

Chronic pain : multiaxial assessment and behavioral mechanisms

Citation for published version (APA):

Lousberg, H. B. (1994). *Chronic pain : multiaxial assessment and behavioral mechanisms*. [Doctoral Thesis, Maastricht University]. Datawyse / Universitaire Pers Maastricht. <https://doi.org/10.26481/dis.19940225hl>

Document status and date:

Published: 01/01/1994

DOI:

[10.26481/dis.19940225hl](https://doi.org/10.26481/dis.19940225hl)

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.

CHRONIC PAIN

Multiaxial assessment and
behavioral mechanisms

CHRONIC PAIN

Druk

Datawyse | Universitaire Pers Maastricht

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Lousberg, Henri Béatrice

Chronic pain : multi-axial diagnostics and behavioral mechanisms / Henri Béatrice (Richel) Lousberg. - Maastricht : Universitaire Pers Maastricht - III. Thesis Maastricht. - With ref. - With summary in Dutch. ISBN 90-5278-124-9
Subject headings: chronic pain ; diagnostics / chronic pain ; operant conditioning.

Het verschijnen van dit proefschrift werd mede mogelijk gemaakt door financiële steun van:

- **Schering-Plough BV**, Postbus 70, 1180 AB Amstelveen
- **Sanofi Winthrop v.o.f.**, Postbus 97, 3140 AB Maassluis
- **Revalidatiecentrum Breda**, 4817 JW Breda
- **Revalidatiecentrum Amsterdam**, 1054 HW Amsterdam

CHRONIC PAIN

Multiaxial assessment and behavioral mechanisms

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Rijksuniversiteit Limburg te Maastricht,
op gezag van de Rector Magnificus, Prof. dr. H. Philipsen,
volgens het besluit van het College van Dekanen,
in het openbaar te verdedigen op
vrijdag, 25 februari 1994 om 14.00 uur

door

Henri Béatrice (Richel) Lousberg



UNIVERSITAIRE PERS MAASTRICHT

Promotor

Prof. dr. M.A. van den Hout

Co-promotores

Dr. N.H. Groenman

Dr. A.J.M. Schmidt

Beoordelingscommissie

Prof. dr. J.J.C.B. Bremer (voorzitter)

Dr. Th. Bouman (Rijksuniversiteit Groningen)

Prof. dr. H. Huijjer Abu-Saad

Prof. dr. H.G. Schmidt

Prof. dr. F.G. Zitman (Katholieke Universiteit Nijmegen)

CONTENTS

1. General introduction	1
1.1 An overview	1
1.2 The multidimensional nature of pain	2
1.3 Psychological models and treatment of chronic pain	4
1.4 Non-optimal treatment results and identification of patient sub-groups	10
1.5 Empirically based pain-patient sub-groups based on the 'MAP'-taxonomy	15
1.6 The MAP taxonomy: evaluative comments	17
1.7 Specification of research questions	18
2. The relationship between spouse solicitousness and pain behavior: searching for more experimental evidence	21
2.1 Summary	21
2.2 Introduction	21
2.3 Method	23
2.4 Results	24
2.5 Discussion	27
3. Operant conditioning of the pain experience	31
3.1 Summary	31
3.2 Introduction	31
3.3 Method	34
3.4 Results	36
3.5 Discussion	39
4. Operant conditioning of pain: a failed replication	43
4.1 Summary	43
4.2 Introduction	43
4.3 Method	46
4.4 Results	48
4.5 Discussion	51
5. Psychometric properties of the Multidimensional Pain Inventory - Dutch Language Version (MPI-DLV)	55
5.1 Summary	55
5.2 Introduction	55
5.3 Method	57
5.4 Results	60
5.5 Discussion	68

6.	MPI-DLV pain patient clusters: an attempt at replication	73
6.1	Summary	73
6.2	Introduction	73
6.3	Method	75
6.4	Results	76
6.5	Discussion	85
7.	Validating the MPI-DLV using experience sampling data	87
7.1	Summary	87
7.2	Introduction	87
7.3	Method	89
7.4	Results	92
7.5	Discussion	94
8.	Profile characteristics of the MPI-DLV clusters	97
8.1	Summary	97
8.2	Introduction	97
8.3	Method	99
8.4	Results	99
8.5	Discussion	103
9.	General discussion	105
9.1	MPI-DLV development and patient sub-types	105
9.2	MPI-DLV development and patient sub-groups: evaluative comments	106
9.3	Operant conditioning of pain: evaluative comments and post-hoc analyses	109
9.4	MPI-DLV patient sub-groups and pain mechanisms: an integration	114
9.5	Hypothetical underlying mechanisms for MPI-DLV patient clusters	116
9.6	General conclusions	119
	References	121
	Summary	133
	Samenvatting	139
	Dankwoord	145
	Curriculum Vitae	147

CHAPTER 1

GENERAL INTRODUCTION

1.1 An overview

Many dissertations on chronic pain start with a brief summary of the economic, social and epidemiological aspects of the subject. This is not surprising when one considers the severity and magnitude of the problem. It has been estimated that, in the United States, more than 15 million people suffer from chronic low back pain, and that 93 million working days are lost per year as a consequence (Aronoff et al., 1983). In addition, reports give the number of migraine patients as 25 million (Paulley and Haskell, 1975). In total, chronic pain is estimated to cost the country 40 billion dollars a year (Aronoff et al., 1983). The situation is no better in Western Europe. In the United Kingdom, for example, the annual cost to the country due to backache is about 1 billion pounds (Waddell, 1982).

Although precise data is lacking for the Netherlands, it is reported that 10.000 new cases (patients who are unable to work because of pain) are registered each year (Ministerie WVC, 1991). All these patients receive a disability allowance. The duration of pain in this patient group is at least 54 weeks; longer than might be expected, given the type of diagnosis made. Another source gives a considerably higher estimate: Foekens (1992) reports 33.445 new cases in 1991. All these patients were assigned to ICD category XIII, characterized by muscular, skeletal or connective tissue pain complaints. Foekens further mentions that 18.295 people who are suffering from back pain receive a disability allowance.

Another aspect of the chronic pain problem is that relating to treatment. Despite our large and rapidly increasing knowledge we (whether scientists, physicians, or health care providers) are still unable to provide optimum treatment for chronic pain. Two possible ways of tackling this problem are described here.

The first one originates in the already long-held view that chronic pain is a specific syndrome (Black, 1975), leading to the point of view that all patients suffering from 'the chronic pain syndrome' can be treated in the same manner. Frequent observations in clinical practice, however, strongly indicate that this is not so. The following is perhaps the most convincing illustration of this. Two chronic pain patients with exactly the same (medical) diagnosis are treated by means of an identical intervention (medication, a comprehensive treatment program supplemented by medication, etc.). At the end of treatment it is quite possible that the first patient will report a clear alleviation of pain and suffering, whereas the second patient might report no lessening of pain or improvement in his situation whatsoever. In the light of this empirical observation and other facts which serve to refute the uniformity proposition, investigators have come to question its validity

(Fordyce, 1976; Kiesler, 1971). Indeed, it has been argued that treatment should be matched to the patient's psychological characteristics (Turk, 1990). In recent years, several attempts have been made to identify homogeneous patient sub-groups, based on specific patient characteristics (e.g., Bradley et al., 1978; Jamison et al., 1988; Keefe et al., 1990). It is assumed that treating each sub-group with a specific treatment program (suited to the characteristics of that sub-group) will produce better treatment results. In recent years, an American research team led by Dennis Turk, published a number of articles on a) the rationale of creating pain patient sub-groups and b) the empirical and statistical development of pain patient sub-groups (for an overview see Turk, 1990). Compared to other attempts at developing patient sub-groups, Turk's system has some clear practical, methodological and theoretical advantages. These will be reported upon later.

A second way of tackling the issue of non-optimal treatment results is to acquire greater insight into the etiological and/or maintaining mechanisms of chronic pain. Finding the answer to such questions as: "Why do patients behave as they do?", or "What kind of biobehavioral mechanisms (Feuerstein et al., 1987) determine chronic pain patient behavior?", would lead to a refinement of theory and thus a better understanding of chronic pain. A greater knowledge of etiological and/or maintaining mechanisms would provide treatment programs with a more scientific foundation, and would therefore lead to better treatment results.

This Chapter presents an overview of the studies conducted in the areas of research described above, and concludes with the formulation of the research questions for the present thesis.

1.2 The multidimensional nature of pain

Assisted by rapid progress in the field of medical technology, a unidimensional, physiologically based theory of pain - the specificity theory - dominated the first decades of this century. In this theory it is assumed that the experience of pain is solely a function of the amount of nociceptive input and extent of tissue damage. It was postulated that pain was a specific sensation and that it was communicated to the brain via an entirely independent sensing system devoted to pain, an idea which had already been put forward by Descartes (1664). Accompanying negative emotions were viewed as a consequence of pain, a reaction to the primary pain sensation (cf. Feuerstein et al., 1986).

