in Psychotherapy and Mental Health Program-Evaluation Research: A Review of Target Complaints, Problem Oriented Record and Goal Attainment Scaling*. Internal report, Virginia Commonwealth University.

Klapp OE (1957): The concept of consensus and its importance. *Sociology and Social Research*, 41: 336-342.

Klerman GL, Vaillant GE, Spitzer RL and Michels R (1984): A debate on DSM-III. *American Journal of Psychiatry*, 141: 539-553.

Korzybski A (1941): *Science and sanity*. Science Press, New York.

Koss MP, Graham JR, Kirkhart K, Post G, Kirkhart RO and Silverberg RS (1983): Outcome of eclectic psychotherapy in private psychological practice. *American Journal of Psychotherapy*, 37: 400-410.

Kuhn TS (1970): *The structure of scientific revolutions (2nd ed)*. University of Chicago Press, Chicago.

Laforge R and Suczek RF (1955): The interpersonal dimension of personality III: An interpersonal checklist. *Journal of Personality*, 24: 34-55.

Laing R, Philipson H, Lee AR (1966): *Interpersonal perception: a theory and a method of research*. Springer, New York.

Lakoff G and Johnson M (1980): *Metaphors we live by*. The University of Chicago Press, Chicago.

Lankton R and Lankton CH (1983): *The answer within: a clinical framework of Ericksonian hypnotherapy*. Brunner/Mazel, New York.

Leary T (1957): *Interpersonal diagnosis of personality. A functional theory and methodology of personality evaluation*. Ronald, New York.

Lehrer K and Wagner C (1981): *Rational consensus in science and society*. A philosophical and mathematical study. Reidel, Dordrecht.

Levin J, Arluke A (1983): Attitude similarity and parent-child attractiveness. *The Journal of Social Psychology*, 120: 223-228.

Lewis GL (1980): The relationship of conceptual development to consensus: an exploratory analysis of three subfields. *Social Studies of Science*, 10: 285-308.

Lieberson S (1969): Measuring population diversity. *American Sociological Review*, 34: 850-862.

- Lilienfeld R (1978): *The rise of systems theory. An ideological analysis*. Wiley, New York.
- Lincoln JR (1984): Analyzing relations in dyads. Problems, models, and an application to interorganizational research. *Sociological Methods & Research*, 13: 45-76.
- Marks G, Miller N (1987): Ten years of research on the false-consensus effect: an empirical and theoretical review. *Psychological Bulletin*, 102: 72-90.
- Markoff J (1982): Suggestions for the measurement of consensus. *American Sociological Review*, 47: 290-298.
- Maturana HR and Varela FJ (1980): *Autopoesis and cognition. The realization of the living*. Reidel, Boston.
- Maturana H (1988): Reality: the search for objectivity or the quest for a compelling argument. *The Irish Journal of Psychology*, 9: 25-82.
- Mavissakalian M (1986): Clinically significant improvement in agoraphobia research. *Behaviour Research and Therapy*, 24: 369-370.
- McCulloch, WS (1965): *Embodiments of mind*. M.I.T. Press, Cambridge, Mass.
- McKinnon NJ, Summers GF (1976): Homogeneity and role consensus: a multivariate exploration in role analysis. *The Canadian Journal of Sociology*, 1: 439-462.
- McLeod JM, Chaffee SH (1973): Interpersonal approaches to communication research. *American Behavioral Scientist*, 16: 469-499.
- Mead GH (1934, 1962): *Mind, self and society. From the standpoint of a social behaviorist*. University of Chicago Press, Chicago.
- Meissner WW (1980): Theories of personality and psychopathology: classical psychoanalysis In: Kaplan HI, Freedman AM, Sadock BJ (eds), *Comprehensive Textbook of Psychiatry* (3rd ed), vol 1, Williams & Wilkins, Baltimore
- Menninger K, Ellenberger H, Pruyser P and Mayman M (1958): The unitary concept of mental illness. *Bulletin of the Menninger Clinic*, 22: 4-12.
- Merton RK (1968): *Social theory and social structure. Enlarged edition*. Free Press, New York.
- Miller LD (1977): Dyadic perception of communicator style. Replication and confirmation. *Communication Research*, 4: 87-112.
- Mischel W (1968): *Personality and assessment*. Wiley, New York.