It is beyond question that research based on this theory has produced useful results for the diagnosis and treatment of acute pain. However, the tenability of this theory, especially in the case of chronic pain, has seemed questionable. It is primarily the contradictory, empirical facts that have greatly challenged the theory. In many - even acute - cases there is no direct relationship between tissue damage and the experience of pain (Wall, 1979). Beecher's study (1956) is often cited in this context. Beecher described the expressions of pain of wounded soldiers at the notorious Anzio Front (a battle at Monte Casino during the Second World War).

Soldiers evacuated from the front line rarely asked for pain medication. This is in contrast to, for example, victims of road traffic accidents with comparable (severe) wounds. Beecher concluded from this that emotions played an important role in the experience of pain.

A second empirical argument directed against purely physiological formulations can be derived from medical treatment. Patients frequently continue to report pain - or the elimination of pain is of only short duration - after neurosurgical interventions by which assumed pain pathways are blocked, preventing the conduction of nociceptive stimulation (Loeser, 1980). Hypnotic analgesia is another example which serves to show that nociception and the experience of pain can be unrelated. According to Hilgard (1975) cognitive mechanisms play an important role in hypnosis. It is hypothesized that, due to an altered subjective awareness of, and reaction to pain, in hypnotic analgesia nociceptive information is registered, but is partially blocked from consciousness. However, the exact mechanism by which hypnosis works remains a matter of some debate (cf. Turner and Chapman, 1982). In addition to these empirical, anecdotal examples, there is also experimental evidence that pain is influenced by psychological factors, such as attention (Arntz, 1991).

The foregoing examples focus mainly on acute pain. It is even more difficult to explain chronic pain problems in terms of the specificity theory. A large discrepancy between pain complaints and 'objective' organic findings is characteristic of the so-called chronic 'benign' pain states. This pain behavior cannot be explained exclusively by underlying somatic pathology. Therefore, treating this type of pain with purely medically oriented regimens seems doomed to failure in many cases (Fordyce et al., 1986).

Taken together, these clinical observations and experimental findings suggest that there is no one-to-one relationship between tissue damage/nociceptive stimulation and pain sensation, and that psychological factors should be taken into account in pain theories. An adequate conceptualization of pain must be multidimensional in nature (cf. Turk et al., 1983)¹.

Pioneering theoretical work was carried out by Melzack and Wall, resulting in the publication of their 'gate control theory' (Melzack and Wall, 1965). In short, this model proposes that there is a spinal gate mechanism in the dorsal horn region of the spinal cord which modulates the transmission of nerve impulses from peripheral fibers to the central nervous system. It is further proposed that this gate

¹This statement, and especially the word multidimensional, needs to be interpreted carefully. A phenomenon is often called multidimensional if it is not completely understood in terms of a unidimensional theory. In actual fact, in these cases the word multidimensional can be considered as a sign of weakness because it is not supported by experimental evidence. In the present case the word multidimensional means that there is experimental evidence that a phenomenon (pain) is influenced / modulated via independent pathways.

mechanism can be influenced by descending nerve impulses (representing psychological processes) from the brain. Criticism of this theory, primarily directed at the presumed physiological and anatomical bases of the theory (Liebeskind and Paul, 1977; Nathan, 1976; Weisenberg, 1977) has resulted in a revised model (Melzack and Wall, 1982). However, the most essential implication of the theory, the multidimensionality of pain, remains unaltered. This was further illustrated by the publication of the McGill Pain Questionnaire (Melzack, 1975). This list, made up of 20 subclasses each consisting of a group of words (pain adjectives), can be used to describe three different qualities of the pain experience. The sensory-discriminative quality refers to sensory qualities of the pain experience in terms of temporal, spatial, thermal and other properties. The second quality is called motivational-affective and describes affective qualities that are part of the pain experience, such as tension and fear. The third quality - cognitive-evaluative - measures the subjective overall intensity of the total pain experience. The McGill Pain Questionnaire was translated into many languages and the three major categories of pain experience could be confirmed in several other studies (Reading, 1984).

Nowadays, the multidimensional view of the phenomenon of pain is generally accepted, as is illustrated by the definition offered by the International Association for the Study of Pain (1979): "An unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage". This position is supported by a large body of literature in which it is argued that in addition to somatic variables (e.g., Frederickson and Chipkin, 1988; Flor and Turk, 1984), psychological factors (for overviews see, Pennebaker, 1982; Sternbach, 1978; Summers et al., 1991), behavioral aspects (e.g., Fordyce, 1974; Groenman et al., 1990; Keefe, 1989; Vlaeyen, 1991), personality factors (e.g., Skevington, 1983; Wade et al., 1992) and socio-cultural factors (e.g., Bates et al., 1993; Lipton and Marbach, 1984; Weisenberg, 1975; Zola, 1966) are also involved in pain.

1.3 Psychological models and treatment of chronic pain

Taking this multidimensional view as their basis, researchers developed psychological models to explain the etiology and/or maintaining mechanisms of chronic pain. Three of these, the operant model, a model marked by antecedent control of pain and the cognitive-behavioral approach will be considered here².

²In addition to these three psychological models, articles have been published showing the importance in pain problems of observational learning (Bandura, 1977b; Craig, 1978, 1988) and family systems theory (Turk et al., 1987b; Flor et al., 1987b). For a more extensive and detailed overview, albeit limited to chronic lower back pain, see Feuerstein et al., 1987.

The operant model

The application of the operant learning theory (Skinner, 1953) to pain problems represented a landmark in the history of pain-theory development (Fordyce, 1976). The introduction of the 'pain behavior' concept (all observable communications of pain and suffering) made it possible to employ the laws of operant learning theory. More specifically, it was hypothesized that patients' pain behavior could be controlled operantly by (in many cases environmental) consequences. Acute pain problems may come under the control of environmental contingencies and consequently lead to a chronic pain problem. Thus, chronic pain behavior can be seen as learned behavior. Maintenance of pain behavior can take place in three ways.

Firstly, by direct positive reinforcement of pain behavior. In order to avoid any possible misinterpretations, 'positive reinforcement' will be defined as follows: Any stimulus is considered to be a reinforcer if it increases the probability of a response. Positive reinforcement refers to a situation where the **presentation** of a stimulus following an operant response, strengthens the probability of that response (Skinner, 1953, p.73). Attention from the spouse, analgesic medication intake with subsequent pain reduction, etc., are positive reinforcers in pain problems.

The second way in which pain behavior can be maintained is by means of negative reinforcement of that behavior. According to Skinner (1953), negative reinforcement refers to a situation where the **removal** of an aversive stimulus following an operant response, strengthens the probability of that response. The non-occurrence of a stressful and conflictous working situation is a potential, negative reinforcer of pain behavior.

The third way relates to a lack of reinforcement of 'sound' or 'well' behavior such as activity, work, etc. This situation may occur when a solicitous spouse or other family member takes over patient's activities and encourages the patient to rest (cf. Groenman et al., 1988).

It is thought that these conditioning principles can result in the continuance of pain behavior after a wound has healed, and in the absence of somatic pathology. Moreover, it is believed that operant conditioning can take place without the individual's being aware of it (Chapman and Turner, 1990). A recent study by Conally and Sanders (1991) lends convincing support to the important role pain behavior plays in chronic pain problems. They found that the level of pretreatment overt pain behavior was associated with poorer subsequent treatment outcome (lumbar sympathetic nerve blocks and interdisciplinary rehabilitation). The predictive capability of pain behavior was greater than pretreatment variables such as duration of pain, number of surgical interventions or receipt of financial compensation.

The therapeutical implications of the operant view are straightforward and are based on the principles of 'shaping' and 'extinction'. Attempts are made to reduce pain behavior by the withdrawal of its reinforcers. An increase of 'well' behavior (in most instances physical activity) is established by systematically giving positive reinforcement of this target behavior. Attention and other reinforcers are no longer

contingent to pain behavior but are contingent to sound/well behavior. The active participation of the patient's partner in the treatment program is usually required. Because operant treatment programs often take place in well-controlled clinical settings, attention is also given to the maintenance of well behavior in the patient's natural environment. Reviewing the literature, Linton (1986) comes to the conclusion that operant treatment programs are effective in increasing activity levels and in decreasing the consumption of medication. Whether these treatment programs also lead to a decrease in reported pain levels is less clear.