- Myers JK et al. (1984): Six-month prevalence of psychiatric disorders in three communities. *Archives of General Psychiatry*, 41, 959-967.
- Neal AG, Groat HT (1976): Consensus in the marital dyad: couples perceptions of contraception, communication, and family life. *Sociological Focus*, 9: 317-329.
- Newcomb TM (1955): An approach to the study of communicative acts. In: Hare AP, Borgatta EF and Bales RF (eds.), *Small groups*: 149-163. Knopf, New York.
- Newcomb TM, Turner RH, Converse PE (1966): *Social psychology. The study of human interaction (2nd ed)*. Routledge & Kegan Paul, London.
- Nunnally JC (1978): *Psychometric theory (2nd ed)*. McGraw-Hill, New York.
- O'Gorman HJ (1975): Pluralistic ignorance and white estimates of white support for racial segregation. *The Public Opinion Quarterly*, 39: 313-330.
- O'Keefe GJ (1973): Coorientation variables in family study. *American Behavioral Scientist*, 16: 513-536.
- Parsons T and Shils EA (eds) (1951): *Toward a general theory of action*. Harvard University Press, Cambridge Mass.
- Parsons T (1960): *Structure and process in modern societies*. Free Press, New York.
- Parsons T (1964): A functional theory of change. In: Etzioni A and Etzioni E (eds): *Social change*. Basic Books, New York.
- Polanyi M (1962): *Personal knowledge. Towards a post-critical philosophy*. Harper & Row, New York.
- Popping R (1983): *Overeenstemmingsmaten voor nominale data (Measures of agreement for nominal data)*. Doctoral thesis, University of Groningen.
- Quesada E, Ventosa A, Rodriguez-Valera F, Megias L and Ramos-Cormenzana A (1983): Numerical taxonomy of moderate halophilic gramnegative bacteria from hypersaline soils. *Journal of General Microbiology*, 129: 2649-2657.
- Rabkin R (1986): Essay: A tower of Babble: the sociology of body and mind. *Family Process*, 25: 153-163.
- Rachman S (1980): Emotional processing. *Behaviour Research and Therapy*, 18: 51-60.

- Ramsay RW and Happée JA (1977): The stress of bereavement; components and treatment. In: Spielberger CD & Sarason IG (eds), *Stress and anxiety*, vol. 4. Hemisphere Publ, Washington D.C.
- Ramsay RW (1979): Rouwtherapie; de gedragstherapeutische behandeling van pathologische rouwproblemen. In: Orlemans JWG (ed): *Handboek voor gedragstherapie*. Van Loghum Slaterus, Deventer.
- Rose AM (1954): *Theory and mind in the social sciences*. University of Minnesota Press, Minneapolis.
- Sanders GS, Mullen B (1983): Accuracy in perceptions of consensus: differential tendencies of people with majority and minority positions. *European Journal of Social Psychology*, 13: 57-70.
- Sartre JP (1938): *La Nausée*. Gallimard, Paris.
- Scheff TJ (1967): Toward a sociological model of consensus. *American Sociological Review*, 32: 32-46.
- Scheff TJ (1973): Intersubjectivity and emotion. *American Behavioral Scientist*, 16: 501-511.
- Scheff TJ (1979): *Catharsis in healing, ritual and drama*. University of California Press, Berkeley.
- Scheler M (1980): *Problems of a sociology of knowledge*. Routledge & Kegan Paul, London.
- Schelling, TC (1963): *The strategy of conflict*. Oxford University Press, New York.
- Schulman ML (1974): Idealization in engaged couples. *Journal of Marriage and the Family*, 36: 139-147.
- Schutz A (1962): *Collected papers 1. The problem of social reality*. Nijhoff, The Hague.
- Schwartzman J (1982): Creativity, pathology and family structure: a cybernetic metaphor. *Family Process*, 21: 113-127.
- Schwartzman J (1984): Family theory and the scientific method. *Family Process*, 23: 223-236.
- Segal L (1986): *The dream of reality. Heinz von Foerster's constructivism*. Norton, New York.