Despite the therapeutic successes, the operant paradigm and treatment of chronic pain has not been without its critics. The first point of criticism is a theoretical one. Promising treatment results from operant pain programs do not necessarily mean that therapeutically reduced pain behavior has, in fact, resulted from operant conditioning; therapeutic effects do not prove causation. In other words, "Support from clinical studies is based on a dangerous type of logic: if treatment works, then the hypothesis must be correct. Obviously, this need not be so" (Linton and Götestam, 1985; see also Linton, 1982). The question as to what exactly is meant by 'pain behavior' forms another point of concern (see the discussion between Schmidt and Fordyce: Fordyce et al., 1985; Fordyce et al., 1988; Schmidt, 1987; Schmidt, 1988). One aspect of their discussion focuses on the significance of a visual analogue score. Is this a form of pain behavior or not? To quote Fordyce et al. (1987): "One may know of a subject's pain by observing his or her facial expressions, posture, vocalizations, peripheral autonomic responses, etc. These are pain 'behaviors'. Even the most mentalistic observer must ask the subject to rate or categorize or describe the pain in some way; these ratings or descriptions are also behaviors". Schmidt (1988), on the other hand, argues that "pain ratings are neither pain behaviors nor nociception but - as an operationalization of pain perception - have a distinct status on the pain dimension between nociception and pain behavior". Thirdly, operant treatment programs primarily focus on pain **behavior** and not on the pain **experience** since the latter is seen as a private, unobservable event and is thus inappropriate as a direct focus of treatment. It is assumed that changing a patient's pain behavior will have a positive effect on the pain experience. Some outcome studies do report a significant reduction in the subjective pain intensity level (e.g., Fordyce et al., 1973; Ignelzi et al., 1977; Sternbach, 1974) but in other studies no significant decrease effect was found (e.g., Kerns et al., 1986). Swanson et al. (1976a, 1976b) found only a 14% decrease in pain ratings. After finishing an operant treatment program patients often say that they can perform more activities and that their quality of life has improved. However, they also report that the intensity of pain has remained unaltered.

A final criticism concerns research technique. In most instances admission to operant treatment programs depends on rigid selection criteria. Two of the requirements, for example, are the presence of observable motor deficits and the participation of the spouse (e.g., Vlaeyen, 1991). The chances of treatment success are thus enhanced.

Antecedent control of pain

In the previous paragraph the maintenance of sustained pain behavior was explained in terms of consequent controlling factors. Although it is assumed that operant factors are much more important in chronic pain states than in acute pain, chronic pain behavior may also be controlled by antecedent factors. This type of pain is sometimes called respondent pain (Chapman and Turner, 1990). The term 'respondent' is not always used consistently³. Feuerstein et al. (1987), for example, clearly defined respondent pain behavior as behavior which is "directly related to the presence of a nociceptive stimulus". According to Gentry and Bernal (1977) the respondent model is an approach in which "pain is viewed both as a response to and an antecedent of physical tension, specifically muscular tension. That is, whenever there is an organic insult (injury) to the body, causing pain, the resulting physiological response may be one of immediate tensing of the muscles surrounding the injured area. Such tensing may in turn increase the subjective experience of pain, which leads to increased tension in the neighboring muscle groups as a means of further immobilizing the injured site. In short, what is developed is a pain-tension-pain cycle" (see also Peters, 1992). It is further argued that immobilization, as a means of pain-reduction, may contribute to muscular atrophy and increasing disability (Turk and Flor, 1984). It is assumed that classical conditioning (Pavlov, 1927) plays an important role in the development of the pain-tension-pain cycle. Increases in muscle tension and sympathetic activation may become conditioned responses and maintain a pain-tension cycle (Flor et al., 1990a). The role that anticipatory anxiety and avoidance of activity plays in this process was elaborated by Philips (1987).

Certain clinical observations seem to suggest that the very experience of pain is directly conditionable in a classical way. For example, a dentist may observe that a frightened patient reports pain before he starts drilling (see Arntz, 1991); i.e., classical conditioning of pain takes place if, after conditioning, a previously neutral stimulus is experienced as painful because it is associated with an unconditioned stimulus (a pain stimulus). Over time, by means of stimulus generalization, a large variety of stimuli which were previously associated with nociceptive stimulation may elicit pain. To take an example, a staircase where someone was once injured may become a conditioned stimulus for pain perception or behavior (see Sanders, 1985). In the present author's opinion, however, the reaction of immediately feeling pain upon seeing a staircase is very unlikely⁴. Moreover, a direct attempt to condition pain experience in a classical conditioning experiment failed (Meijboom, 1992). In fact, the opposite effect (analgesia) was found. After conditioning

³This is why this paragraph is not entitled 'the respondent model of pain'.

⁴On the other hand, it is more likely that the sight of the staircase will lead to an increase in anxiety. Thus, it is not the experience of pain, but the anxiety which is conditioned classically, and this is completely in line with Mowrer's statement that fear is the conditioned form of the pain reaction (Mowrer, 1939).

with a painful electric shock, subjects reported less pain in the extinction phase than did subjects in the control group.

Despite the disputed theoretical foundation of the respondent model, the effects of therapies based on it - relaxation techniques (Bernstein and Borcovic, 1973; Jacobson, 1938) and biofeedback - have proved promising (Blanchard et al., 1982; Cott et al., 1992; Murphy et al., 1989; Turner and Chapman, 1981).

The cognitive-behavioral perspective

Cognitive-behavioral concepts of pain are closely linked to the gate control theory. The basis of this theoretical perspective is that emotional and behavioral reactions to situations are determined by the cognitions people maintain in those situations (Goldfried, 1977). Maladaptive cognitive processes such as 'catastrophizing' may lead to or exacerbate feelings of helplessness and hopelessness. In a general sense, a patient's physical, psychosocial and occupational functioning can be affected by negative cognitions. These notions are supported by (experimental) studies (e.g., Demjen and Bakal, 1986; Philips, 1987; Schmidt, 1985). In attempts to predict pain and disability in chronic back pain and rheumatoid arthritis patients, analyses revealed that cognitive variables explain significantly more variance (44% to 54%) than disease-related variables (12% to 28%) (Flor and Turk, 1988).

The aim of the cognitive-behavioral approach to pain is to change both the patient's view of his pain and his habitual maladaptive ways of coping with it (Turk et al., 1983). In the cognitive-behavioral treatment of pain, attempts are made to increase the patients' control over the pain experience and over their lives. This is accomplished by giving the patients information on the multidimensional nature of pain, by setting goals, by challenging ideas of helpless passivity, by re-conceptualization, by the teaching of various coping skills (relaxation and attention diversion), and by increasing activity. The efficacy of cognitive-behavioral treatment strategies has been demonstrated in several controlled studies (see Keefe, 1992). Spence (1989), for example, found significant effects on variables like anxiety, depression, coping strategies, impact on daily life, pain and distress caused by pain. What is more, these effects were found to have been maintained at 6 months' follow-up.

Cognitive-behavioral treatment has worked successfully in conjunction with biofeedback and other treatment techniques (Tan, 1982), and research has been carried out to compare the differential efficacy of cognitive-behavioral treatment versus operant-behavioral treatment. In a controlled study Turner and Clancy (1988) compared the operant and cognitive-behavioral treatment of chronic low back pain patients. It was found that post-treatment gains (decreased physical and psychosocial disability) had been maintained at a one year follow-up in both groups. However, the type of improvement in the two treatment conditions was different. Operantly treated patients showed a greater initial improvement, whereas the cognitive-behaviorally treated group showed steady improvement over the one-year follow-up period.

In each of the treatment interventions described above, the emphasis was placed on specific pain-related aspects: behavior, muscle tension, cognitions. During the past few decades another treatment strategy has become increasingly popular; the 'eclectic' or 'comprehensive' approach, practised in 'multidisciplinary pain centers'. This kind of treatment comprises a wide range of treatment techniques, stemming from various theoretical perspectives. Indeed, the ingredients of comprehensive treatment programs differ from study to study. The justification for the comprehensive approach is that some components will be effective for some patients, whereas other patients will probably benefit from other combinations. The programs mostly consist of a combination of physical therapy, relaxation and biofeedback, psychotherapy, educational therapy, and spouse participation. They are also frequently performed in an operant and/or cognitive-behavioral sphere.

Recently, Flor and associates (1992) conducted a meta-analytical review which investigated the efficacy of multidisciplinary pain treatment centers. Based on the data reported in sixty-five multidisciplinary treatment studies, and making use of 'meta-analysis' statistical techniques (Glass et al., 1981), it was found that multidisciplinary treatment methods were superior to no treatment or to waiting-list controls. One very interesting finding was that multidisciplinary treatment was superior to single-discipline, unidimensional treatment, such as medical treatment or physical therapy.

Until recently, the lack of a clear, sharp multidimensional model has been a major point of concern with regard to the eclectic strategy. In the past few years some attempts have been made to integrate notions of behavioral, psychological and psychobiological perspectives of pain, resulting in a multidimensional model (Feuerstein et al., 1987; Flor et al., 1990a).

Some unidimensional and multidimensional models of pain have been presented above. Treatment approaches based on these models have proved to be effective in reducing pain behavior, psychological distress, etc. Yet these positive effects have to be considered in context. Treatment effects are almost always computed on averaged data on a (large) sample of patients. This aggregate effect does not provide insight into how this effect is made up. As was mentioned in section 1.1, within a 'successfully' treated group, there is always a certain number of patients who are less responsive or who do not respond at all. For example, in a study by Vlaeyen (1991), of the nineteen patients who had followed an inpatient operant treatment program, three (16%) were classified as 'poor persisters' at six months' follow-up. These were patients who met fewer than 4 of the 11 'target' variables. A somewhat lower number of patients (2 out of 18: 11%), who had followed an operant-cognitive treatment program, were also scored as 'poor persisters'. This is known as 'the phenomenon of non-optimal treatment results' and, whether unidimensional, multidimensional or eclectic in nature, all treatment strategies are acquainted with it.

As was stated at the beginning of this Chapter the scientific problem of non-optimal treatment results can be attacked in different ways. One strategy would be

to gain more insight into the basic mechanisms involved in the development and maintenance of pain. Let us take the example of a researcher wishing to investigate the phenomenon of non-optimal treatment results in operant treatment programs. The theoretical basis of an operant treatment program is that (pain) behavior can be influenced by means of extinction and shaping⁵. As was discussed previously, operant treatment programs succeed fairly well in doing this, but the results are not optimal. One of the researcher's objectives would therefore be to increase his knowledge of the developmental mechanism of chronic pain behavior. Although both the manner in which pain behavior in an acute state can be influenced by operant factors and the way in which it can be maintained are extensively documented (Fordyce, 1976), only a few experimental studies have been published that support the notion that pain behavior can be conditioned operantly (e.g., Block et al., 1980; Linton and Götestam, 1985). Indeed, Turk and Flor (1987) had some critical observations to make on this point: "Pain behaviors are multiply determined (physical, psychological, as well as environmental factors) and one cannot assume that they are the result exclusively of learned patterns following positive or negative reinforcement". The researcher might therefore decide that the issue of the operant conditioning of pain behavior requires further study. The knowledge acquired from these studies may lead to some adjustments in the treatment program, which might ultimately lead to better treatment results.

Another way of tackling non-optimal treatment results relates to patient sub-groups. There are three lines of research in this field. The first is characterized by an attempt to predict treatment success (which patient sub-group will be successfully treated and which will not?). The second deals with the identification of patient sub-groups based on a theoretical model, and the third is marked by the identification of sub-groups using empirical methods. The last two will be discussed in some detail.

1.4 Non-optimal treatment results and identification of patient sub-groups

Predictors of treatment success

Knowing that some patients benefit from a treatment modality whilst others do not, or do so only marginally⁶, researchers have tried to identify variables which predict success or failure. In most cases multivariate (logistic regression) techniques are applied. In many cases treatment effect is assessed by means of a 'dragnet procedure'. That is, treatment effect is computed on the basis of the changes in a large number of variables where an effect might have been expected. It should be noted that these kinds of study have a retrospective orientation, in that it is

⁵This does not necessarily mean that pain behavior is learned in an operant way!

⁶Throughout this dissertation the term 'marginal' is used to label P-values between 0.05 and 0.10.

determined, post treatment, whether patients are typical cases of success or failure. Investigations are then carried out to see if pre- or begin-treatment variables, which can adequately predict success or failure can be determined. Kores et al. (1990), for example, found that higher pre-treatment self-efficacy scores were associated with a higher improvement level, a better overall functioning and greater reductions in chronic illness behavior at 7 months' follow-up. Härkäpää et al., (1991) investigated the role of health locus of control beliefs and psychological distress as predictors of treatment outcome in chronic low back pain patients. They reported evidence that internal locus of control and treatment success are positively associated. In addition, psychological distress was significantly associated with poorer accomplishment of back exercises. The negative effects of psychological distress on treatment outcome are also reported by other researchers (Lee et al., 1989; Polatin et al., 1989). On the other hand, Brennan et al., (1986) and Watkins et al., (1986) failed to find such a relationship. The negative effect of psychological distress in the treatment of pain patients is further questioned by King and Snow (1989) and Kleinke and Spangler (1988). In these studies it was concluded that patients with higher levels of psychopathology, as measured using the MMPI, are less likely to drop out of treatment than those who report a lower level of psychopathology.

In sum, variables such as self-efficacy and health locus of control seem to predict favorable treatment outcome. The role of psychological distress as a predictive variable is less clear. A logical sequel to the 'predictor of treatment success' research strategy is to use the information obtained for the creation of selection criteria and/or in the adjustment of treatment components. To the author's knowledge, however, no study has been carried out in the field of chronic pain to test the value of predictive variables in a prospective, hypothesis testing, design.

Deductively and empirically derived patient sub-groups

The idea that chronic pain patients do not form a homogeneous group, particularly in somatic terms, is not new. It has long been realized that somatically different types of pain (e.g., disk herniation, migraine, fibromyalgia, etc.) require different types of treatment, each placing the emphasis on the specific somatic aspects. Besides this somatic subdivision, the existence of sub-groups based on personality, psychosocial and behavioral aspects is frequently claimed in current literature on pain.

There are, roughly, two ways to create a subdivision, classification or taxonomy. The first is 'deductive'. Based on specific theoretical concepts, assumptions are made about the existence of sub-groups and mechanisms underlying each sub-group. Two examples of this deductive method are the classification system published by the International Association for the Study of Pain IASP (Merskey, 1986) and the three-factor model of pain (Vlaeyen et al., 1989b).

The design of the classification system, as proposed by the IASP, is based on contemporary somato-medical knowledge. The pain problem is classified according to a specific combination of judgments on five axes: (1) body region, (2) system

involved, (3) temporal characteristics of pain and patterns of occurrence, (4) intensity and duration, and (5) etiology. For example, severe tension headache is coded 9033.97: body region: head, face and mouth; system involved: musculo-skeletal system and connective tissue; temporal characteristics of pain: continuous or almost continuous, fluctuating severity; intensity and duration of pain: more than 6 months; etiology: psychophysiological. The IASP system seems to work reasonably well as far as the classification of medical information is concerned. However, if this medical-somatically oriented classification system is considered from a multidimensional point of view, the system has its shortcomings; e.g., it is not possible to include psychosocial and behavioral pain-related aspects⁷.

The three-factor model of pain (Vlaeyen et al., 1989b) states that pain should be considered as an emotion. Analogous to the three-system model of fear (Hugdahl, 1981), three pain response systems are distinguished: overt-motor, verbal-cognitive and physiological. It is hypothesized that operant conditioning processes influence the overt-motor system. Similarly, respondent conditioning is considered to be an underlying mechanism for the physiological response system. With respect to the verbal-cognitive response system, attribution theory (Weiner and Graham, 1984) and Bandura's self-efficacy theory (1977a) are presumed to play an important role here. Suijker (1989) developed a questionnaire based on the three-system theoretical model, to assess the relative importance of the three response systems. Confirmatory factor analysis, however, failed to reveal the three response factors. In discussing the lack of confirmation from these results, it was suggested that the measurement of overt behavior and tension level by means of a self-report questionnaire is doubtful. An alternative and perhaps better way of assessing the role these three response systems play in individual patients, is to use different measurement techniques. The pain behavior scale (Vlaeyen et al., 1990b) or a pain behavior observation method (Keefe and Block, 1982) could be applied to obtain quantitative results on the overt response system. Questionnaires measuring pain cognitions (e.g., Flor et al., 1993; Philips, 1989; Vlaeyen et al., 1990a) could be used to map out the verbal cognitive system. In order to get more information about the physiological response system, EMG measurements, for example, could be used. After the assessment phase, a profile could be drawn up, in which the 'contribution' of each response system was reported. In other words, a three dimensional pain space is evolved (cf. Duncan et al., 1978). The three axes represent the three pain systems. The vector 'P' points to a certain place in the three dimensional space. Vector 'P' may be viewed as the geometric representation of the assessment results. A patient with a high score on the overt dimension (thought to be influenced by operant factors) is held to benefit most from a

⁷In this context it should be mentioned that there is another, more specific classification system, comparable to that of the IASP. Olesen (1988) proposed a system for classifying headache disorders. There are twelve categories, supplemented by a non-classifiable one. This system can, however, also be criticized using the same arguments as apply to the IASP classification system.

behavioral modification program based on such operant learning principles as shaping and extinction. A treatment module emphasizing a change in distorted pain cognitions is indicated in patients scoring high on the verbal-cognitive dimension. In patients showing high habitual tension and/or high stress reactivity levels, the pain-tension cycle can be interrupted using relaxation techniques (e.g., Bernstein and Borkovic, 1973; Jacobson, 1938).

In contrast to the IASP taxonomy, the three-system model represents a strongly heuristic model for pain research. Despite this positive evaluation of the three-system approach, there are some concerns. Firstly, it is not clear how knowledge about somato-medical findings or some of the psycho-social factors (e.g., life events, Craufurd et al., 1990) should be placed within the model. Secondly, having determined the importance of each of the three dimensional pain space vectors, it is difficult to develop strict criteria for matching 'vector importance' to appropriate treatment strategy.

The second way of creating sub-groups is a more or less empirical one. In the absence of clearly defined specific patient sub-groups, cluster-analysis techniques are used to create homogeneous ones. In order to do this the average between-group distance is maximized, and the average within-group distance is minimized. Patients are thus grouped in such a way that those within a group share as many characteristics as possible, whilst at the same time differing as much as possible from those in other groups. The choice of variables, patient characteristics, used to create the pain patient sub-groups depends upon the researcher's theoretical background. By far the largest amount of research into the identification of patient sub-groups has centered on personality, making use of the Minnesota Multiphasic Personality Inventory (MMPI). The worldwide use of this instrument and the long research tradition in this area warrant detailed discussion here.