Consensus in Psychotherapy

- Shannon CE and Weaver W (1949): *The mathematical theory of communication*. University of Illinois Press, Urbana, Illinois.
- Singer JL and Pope KS (eds) (1978): *The power of human imagination. New methods in psychotherapy*. Plenum, New York.
- Smith, ML, Glass GV, Miller TI (1980): *The benefits of psychotherapy*. John Hopkins University Press, Baltimore.
- Sneath PHA and Sokal RR (1973): *Numerical taxonomy*. Freeman, San Francisco.
- Snijders TAB, Dormaar M, Van Schuur WH, Dijkman-Caes C and Driessen G (1989): Distribution of some similarity coefficients for dyadic binary data in the case of associated attributes. *Journal of Classification*, accepted for publication.
- Sokal RR and Michener CD (1958): A statistical method for evaluating systematic relationships. *University of Kansas Scientific Bulletin*, 38: 1409-1438.
- Spencer Brown G (1969): *Laws of Form*. Allen & Unwin, London.
- Spillman B, Bezdek J, Spillman R (1979): Development of an instrument for the dynamic measurement of consensus. *Communication Monographs*, 46: 1-12.
- Strauss JS and Carpenter WT (1981): *Schizophrenia*. Plenum, New York.
- Strupp HH and Bergin AE (1969): Some empirical and conceptual bases for coordinated research in psychotherapy: a critical review of issues, trends and evidence. *International Journal of Psychiatry*, 7: 18-90.
- Strupp HH and Hadley SW (1977): A tripartite model of mental health and therapeutic outcomes. *American Psychologist*, 32: 187-196.
- Sullivan HS (1974): *Schizophrenia as a human process*. Norton, New York.
- Stephenson W (1953): *The study of behavior*. University of Chicago Press, Chicago.
- Thomas WI, Thomas DS (1928): *The child in America: Behavior problems and programs*. A. Knopf, New York (p. 572).
- Thompson KF (1979): The case against relaxation. In: Burrows GD et al. (eds), *Hypnosis*. Elsevier/North Holland Biomedical Press, Amsterdam.
- Tournier M (1967): *Vendredi on les limbes du Pacifique*. Gallimard, Paris.
- Tversky A (1977): Features of similarity. *Psychological Review*, 84: 327-352.

- Varela FJ (1976): Not one, not two. *Coevolution Quarterly*, 11: 62-67.
- Velden K van der, Dijck R van (1977): Wat is directieve therapie? In: Velden K van der (ed), *Directieve therapie I*. Van Loghum Slaterus, Deventer.
- Velden K van der, Habekotté S (1980): Vaak, kort en eventueel op ongelegen tijden. In: Velden K van der (ed), *Directieve therapie II*. Van Loghum Slaterus, Deventer.
- Velden K. van der (1987): De les van Charcot (The lesson from Charcot). *Directieve Therapie en Hypnose*, 7: 5-21.
- Vertommen H and Rochette F (1979): *Vragenlijst interpersoonlijke relaties. Handleiding*. Swets & Zeitlinger, Lisse.
- Von Bertalanffy L (1979): *General systems theory: foundations, development, applications*. Rev. ed. Braziller, New York.
- Von Glaserfeld E (1984): An introduction to radical constructivism, p. 17-40. In: Watzlawick P (ed): *The invented reality*. Norton, New York.
- Wackman DB (1973): Interpersonal communication and coorientation. *American Behavioral Scientist*, 16: 537-550.
- Washington HG (1984): Diversity, biotic and similarity indices. A review with special relevance to aquatic ecosystems. *Water Research*, 18: 653-694.
- Watzlawick P, Beavin JH, Jackson DJ (1967): *Pragmatics of human communication; A study of interactional patterns, pathology and paradoxes*. Norton, New York.
- Watzlawick P, Weakland JH and Fisch R (1974): *Change; principles of problem formation and resolution*. Norton, New York
- Watzlawick P (1984): *The invented reality. How do we know what we believe we know? Contributions to constructivism*. Norton, New York.
- White JM (1982): Dyadic systems analysis. *Behavioral Science*, 27: 104-117.
- White JM (1985): Perceived similarity and understanding in married couples. *Journal of Social and Personal Relationships*, 2: 45-57.
- Williams JBW (1985): The multi-axial system of DSM-III: Where did it come from and where should it go? *Archives General Psychiatry*, 42: 181-186.
- Wishart (1978): *Cluster user manual (3rd ed)*. Program Library Unit, Edinburgh University, Edinburgh.

Consensus in Psychotherapy

Witzlum E, Hart O van der & Friedman B (1988): The use of metaphors in psychotherapy. *Journal of Contemporary Psychotherapy*, 18: 270-290.

Zeig JK (ed) (1980): *A teaching seminar with Milton H. Erickson, M.D.* Brunner/Mazel, New York.