Early MMPI personality research contributed to the concept of pain patient homogeneity. It was frequently found that a group of pain patients' mean MMPI profile revealed a specific pattern: the so-called "conversion-V". This V-shaped profile consisted of elevated scores on the first and third MMPI scale (the Hypochondriasis and Hysteria scale respectively) and a relatively low score on scale two, the Depression scale. The other MMPI scales showed mean scores. In short, this profile suggested a tendency to react to psychological distress with physical symptoms. It was usually assumed that psychological problems were denied and repressed. This MMPI 'conversion-V' profile was reported worldwide (for an overview see Love and Peck, 1987) and the finding led to the supposition that all chronic pain patients were characterized by similar psychological functioning. Motivated, however, by the contradictory clinical findings of Sternbach (1974), Bradley and co-workers (1978) did some pioneering work, testing the variability within the mean conversion-V profile. The results of a multivariate hierarchical cluster-analysis clearly showed that the conversion-V pattern comprises different, distinct sub-profiles. Four female clusters and three different male

clusters were found. These findings were replicated in three cohorts (patients seen in three consecutive years) and it was concluded that this sub-grouping might have important implications for practitioners' choice of treatment modalities. A number of researchers have replicated these findings, the only difference being that in most cases four, instead of three, male clusters were found (e.g., McCreary, 1985; McGill et al., 1983; Schipper, 1992). Costello et al. (1987) carried out a MMPI meta-clustering by combining the results of 10 investigative teams. Four MMPI types emerged which were originally labelled as P-A-I-N. In the 'P' MMPI profile there were elevated scores on almost all the scales, hence the designation 'psychopathological'. Type 'A' was characterized by the conversion-V. In the 'I' profile all three scales of the neurotic triad were elevated. Type 'I' patients appeared to benefit from psychological treatment. And finally, since none of the scales in the 'N' profile showed elevated scores, this personality profile was designated 'normal'. The authors also proposed simple classification rules (without the use a computer) for assigning individual patients to one of the four P-A-I-N cluster types.

Research using the MMPI as an instrument for classifying chronic pain patients has not gone uncriticized. Reviewing the P-A-I-N classification system, Robinson et al. (1989) concluded that Costello's proposed classification rules were too restrictive. They failed to classify the MMPI profile of 87 out of 125 patients. Similar observations were reported by Kole-Snijders (1988). On the other hand, a pilot study by Lousberg (1989) showed that small adjustments⁸ to Costello's criteria increased the number of classifiable MMPI profiles from 25% to 72%. Another critical remark concerning the use of the MMPI in pain research relates to the problems of interpreting elevations in clinical scales. For example, a conversion-V profile is interpreted as being a cause of the chronic pain problem. The tendency to react to psychological distress with physical complaints is assumed to be a personality trait which is already present **before** the development of the pain problem. However, Roberts and Reinhart (1980) reported evidence that the conversion-V is a **response** to chronic pain rather than a cause. Finally, there is a crucial point of concern in connection with the predictive validity of the MMPI sub-groups. Up to now, no convincing evidence has been reported that there is a relationship between MMPI sub-group profiles and treatment outcome (Guck et al., 1988; Moore et al., 1986).

In recent years attempts have been made to develop empirically derived pain patient sub-groups using other instruments, such as the SCL-90 (Butterworth and Deardorff, 1987; Groenman et al., 1993; Jamison et al., 1988; Schwartz and Degroot, 1983). Based on data acquired from a standardized observation protocol assessing pain behavior in low back pain patients, Keefe et al. (1990) identified four patient sub-groups which could be replicated in a second independent sample.

⁸The clinical scales' criteria have to be extended to a range of 4 T-points (2 above and 2 below the criterium score). Also, MMPI profiles with a K-score of up to 75 have to be considered as valid.

In principle, then, all kinds of questionnaires or other instruments can be used to create sub-groups. The value of clustering results, however, depends largely on the qualities of the instrument being used. One of the primary requisites is that the instrument should have sound psychometric properties. In addition, an adequate norm group, in this case chronic pain patients, is essential. If reasonable sub-groups are established, the clinical usefulness has to be demonstrated by assessing the predictive qualities. The MMPI and SCL-90 do not meet these requirements sufficiently. The Keefe et al. (1990) study is a step forward since the psychometric properties of the assessment instrument are good and the sub-groups have been replicated. However, the predictive validity of this classification system remains to be proved.

1.5 Empirically based pain-patient sub-groups based on the 'MAP'-taxonomy

As is argued by Turk (1990) a common shortcoming in all previously mentioned empirically based classification attempts is the fact that in each case only one single pain related factor is focused on. In other words, all these studies suffer from a unidimensional approach to pain (personality, general complaining behavior, pain behavior). Stressing the multidimensional nature of pain, Turk and Rudy (1986) proposed a Multiaxial Assessment of Pain (MAP) approach. This assessment approach aims at an integration of the information obtained from three axes: 1) a physical-medical axis, 2) a psychosocial axis and 3) a behavioral axis. The emphasis was on obtaining quantifiable and objective data on each axis. For MAP-axis one, Rudy et al. (1990) developed a standardized Medical Examination and Diagnostic Information Coding System (MEDICS), resulting in a 'pathology index'. The (West-Haven Yale) Multidimensional Pain Inventory (WHY)MPI was developed (Kerns et al. (1985)) to measure pain-relevant information from MAP-axes two and three. The MPI is divided into three parts. MPI-part one measures pain-relevant psychosocial information, and in MPI-parts two and three pain-relevant behavioral information is assessed. The psychometric properties of this questionnaire are good, and sensitivity to treatment changes has also been demonstrated. Applying the K-means cluster-analysis technique, Turk and Rudy (1988) reported three distinct pain patient cluster types, based on the nine MPI scales. The first patient type, labelled as 'Dysfunctional', is characterized by high scores on subjectively perceived pain intensity level, interference of the pain with the daily life, psychological distress, solicitous responses from the partner and environmental support. The level of life control is low. In addition, the Dysfunctional patient type is marked by a relatively low activity level. The second patient type, the 'Interpersonally Distressed' one, reports somewhat lower levels of pain, interference and affective distress, than the Dysfunctional type. These patients feel they gain little support from their environment. They report a high level of punishing responses to pain behavior from their partners. The third patient type, the 'Adaptive Coper',

reports the lowest level of pain, interference and psychological distress. They also seem to be the most active.

Attention has been deserved to rule out potential confounds of this clustering. If the differences between the MPI clusters could have been explained by variables such as age, sex, duration of pain or measures of physical pathology, the MPI classification would add little or no information. However, no difference was found between the MPI clusters with regard to duration of experiencing chronic pain, age or sex, and, although a measure of physical pathology (Rudy et al., 1990) correlated significantly with some MPI scales, the cluster differences remained significant after correction for this effect. 'New' individual patients can be assigned to one of the three MPI clusters on the basis of multivariate generalized squared distances and Bayesian posterior probabilities (Tatsuoka, 1988). In less technical terms, the distance between an individual MPI profile and each of the three cluster means is computed. Intuitively, it would seem obvious that a short distance from a cluster center would represent a typical example of that MPI cluster. A large distance, on the other hand, would indicate that the individual had no or few characteristics in common with that cluster. It is possible to calculate an a-posteriori probability of belonging to a cluster. Since the expected chance value of classifying a profile to one of the three MPI clusters is 0.33, a profile is classified if the posterior probability is at least 0.66 (Rudy, 1989). In general, it can be stated that the shorter the distance to a cluster center, the more likely is the chance that a patient will be assigned to that cluster. This is depicted in Figure 1.1. In this figure the circles represent 99% confidence intervals around the cluster centers. As can be seen, there are also overlapping areas. Patients situated in these areas have characteristics belonging to more than one patient cluster. In the case depicted, it seems likely that case 'X' will be classified to the Dysfunctional cluster.

The accuracy of the MPI classification system has proved to be very high (Turk and Rudy, 1988). Only 5 of a sample of 100 pain patients could not be classified to one of the three MPI clusters. The mean posterior probability was also very high: 0.93, 0.90 and 0.94 for Dysfunctional, Interpersonally Distressed and Adaptive Copers patients respectively.

The next step was to assess the generalizability, robustness, of the MPI classification system (Turk and Rudy, 1990). Three groups were investigated: chronic low back pain patients, headache patients and temporomandibular disorder patients. The results indicated that the MPI-scale co-variance matrices were not significantly different between the three

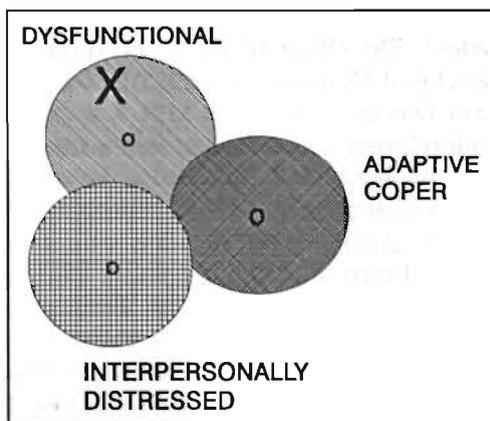


Figure 1.1 Overlapping MPI pain clusters.

groups. In each of the 'somatically' different groups the three MPI clusters were observed, albeit in different proportions (in chronic low back pain patients the proportion of Dysfunctional patients was significantly higher). It was concluded that "psychosocial responses associated with chronic pain are common to diverse samples of pain patients despite differences in demographic characteristics and medical diagnosis" (Turk and Rudy, 1990).

Up to this point, only psychosocial and behavioral pain-relevant aspects (MAP axes 2 and 3), as measured using the MPI, have been integrated. A polydiagnostic approach was then suggested in order to complete the MAP taxonomy with somato-medical diagnostic data (Turk, 1992). The proposition was to combine the results of two classification systems: the MPI taxonomy and the above-mentioned IASP classification system. For example, a patient could be classified as Interpersonally Distressed on the MPI, and severe tension headache (code 033.97, see above) could be the classification on the IASP system. Appropriate multidimensional therapeutic interventions can then be carried out on the basis of these MAP classification results. A second patient may have the same IASP classification, but a different MPI classification, e.g., Dysfunctional. Both patients could be given the same treatment with respect to medical intervention, but, as far as cognitive-behavioral interventions are concerned, it may be advisable to offer the second patient a different treatment package.

1.6 The MAP taxonomy: evaluative comments

A complete MAP taxonomy is made up of the results from two complementary classification systems (those of the IASP and the MPI). The IASP classification system can be considered a deductive classification system. The field may profit more from a deductive-like (e.g., Lakatos and Musgrave, 1974; Popper, 1972) than from an inductive approach (Stove, 1973). A lack of consensus on a deductive classification system regarding psychosocial and behavioral pain-relevant aspects justified the **more or less** empirical MPI taxonomy. 'More or less', because the choice of variables used to create the MPI sub-groups was not a random venture. The cognitive-behavioral theory of pain (Turk et al., 1983) formed the foundation for the construction of the MPI and subsequent patient clustering. Variables such as number of attic-windows in the patient's house were not included in the classification procedure, as might have happened, were the approach purely empirical.

A salient feature of the MPI taxonomy is its descriptive, broadly defining and integrative nature⁹. The distinctiveness of the three MPI profiles facilitates a clear and unambiguous interpretation, and the accuracy with which individual patients

⁹Therefore, the MPI clusters cannot be considered diagnostic categories. As a matter of fact, a diagnostic classification system aims to provide information on etiology, prognosis and treatment.

can be classified as belonging to one of the three empirically derived clusters is high.

This is, then, the current 'state of the art' regarding pain patient sub-group development. In contrast to other attempts at developing empirical sub-groups, the MAP line of research described above and initiated by Turk and colleagues, is considered to be the most promising, because of the good psychometric qualities of the MPI, the short time needed to fill in the questionnaire, and its ability to classify patients. Another argument pertains the importance of an international agreed upon as to the (IASP) medical somatic classification system (to be used in combination with the MPI classification). Finally, the multiaxial assessment approach matches completely the multidimensional nature of pain.

With respect to the therapeutic implications of the MAP system the following can be said: A first logical consequence is to treat chronic pain patients in a multidimensional way. It would seem reasonable to develop three specific cognitive-behaviorally oriented 'comprehensive' treatment modules, based on the three MPI sub-groups. Each module should contain components appropriate to the optimal treatment of the particular MPI sub-group. Incorporation of specific somatic interventions, guided by the IASP classification, with a specific cognitive-behavioral intervention, would complement treatment strategy based on the MAP approach.

1.7 Specification of research questions

This Chapter has focused on the phenomenon of non-optimal treatment results and ways of dealing with it. The present thesis elaborates on two issues discussed earlier. The first concerns the operant conditioning of pain; the second, the MPI patient sub-groups. More specifically, the following issues and research questions are evaluated.

Operant conditioning of pain (Chapters 2-4)

The study by Block et al. (1980) showed that the patient's pain report depended on whether he/she was observed by his/her spouse or ward clerk. The results concurred with operant notions of pain behavior (for a detailed description of this study, see Chapter 2).

In Chapter 2 an experiment is described which is comparable to that of Block et al., (1980). The most important modification is that relating to the dependent variables. That is, pain behavior is not only measured by a visual analogue scale (VAS) (as was done in the Block study) but is also operationalized as poorer endurance during a walking-to-tolerance treadmill test for chronic low back pain patients. The aim of this study is to replicate the VAS effect of the Block study, and also to generalize this effect to a more overt form of pain behavior.

Chapter 3 presents an experiment comparable to that carried out by Linton and Götestam (1985). There were two central points of interest.

- A) The replication of one of their findings; i.e., pain report can be conditioned operantly to a higher level depending on the experimenter's reaction to the test subject's VAS pain report.
- B) Can psychological and psychophysiological variables which are presumed to be correlated with the very experience of pain, also be conditioned in an operant way?

Based on the results of the experiment in Chapter 3, a second experiment was carried out, aimed at replicating the results of the first experiment. This second experiment is discussed in Chapter 4. Investigations were also carried out to determine whether it was possible to condition pain report not only to higher levels (Chapter 3), but also to lower levels.

Pain patient sub-groups (Chapters 5-8)

Inspired by the work of Turk and co-workers, the author decided to translate the Multidimensional Pain Inventory (Kerns et al., 1985) into Dutch and to investigate the psychometric qualities of the MPI-Dutch Language Version (MPI-DLV). Besides being assessed in a 'traditional' concurrent manner, the validity of the General Activity scale was also assessed experimentally (cf. Pearce and Morley, 1989). Based on the data from a bicycle ergometer treadmill test, it was hypothesized that a positive relationship would be found between the MPI-DLV General Activity scale and performance level on the bicycle ergometer test. This study is described in Chapter 5.

The identification of pain patient sub-groups, based on the MPI-DLV, is presented in Chapter 6. The same statistical technique was applied as was used by Turk and Rudy (1988).

The normal method of validity assessment in questionnaire research is to compute correlations with the content of comparable scales from other questionnaires. In addition to this, an experience sampling procedure was carried out to test the ecological validity of the MPI-DLV (Chapter 7).

Chapter 8 describes a study in which additional profile characteristics (analgesic medication intake, intelligence, and personality characteristics) of the MPI-DLV clusters were investigated. This Chapter is partly hypothesis-testing (analgesic intake) and partly explorative (intelligence and personality) in nature.

Chapter 9, finally, presents a general discussion of the findings of Chapters 2-8. In addition, an attempt is made to integrate the findings of the two central parts of the thesis: operant conditioning of pain and the MPI-DLV clusters.

All these Chapters are adapted forms of already published or submitted articles. Consequently, some of the points already discussed in this Chapter can be found in a reformulated, and in most instances more detailed form in the other Chapters.

THE RELATIONSHIP BETWEEN SPOUSE SOLICITOUSNESS AND PAIN BEHAVIOR: SEARCHING FOR MORE EXPERIMENTAL EVIDENCE¹

2.1 Summary

In this study 42 chronic back pain patients participated twice in a treadmill test. During one of these two sessions, the partner was present. Walking time, pain intensity ratings, and heart rate were measured before and after the tests. From the results of previous studies it was expected that in the presence of a relatively solicitous spouse, patients would report more pain, would have a shorter walking time, and would exert themselves less physically. Spouse solicitousness was measured in two ways: from the patient's perspective as well as from that of the spouse. Results obtained on the basis of the patient's interpretation of their partner's responses are not in accordance with previous findings. Results obtained from the spouse's view demonstrate, however, that patients with solicitous spouses do, in fact, report more pain and walk for a shorter duration in the presence of the spouse than patients with relatively non-solicitous spouses. Theoretical and practical implications are discussed.

2.2 Introduction

Based on the work of Fordyce (1976), it has long been held that pain behavior is influenced by environmental contingencies, an idea that is derived from operant learning theory. Evidence has been published that appears to confirm this notion.

A direct test of the operant model of pain behavior was carried out by Linton and Göttestam (1985). On repeated presentation of experimental pain stimuli of constant physical intensity, pain intensity reports of healthy test subjects proved to vary systematically, depending on negative or positive feedback given by the experimenter upon the report of the test subject. A report of increasing pain could even be provoked during diminishing intensity of the experimental pain stimulus by

¹Adapted from Lousberg, R., Schmidt, A.J.M. & Groenman, N.H. (1992). The relationship between spouse solicitousness and pain behavior: searching for more experimental evidence. *Pain*, 51, 75-79.

positive verbal reinforcement of an increasing pain-report. Flor et al. (1987), using a questionnaire approach, found that the best predictor of the pain and activity levels reported by the pain patient was the patient's perception of spouse reinforcement, followed by the spouse's self-reported responses to the patient's pain. Data by Gil et al. (1987) likewise seem to support the operant paradigm. They demonstrated that pain behavior varied as a function of level of satisfaction with social support. Chronic pain patients reporting high satisfaction with social support exhibited significantly higher levels of such individual pain behaviors as guarding, rubbing, and bracing.