Ziman J (1968): *Public knowledge*. Cambridge University Press, London.

Zuckerman M, Mann RW, Bernieri FJ (1982): Determinants of consensus estimates: attribution, salience, and representativeness. *Journal of Personality and Social Psychology*, 42: 839-852.

Index

Adaptation	133, 141
- at the dyadic level	140
- at the individual level	140
Social	140
Affect	
-ive and cognitive sides of relationship	30
Level of	53
Affiliation	135
Analogy	131
Attractiveness, parent-child	53
Attribute	17, 18, 92, 108, 110-112, 116
Attribute, personal	56
Attribution	21, 54, 55
Classification	
Basic criteria in	130
Functional	130
Historic	131
Psychiatric	130-136
Taxonomic	133
Clinical significance	83
Clinically significant change	85-90
Coefficient	
Correlation	50, 52, 108
Dice	110
Kappa	116, 117
Relationship expressed as	18
Reliability	52
Similarity	110-113, 116, 129, 132
Simple matching	110
Cognition	30
Cognitions	34, 35, 93
Common factors	84
Communication	23, 27, 80
- and co-orientation	28, 109
- and consensus	28
- and metacommunication	19
Function of	34
Function of - in balance model	30
Marital	34
Complementarity	
Principle of	143
Concept	18
Confidence interval	110

Consensus in psychiatry

Consensus	14, 15, 129
- and co-orientation	28, 29, 35-41
- and dissensus	48
- and Jaccard index	114
- and outcome	101, 102
- and public opinion research	47
- as accuracy	43
- as agreement	28, 36, 96, 97, 99, 103, 109
- as congruency	96, 97, 99, 103
- as experience	19
- as realization	28, 39
- as understanding	28, 37, 38, 96, 97, 99, 101, 103
- as unifying concept	26
- in psychiatric setting	48
- in sociology of science	45
- in therapeutic relationship	95
as a relationship	17
Dynamic approach to	57
Dynamic approach to measurement of	57
Experienced	95, 97-100, 103
Interpersonal	53, 96, 99, 101
Intrapersonal	53, 96
Main terms for levels of	42
Probabilistic approach to measurement of	56
Significance of observed values of	115
Conservative bias	47
Construction	77, 78, 125
Construction of reality	21, 81
Context	76
Cooperation	19, 20, 27, 96, 105
Coordination	19, 20, 23, 28, 129
Co-orientation	
- model; see Model	
and consensus; see Consensus	
and communication; see Communication	
Culture	
Western	27
Cybernetics	12, 20, 80
Dimension	
Affiliation	135
Dominance-submission	135
Distribution	
Independent	37
Observed	37
Permutation	114
Probability	114
Domain	
Consensual	20, 127
Dominance	55, 135
- and submissiveness	136

Dyad	109, 113, 117
-ic structure	31, 32
Marital	49, 50
Patient-therapist	105, 106
Emotion	18, 61
- and human relationship	19
- as definition of relationship	19
Extinction of	67
Epistemology	
Cartesian	24
Exposure technique	61, 64, 66
Factors	
Common - in psychotherapy	143
Facts, scientific -	
as relationships	17
Feature	
Common	108, 118, 132
Distinctive	108, 118, 132
Feature space	118
Fitting	67, 78, 136
- explanations	78
Heuristic	
- principle	18, 143
Homogeneity	
demographic	50
Homology	131
Hypnosis	71
Self-	76
Hypnotherapy	71-79
Hypothesis, null	113
Idealization	49
Ignorance	28
Ignorance, pluralistic	28
Interaction	
Range of	133, 134
Intersubjectivity	28, 58
- and emotion	44
Jaccard	
-values	38
Jaccard index	108, 109, 111, 115
- and consensus	36, 99, 109, 114
coefficient, similarity	36
Expected distribution (math.)	113
Metaphor	27, 92, 118, 119, 129
- and metonymy	126
-ic domain	74, 122, 126
-ic Kernel statement	121
-ic strategy	74
Individualized	121-127
similarity in	107, 118
Therapeutic	118-120