Note that the studies of Flor et al. (1987) and Gil et al. (1987) were based on self-report. A further step was taken by Block et al. (1980). Their study was carried out as part of the diagnostic intake procedure in a clinical setting. Results showed that chronic pain patients with relatively non-solicitous spouses reported significantly lower pain levels in a spouse-observing condition (behind a one-way mirror) than in a neutral observer condition. Also, patients who reported that their spouses were relatively solicitous in responding to pain behavior reported marginally higher levels of pain in a spouse-observing condition than in a neutral observer condition. This study experimentally demonstrated the relationship between the pain report of chronic pain patients and environmental influences, in this case, the influence of the partner.

The purpose of the present study was to experimentally investigate the relationship between direct environmental influences (spouse solicitousness) and the observable pain behavior of chronic pain patients using the methods employed by Block et al. (1980) but with the following modifications:

1. In order to maximalize the effect of solicitousness vs. non-solicitousness, the partner was physically present during the experimental session.
2. Pain behavior was measured not only by way of a VAS-intensity score but was also operationalized as poorer endurance in a walking-to-tolerance treadmill test for chronic low back pain patients. As this test places a great deal of stress on the back, back pain patients will terminate this endurance test sooner than non-back patients, something which may be seen as a form of back pain behavior (Schmidt, 1985).
3. The study was carried out outside the clinical setting and was independent of any form of treatment.
4. Spouse solicitousness was measured in two ways, namely, from the perspective of the patient and that of the partner (Flor et al., 1987).

In this study, the procedure of Block et al. (1980) was followed. The patient performed a pain behavior-provoking test (in this case, the treadmill test) twice and the partner, whose solicitousness was measured, was present at one test and absent at the other. Based on the findings of the aforementioned studies, it was expected that when compared to patients with non-solicitous spouses, patients with solicitous

spouses would: (1) walk for a shorter duration, (2) report a larger increase in pain intensity (scored on a VAS), and (3) exert themselves less physically in the presence of the spouse than in the absence of the spouse. Moreover, it was expected that given the dependent variables (VAS-intensity score, time of walking, and heart-rate), both measurements of spouse solicitousness would produce comparable results.

2.3 Method

Subjects

Forty-two chronic low back pain (CLBP) patients and their partners participated in the study. Forty-one patients were married, one female patient lived together. There were 32 men and 10 women. CLBP patients were recruited by means of an advertisement in a local newspaper.

Selection criteria included:

- age between 25 and 55 years;
- low back pain for at least six months;
- no cardiovascular disease, in view of the physical exercise in the research protocol;
- no distinct limitations on motor performance or difficulties of walking as a result of disease or surgery;
- no chronic pain complaints in partners of patients.

One female patient was not able to perform the treadmill test according to the instructions. She was excluded from the data analyses.

Procedure

The chronic back pain patients participated twice in a working-to-tolerance treadmill test with an intermission of 1½ weeks to exclude physical training effects. Treading speed was maintained at 5 km/h. The grade of ascent was set at 5% for the first minute and was increased 1% at 1 minute intervals.

All subjects were informed that the purpose of the experiment was to study physical reactions during back-stressing physical exertion. Prior to test one, a short explanation was given about the procedure during the treadmill test and the informed consent was signed. Debriefing took place following the second test. Also prior to test one, subjects were given the opportunity to get used to treadmill walking (speed = 3 km/h; grade = 5%) for about 2 minutes, depending on individual skill. Subjects performed the treadmill test in sportswear, and watches, if worn, were removed. Any recording equipment that might have given exertion-contingent feedback was placed outside the field of vision of the patients. The two experimenters gave no exertion-related verbal or non-verbal feedback or information preceding, during, or immediately after the treadmill tests. Subjects' requests

were referred to the debriefing period. Treadmill test instructions were to continue walking on the treadmill until the need was felt to stop due to pain or fatigue. Subjects were instructed to say "stop" when they wished to finish the protocol. During both treadmill test conditions the two experimenters were physically present but were out of the patient's view.

To control for a carry-over effect, the order of cue conditions was balanced across subjects. That is, in one half of the subjects the partner was present during the first session and in the other half during the second session. At the beginning of the treadmill test, the partner was placed within the field of vision of the patient and the two were allowed to talk to one another during the test. During the breaks in the protocol -changing and showering of the subjects-, the partner filled in a 'solicitousness' questionnaire. The patient filled in the patient version of this list in the condition during which the partner was absent.

Measures

1. Solicitousness was measured in two ways, using two parallel questionnaires. The first one, the patient version, was derived from section 2 of the West Haven-Yale Multidimensional Pain Inventory (WHYMPI) (Kerns et al., 1985). This scale contains items that assess a patient's perceptions of the range and frequency of responses by the partner to pain behavior and suffering (according to the method of Block et al. (1980)). The items on the spouse version are the same as those on the patient version with respect to the content, but they assess the reactions of the spouse to the patient's pain behavior. For example the item "My partner takes over my chores" from the patient version was replaced by "I take over my partner's chores" in the spouse version.

2. Treadmill test variables included VAS back pain intensity scores, both pre and post-test, and walking time (in seconds). In order to determine the patient's physical exertion level, heart rate was measured by a sport tester (type POLAR). The number of beats per minute was measured twice: just before starting the treadmill test and immediately after the patient said "stop".

3. Biographical variables, including a pain history interview, were also measured.

2.4 Results

Due to technical problems with the sport tester, heart rate data were missing for one subject. The multivariate analyses were, therefore done on data from 40 subjects. Before testing the hypotheses the dependent variables were examined for normality. They were all normally distributed, allowing for parametric testing.

Reliability of the solicitousness measures

The internal reliability of the patient version -computed from the 40 patient lists in this experiment- was adequate ($\alpha=0.64$). This reliability estimation is based on

a relatively small number of cases. As a result the obtained alpha may be an underestimation. This notion was confirmed by recent psychometrical research on the translated (Dutch) version of the MPI (Lousberg et al., 1993c). Analyses were performed on 473 completed lists. Results of confirmatory factoranalyses and reliability measures showed that both the structure of the MPI as well as the internal consistency of the scales are comparable to the results reported by Kerns et al. (1985). The reliability of the solicitousness scale was good ($\alpha=0.76$). The reliability of the spouse version was measured with the 40 completed lists from this experiment. The reliability coefficient was moderate ($\alpha=0.43$). Again, the small number of cases, may account for the low alpha of the spouse version.

Spouse response scores as measured by the patient version

The median of the spouse response score was 14. Patients whose scores were below the median were assigned to the non-solicitous group (mean=9.1, s.d.=4.3, N=19); those with a score higher than 14 formed the solicitous group (mean=21.1, s.d.=4.59, N=21).

Table 2.1. Patient characteristics in the (non)-solicitous group based on spouse response scores as measured by the patient version

	Non-solicitous spouse		Solicitous spouse	
	mean	s.d.	mean	s.d.
Age (years)	48.0	8.44	45.3	4.93
Duration of complaints (years)	16.6	10.02	16.3	9.62
Spouse response score	9.1	4.30	21.1	4.59
Sex	15 men; 4 women		17 men; 4 women	

Table 2.2. Patient characteristics in the (non)-solicitous group based on spouse response scores as measured by the patient version

	Non-solicitous spouse (N=19)		Solicitous spouse (N=21)	
	mean	s.d.	mean	s.d.
Time *				
-alone	734.7	376.9	748.5	334.6
-with partner	727.2	402.7	735.1	349.5
Pain **				
-alone	12.9	29.1	13.9	19.9
-with partner	17.9	30.2	16.2	21.7
Heart rate ***				
-alone	75.2	25.5	69.3	16.3
-with partner	72.3	23.2	69.1	26.0

* : time of walking in seconds.

** : difference in pain VAS-score (after - before test).

*** : difference in heart rate in beats / min. (after - before test).

In Table 2.1 some characteristics of the solicitous and non-solicitous groups are mentioned. The groups do not differ significantly in age ($t(1,38)=1.26, p=0.22$), duration of complaints ($t(1,38)=0.09, p=0.93$), or sex ($\chi^2=0.025, p=0.87$).

In Table 2.2 the means for walking time, VAS-pain difference, and in heart rate increase are shown. It appears that the responses in both groups are the same: more pain is reported, the walking time is shorter, and the physical exertion level is less in the spouse condition.

To assess the effects of the two cue conditions, a repeated measures design (MANOVA) was performed. The between-group factor was group (solicitous spouse vs. non-solicitous spouse). The within-subject factor was cue condition (presence/absence of the partner). The three dependent variables were walking time, pain difference score, and heart rate difference score. The multivariate interaction effect was not significant ($F(\text{Wilks})=1.42, p=0.25$). Moreover, none of the univariate analyses contributed significantly to this interaction effect (All $F_s(1,36) < 0.35, P_s > 0.55$). It can thus be concluded that there are no different response patterns between the two groups.