Consensus in psychiatry

Model	
co-orientation	28, 29, 96
Ratio - of similarity	108
Model, balance or consistency	30-35
Mourning	61, 66, 68
Myth, therapeutic	77
Objective description	12
Observer	20, 54, 116, 143
Observer dependence	12, 13
Operationalization	
- of consensus in a group	36, 39-41
Outcome of psychotherapy	83-91
Paradox	58
- in psychotherapy model	142
Perception	
- as beliefs	30
Looking-glass	47
Perception, interpersonal	47
Issues in	29
Psychological measurement	
Critique of	55, 92
Q-methodology	18, 52
R-methodology	18
Recognition	25, 44, 58
Relationship	17-19
Complementary	103, 135, 136, 140
Consensual	26
Control of	80
Definition of	134, 135
Equilibrium in complementary and symmetrical -s	141
Structure of	140
Symmetrical	135, 136, 140
Therapeutic	80
Scientific facts as -; see facts and emotion; see emotion	
Reliable change	83
Reliable clinical change	86, 87, 89-91, 101
Ritual, therapeutic	61, 66, 68
Role-taking ability	52
Similarity	107
- and difference	66
see also; Coefficient, Metaphor	
Social desirability	47, 49, 55, 93
Unit, operational taxonomic	110, 113
Validation of DSM-III-R	132
Validity	
Construct	106
Predictive	106

Appendix: Instruments for consensus measurement

MD\SP-309.89
6 maart 1989

INTERPERSONAL PERCEIVED CONSENSUS CHECKLIST - PATIENT FORM

Patient identification:

Date:

Name therapist:

After ... sessions

Please put a mark on the line below each statement.
Your mark indicates to what degree you endorse the statement.

1a. My somatic complaints are a major problem to me.

agree

disagree

1b. How do you guess your therapist would answer this question regarding you?

agree

disagree

2a. A major problem to me is that I have to make certain changes within myself.

3a. A major problem to me is that they should treat me differently at home.

4a. A major problem to me is that I should act differently at home.

5a. A major problem to me is that my partner and me should change our ways of behaving towards each other.

6a. A major problem to me is that other people (outside home) should behave differently towards me.

7a. A major problem to me is that I should act/behave differently towards other people.

8a. A major problem to me regards the practical level (housing, work, finances, etc.).

Items 2-8 are presented in the same form as item 1.

INTERPERSONAL PERCEIVED CONSENSUS CHECKLIST - THERAPIST FORM

Patient identification:

Date:

Name therapist:

After ... sessions

Please put a mark on the line below each statement.

Your mark indicates to what degree you endorse the statement.

1a. In my view his/her somatic complaints are a major problem.

agree

disagree

2a. How do you guess your patient would answer this question?

agree

disagree

2b. In my view it is a major problem that he/she should make certain changes within him/herself.

3a. In my view it is a major problem that at home they should treat him/her differently.

4a. In my view it is a major problem that he/she should act differently at home.

5a. In my view it is a major problem that he/she and his/her partner should change their ways of behaving towards each other.

6a. In my view it is a major problem that other people (outside home) should behave differently towards him/her.

7a. In my view it is a major problem that he/she should act/behave differently towards other people.

8a. In my view a major problem regards the practical level (housing, work, finances, etc.).

Items 2-8 are presented in the same form as item 1.

INSTRUMENT FOR INTERPERSONAL PERCEIVED CONSENSUS CHECKLIST (IPC)

The respondent (patient/therapist) should not be expected to share your distinctions between complaints, problems, needs, demands, hopes, expectations, goals. Any topic that he/she wants to be different, or thinks that it should be, but does not know exactly how to alter or is still unsuccessful in achieving, may be called "a problem". The adjective "major" is merely added to distinguish problems from difficulties: topics the patient/therapist is uncomfortable with, but accepts as part of his/her life or sees as unalterable.

The 8 problem categories as mentioned are a condensed combination of goal-attainment categories (Garwick & Brintrnall, 1977) and health problem classification (Regier et al., 1982).

1. Ask for somatic complaints and attributions of patient/therapist, e.g.: depressed because of somatic complaints or having these complaints because depressed, somatic disease, "it runs in the family", unknown, hypochondric or "weird" attributions. If so, bring up these attributions, if necessary, under other headings (e.g. 2, 3 and 6) to check if a problem also exists at this level. Check if patient/therapist feels, regardless of attributed cause, that the somatic complaints should be focus of attention in treatment and therefore may be called a "major problem".
2. Ask question on:
 - aggression (anger, irritability, antisocial behavior etc.)
 - abuse (dependency problems)
 - cognitive abilities (attention, concentration)
 - decision making, autonomy
 - depression (sleeplessness, loss of appetite and interest, self-critical, despair, mourning)
 - psychopathological symptoms if complained about: phobias, delusions, hallucinations, obsessions, compulsions
 - lack of assertiveness
 - self-image, self-confidence
 - "not able to cope", shame, guilt, homesickness
3. a) Which persons does the patient/therapist consider as "home"?
 - b) Differently: what should be done more (or less) toward him at home?

More positive behaviors (support, understanding, respect, attention, warmth, etc.), less negative behaviors (aggression, ignoring, criticism, threats etc.), more (or less) contact (leaving alone, more or less invitations to common activities, being talked to

more or less), other behaviors (granting of privileges, power, influence, sex, obedience).

4. Differently toward whom especially?

"Differently" can be asked questions as specified in 3.

5. Again the same questions as in 3 and 4.

The crucial difference is that here the patient/therapist sees the problem(s) as interactional, relational instead of the one-sided interpunction in 3 and 4.

6. "Other people" - who especially? When, where, what?

"Differently" can be questioned as specified in 3.

7. Same as 6, but patient/therapist sees interpunction ("locus of control") the other way round.

8. Topics may be:

- education, training
- finances, social security benefits
- being cared for (hostel, guesthouse)
- housing (wants to move, home of his/her own)
- regulations (unemployment, disablement)
- treatment conditions (also support or admittance during crisis)
- juridical (esp. reduction of sentence)

Regier, DA, Burke JD, Burns BJ, et al. (1982): A proposed classification of social problems and psychological symptoms for inclusion in a classification of health problems. In: M. Lipkin jr., K. Kupka: Psychosocial factors affecting health. Praeger Publ., New York.

Garwick GB, Brintnall JE (1977): An introduction to goal attainment scaling, catalogue-assisted and the idea-book: 550 indications for use in setting goals. Technical assistance for program evaluation, Minneapolis.

EXPERIENCED CONSENSUS - PATIENT FORM

Patient identification:
Name therapist:

Date:
After ... sessions

1. To what extent - in your own opinion - your therapist is acquainted with the complaints or problems that are the reasons for your coming here?

☐☐☐☐☐

not at all
or hardly

moderately
or reasonably

quite well
to completely

2. Do you feel your therapist understands you?

☐☐☐☐☐

yes, completely
yes, quite
well

moderately
or reasonable

no, not at all
or hardly

3. Do you feel your therapist understands what's really bothering you?

☐☐☐☐☐

not at all
or hardly

moderately
or reasonably

quite well
to completely

4. Do you feel you understand your therapist?

☐☐☐☐☐

not at all
or hardly

moderately
or reasonably

quite well
to completely

5. Do you feel your therapist and yourself have a different view on you and your complaints or problems?

☐

yes

☐☐

moderately
or reasonably

☐☐

no
or hardly

6. Do you and your therapist agree on the exact nature of your complaints or problems?

☐

yes

☐☐☐☐

no

7. Do you and your therapist agree which problem or complaint is the most important?

☐

yes

☐☐☐☐

no

8. Do you and your therapist agree which problem or complaint should be the first to be worked on?

☐

yes

☐☐☐☐

no

EXPERIENCED CONSENSUS - THERAPIST FORM

Patient identification:

Date:

Name therapist:

After ... sessions

1. To what extent are you - in your own opinion - acquainted with the complaints or problems that are the reasons for your patient's coming here?

☐☐☐☐☐not at all
or hardlymoderately
or reasonablyquite well
to completely

2. Do you feel you understand your patient?

☐☐☐☐☐yes, comp-
letely
yes, quite
wellmoderately
or reasonablyno, not at all
or hardly

3. Do you feel you understand what's really bothering your patient?

☐☐☐☐☐not at all
or hardlymoderately
or reasonablyquite well
to completely

4. Do you feel your patient understands you?

☐☐☐☐☐not at all
or hardlymoderately
or reasonablyquite well
to completely

5. Do you feel your patient and yourself have a different view on him/herself and his/her complaints or problems?

☐

yes (almost)
completely

☐☐

fairly,
quite

☐☐

no, not at all
or hardly

6. Do you and your patient agree on the exact nature of your complaints or problems?

☐

yes

☐☐☐☐

no

7. Do you and your patient agree which problem or complaint is the most important?

☐

yes

☐☐☐☐

no

8. Do you and your patient agree which problem or complaint should be the first to be worked on?

☐

yes

☐☐☐☐

no

INSTRUCTIONS FOR EXPERIENCED CONSENSUS QUESTIONNAIRE

Please stress that we are interested in the respondent's direct, subjective experience or impression. None of his/her opinions or feelings will be made known to anyone, especially not to his/her therapist or patient.

If needed, point out that there are subtle distinctions in the questions (e.g. 7 and 8) that he/she is requested to take into consideration.

GENERAL INFORMATION AND PRELIMINARY DATA ON RELIABILITY OF EXPERIENCED CONSENSUS QUESTIONNAIRES AND IPC-CHECKLIST:

In our field study a qualified research assistant collected data from patients and therapists. The instructions as provided here, had been written for this assistant.

The experienced consensus questionnaires were tested on 20 patients and their therapists for internal consistency. The questionnaire form used at that time consisted of only 6 out of the preset 8 items (item 4 and 8 were added later on). Cronbach's coefficient alpha was .60 for the patients and .58 for the therapist. When recoded to 3 out of 5 response categories (omitting the little used category "no" and the second one next to "no") the coefficient alpha for both groups rose to .72. Considering the test-length and the number of respondents this level is satisfactory.

The Interpersonally Perceived Consensus Checklist showed alpha-coefficients of .48 and .54 (patient-form, "a" and "b" questions) and of .52 and .58 (therapist-form, "a" and "b" questions). These low values were no reason to change the checklist as the different problem areas were not expected to be associated to a high degree.

in the computation of agreement and understanding with the Jaccard coefficient (Jaccard, 1908; Snijders et al., 1989), the IPC-scores were dichotomized on 25 mm from "agree". 25 mm corresponded to the average item-mean.

Summary

After an introductory overview (**Chapter 1**), I offer some theoretical remarks on the social sciences (**Chapter 2**), and discuss selected models of consensus and relevant research from sociological and psychological fields. Theory and research deriving from the co-orientation model and the balance model appear to be particularly pertinent to psychiatric practice (**Chapter 3**).

Chapter 4 and 5 contain case descriptions illustrating that consensus is but a part of a more encompassing concept of how patient and therapist arrive at a mutual fit. The treatment of a female patient with mourning problems (chapter 4) by means of a leave-taking ritual shows how the personal style of a patient and therapist fit in with course and form of the therapy. In a discussion of the hypnotherapy of an overworked, depressed male patient (chapter 5), the idea of 'fitting' - conceived as being both similar and dissimilar - is applied to the explanations, associations and relationship definitions of patient and therapist.

Chapter 6: Reliable clinical change is proposed as a means of assessing improvement in psychotherapy. The use of this twofold criterion with an aspecific outpatient population demonstrates that the distinction between patients and non-patients corresponds little to differences in scoring on a checklist of psychopathology.

Chapter 7 is a research report on the relationships of consensus between patient and therapist with the outcome of short-term treatment. The results show that three different aspects of consensus assessed after the first two sessions - namely agreement, the understanding of the patient by the therapist and consensus experienced by the patient - can partly predict the success of the treatment as assessed after a period of six months.

Chapter 8: Consensus can be defined as the degree of similarity between two persons over a set of features. The distribution of the Jaccard similarity coefficient under a suitable null-hypothesis of independence can be well approximated in a mathematical way. Based on their similarity coefficients a statistically significant (although perhaps clinically irrelevant) number of patient-therapist dyads evidenced a degree of consensus greater than that which could be attributed to chance.

Chapter 9: Reflecting on the use of metaphors in hypnotherapy, I submit that the similarities between dealing with problems and the patient's positive assets are the reason why metaphors can help the patient to overcome his difficulties. Theoretical considerations are offered on how to understand the impact of metaphors on human behavior and experience.

Chapter 10: The conclusion of the present thesis recommends the application of measures of similarity like those I have presented to compute consensus, for

assessing interpersonal relationships in general. This suggestion is further developed with proposals for an unorthodox way of determining reliability and validity of diagnostic categories according to the DSM-III-R and for a classification scheme of interactional diagnoses based on a criterion of adaptation.

Samenvatting

Na de inleiding (**Hoofdstuk 1**) en een korte wetenschapstheoretische inleiding (**Hoofdstuk 2**) worden enkele theoretische modellen over consensus besproken en ruim twintig sociologische en psychologische onderzoeken (**Hoofdstuk 3**). Vooral theorie en onderzoek over consensus volgens het coöriëntatiemodel en het balansmodel blijken relevant te zijn voor de psychiatrische praktijk.

Hoofdstuk 4 en 5 geven enkele gevalsbeschrijvingen om te illustreren dat consensus tussen patiënt en therapeut een deel is van het veelomvattende begrip van het bij elkaar (leren te) passen van patiënt, therapeut, probleem en behandelingstechniek. De behandeling van een vrouw met rouwproblemen (hoofdstuk 4) met behulp van een afscheidsritueel toont hoe de persoonlijke stijl van een patiënt en therapeut passen bij het verloop en de vormgeving van een therapie. Bij de bespreking van de hypnotherapeutische behandeling van een overwerkte en depressieve man (hoofdstuk 5) wordt het idee van 'bij elkaar passen' als tegelijk overeenkomen en niet overeenkomen, toegepast op de door patiënt en therapeut gehanteerde verklaringen, verbanden en relatie-definities.

Hoofdstuk 6: Betrouwbare klinische verandering is een maat om te bepalen wanneer van verbetering kan worden gesproken bij psychotherapeutische behandelingen. De toepassing van dit tweeledig criterium bij een specifieke ambulante patiëntenpopulatie laat o.a. zien dat het onderscheid tussen patiënten en niet-patiënten weinig overeenkomt met verschillen in uitslag op een klachtenlijst.

Hoofdstuk 7: Een onderzoeksverslag over de relaties van consensus tussen patiënt en therapeut als maat voor de kwaliteit van een psychotherapeutische relatie met het resultaat van kortdurende behandelingen. De resultaten laten zien dat drie verschillende aspecten van consensus vastgesteld na de eerste twee sessies - te weten overeenstemming, een goed begrijpen van de patiënt door de therapeut en consensus zoals ervaren door de patiënt - in zekere mate het succes van een behandeling, zoals vastgesteld na zes maanden, kunnen voorspellen.

Hoofdstuk 8: Consensus kan worden gedefinieerd als de mate van overeenkomst tussen twee personen over een aantal kenmerken. De onder de nulhypothese van onafhankelijkheid te verwachten verdeling van de Jaccard index als overeenkomstmaat kan uitstekend mathematisch worden benaderd. Een statistisch significant, maar voor klinische relevantie gering aantal patiënt-therapeut dyades blijkt bij de hier gehanteerde vaststelling van consensus een mate van overeenkomst te bezitten die niet aan kans is toe te schrijven.

Hoofdstuk 9: De toepassing van metaforen in hypnotherapeutische behandelingen leiden tot de conclusie dat overeenkomsten tussen het oplossen van problemen en zijn positieve kwaliteiten de patiënt kunnen helpen om zijn problemen te boven te komen. Theoretische overwegingen maken de invloed van metaforen op gedrag en beleven begrijpelijk.

Hoofdstuk 10: In de conclusie van dit proefschrift wordt de toepasbaarheid van overeenkomstmaten, zoals hier gebruikt voor het meten van consensus, betoogd voor een alternatieve wijze van het vaststellen van interpersoonlijke relaties. Voorstellen voor het vaststellen van validiteit en betrouwbaarheid van diagnose-categorieën volgens de DSM-III-R en voor een stelsel van interactionele diagnostiek gebaseerd op een adaptatie-criterium ondersteunen dit betoog.

Curriculum vitae

Jacobus Marius Maarten Dormaar was born August 29 1939, in Bloemendaal, the Netherlands. After graduating from high school at the Magister Alvinus Gymnasium in Sneek (Gymnasium Bèta, 1959), he studied medicine at the University of Leiden from 1959 until 1967.

After additional training in surgery, obstetrics and tropical medicine in Wageningen and Amsterdam, he worked from 1968 until 1971 as a medical officer at a comprehensive medical service in a rural district of Ethiopia. From 1971 until 1975 he was trained in psychiatry at the psychiatric department of the University of Groningen (head: Prof. Dr. W.K. van Dijk).

From 1975 until 1980 he worked as a social psychiatrist at the Social Psychiatric Service of Groningen, in a Community Mental Health Center in Friesland and in private practice. In november 1980 he joined the Community Mental Health Center 'RIAGG Maastricht e.o.' (director: H.B. Pomerantz) and the department of Social Psychiatry of the University of Limburg, Maastricht, set up by Prof. Dr. M.A.J. Romme (current chairman: dr. H.F. Kraan), where he did his research under the direction of prof. dr. M.W. deVries.