Spouse responses scores as measured by the spouse version

The median of the spouse response score was 20. Patients whose partners had a score below the median formed the non-solicitous group (mean=15.7, s.d.=3.66, N=19); those whose partners had a score above 20 formed the solicitous group (mean=24.0, s.d.=3.46, N=21). In Table 2.3 some characteristics of the solicitous and non-solicitous groups are mentioned.

Table 2.3 Patient characteristics in the (non)-solicitous group based on spouse response scores as measured by the spouse version

	Non-solicitous spouse		Solicitous spouse	
	mean	s.d.	mean	s.d.
Age (years)	47.9	7.25	45.1	6.28
Duration of complaints (years)	14.3	9.32	19.3	9.68
Spouse response score	15.7	3.66	24.0	3.46
Sex	13 men; 6 women		19 men; 2 women	

The groups do not significantly differ in age ($t(1,38)=1.52, p=0.14$), duration of complaints ($t(1,38)=1.48, p=0.15$), or sex ($\chi^2=2.734, p=0.10$). In Table 2.4 the means and standard deviations are indicated for the dependent variables. Contrary to the results of the partner version, different response patterns were found. Patients in the solicitous group reported a greater increase in pain, had poorer endurance, and seemed to exert themselves less physically during the spouse condition. The opposite effect can be seen in the non-solicitous group. The multivariate interaction effect (conditions x solicitousness) was marginally significant ($F(\text{Wilks})=2.50, p=0.08$). Univariate analyses revealed that increase in pain contributed significantly to the overall conditions x solicitousness interaction effect

($F(1,36)=4.69$, $p=0.01$). Walking time also contributed significantly to the interaction effect ($F(1,36)=2.62$, $p=0.05$). However, the univariate effect of the increase in heart rate was not significant ($F(1,36)=0.78$ $p=0.38$).

Table 2.4 Means and standard deviations of walking time and differences in pain ratings and heart-rate scores (spouse version)

	Non-sollicitous spouse (N=19)		Sollicitous spouse (N=21)	
	mean	s.d.	mean	s.d.
Time *				
-alone	745.5	360.4	738.7	350.6
-with partner	764.6	399.4	701.2	350.1
Pain **				
-alone	13.9	27.0	13.0	22.4
-with partner	12.5	28.3	21.1	23.2
Heart rate ***				
-alone	70.7	23.3	73.4	19.4
-with partner	71.7	23.2	69.7	26.0

* : time of walking in seconds.

** : difference in pain VAS-score (after - before test).

*** : difference in heart rate in beats / min. (after - before test).

As for the discrepancy between the results of spouse and patient versions, a post hoc analysis was performed to measure the degree of agreement between the two scales. The correlations between the corresponding items on the patient version and those on the spouse version were very low (r 's ranging from -0.06 to 0.29 with p -values ranging from 0.36 to 0.03), indicating different ideas about spouse sollicitousness. Finally, a t-test was conducted to investigate whether the spouse response score differed significantly between spouses and patients. The mean spouse response score as measured by the patient version was 15.4. This was significantly lower than the mean score from the spouse version, 20.05 ($t=3.11$, $p=0.003$). It should be noted that higher scores mean more sollicitousness.

2.5 Discussion

It is not easy to draw unequivocal conclusions from the present results. One factor was already mentioned: the low reliability of the spouse version. Increasing the N of completed spouse lists will, because of a lower standard error, result in a more precise estimation of the internal reliability coefficient. However, this new estimate will not necessarily lead to an enlargement of the alpha coefficient, as was the case with the alpha of the patient version. Second, the results of the patient version do not agree with the findings of the Block et al. study (1980). On the other hand the results obtained from the spouse version are clear. Taking a number of differences between this study and that of Block into account (setting, dependent

variables), the results of the VAS-pain scores are consistent with those in Block's experiment. Furthermore, and perhaps most importantly, the results of the spouse version imply a generalization of the Block findings. Not only pain intensity ratings (VAS) but also more observable pain behaviors (as measured by walking time) are subject to environmental influences. Although this study is not such a direct test of the operant paradigm as the Linton and Götestam study (1985), the data from the spouse version generally support predictions made from operant conditioning mechanisms (Fordyce, 1976).

The contradictory results may be explained by the low correlations between the patient and the partner version items. This low correlation can be interpreted in two ways: it could be a reflection of the moderate reliability coefficient of the spouse version; on the other hand it is possible that the perceptions held by the patient and spouse differed strongly, at least in this study.

The fact that only the results of the partner version are consistent with previous research is not in line with cognitive-behavioral statements on pain assessment. In this view, emphasis is placed on the assessment of the patient's own perspective of his pain, the reactions of others to his pain, etc. (Turk et al., 1983). However, the present results suggest that the partner's opinion about his/her reactions to the pain behavior of the patient is also important and perhaps more objective. Another explanation for the difference in findings from Block's study may be the fact that patients in this study were recruited from a newspaper advertisement whereas those in Block's study were recruited from a chronic pain program. Therefore these two samples of chronic pain patients may not have the same characteristics. Although our patient sample is comparable to that of Block et al. (1980) with regard to the overall mean age and pain duration, the differences in age and pain duration between the solicitous and non-solicitous group, as reported by Block, were not found in the present study.

In a different, recent study by Turk and his colleagues (Turk and Rudy, 1988), more evidence is found for the relationship between solicitousness of the patient's spouse, pain intensity ratings, and activity level. In search of subtypes of patients, thereby considering psychosocial and behavioral assessment data, they performed a cluster analysis on the scales of the Multidimensional Pain Inventory. Results indicate that a heterogeneous group of chronic pain patients can be subdivided into three subtypes. Considering the scales of the patient subtype labeled 'Dysfunctional', a similar pattern of interrelationships emerged, as is found in both present and previous studies. Relatively high pain intensity levels accompany high solicitous response scores and a low general activity level. The other two patient types are characterized by a more or less inverse relationship between these scales. Combining Turk's (1988) results with Block's findings (1980) and with our spouse version findings, it could be argued that operant conditioning factors play an important role as an underlying mechanism of the 'Dysfunctional' type. If the suggested relationship between operant factors and the Dysfunctional type is true, the present findings may not only have a theoretical value with respect to the understanding of maintaining mechanisms of chronic (back) pain problems, but may also have

therapeutic implications. If, in a diagnostic phase, a patient is assigned to one of the three subtypes, the Dysfunctional type might be expected to respond favorably to an operant treatment approach.

In conclusion, some of the present results support operant notions and earlier experimental evidence. Moreover, it appears (cf. the work by Turk cited above) that experimental data may be consistent with diagnostic/psychometric data. It should be noted again that the moderate reliability of the spouse version prevents, at least at this moment, definite conclusions. Further research in this area would no doubt prove to be rewarding.

OPERANT CONDITIONING OF THE PAIN EXPERIENCE¹

3.1 Summary

The aim of the present study was to replicate previous findings on operant conditioning of pain report. In addition, an attempt was made to determine whether operant conditioning of pain report reflects the pain experience. This was done by measuring a psychophysiological parameter related to the pain experience - skin conductance response - and by carrying out a magnitude matching procedure. Subjects received 12 painful electric shocks of equal intensity. Reports of pain increase from subjects in the experimental group (N=23) were verbally rewarded; reports of pain decrease were punished. The pain ratings of subjects in the control group (N=21) were not reinforced. The results indicate that the attempt at replication succeeded. There was a clear increase in pain report in the experimental group. The skin conductance responses also showed a significant increase. The results of the magnitude matching were in accordance with the predicted effect. The increase in pain report could largely be explained by the punishment of decreases in pain ratings. Clinical implications are discussed. It is concluded that the present results support the notion that operant conditioning influences the experience of pain.

3.2 Introduction

The crux of operant learning theory is that behavior frequency is influenced by environmental contingencies. Behavior that has positive consequences (i.e., that is followed by a reward or the withdrawal of an aversive event) is likely to increase in frequency. On the other hand, if behavior is followed by an aversive consequence, or if it is followed by the omission of the expected reward, the frequency of this behavior will tend to decrease.

One of the behavioral problems to which operant techniques are often applied, is chronic benign pain. Chronic benign pain has frequently been described as a state of continuous pain for at least 6 months, characterized by a discrepancy between observed patient pain behavior and an underlying somatic/medical dysfunction

¹Submitted for publication by Lousberg, R., Schmidt, A.J.M., Groenman, N.H. & Gielen, A.A.C.M. (1993a).

(Feuerstein et al., 1986). This discrepancy can be explained in terms of operant learning. Pain is a special type of behavior. Therefore, pain behavior that is rewarded is likely to increase in frequency. Pain behavior such as resting, limping, grimacing, etc. may remain after the "normal" healing time has passed, to the degree that the relevant behaviors have had by positive consequences.

Intervention programs have been based on the premise that pain behavior is operantly conditionable (Fordyce, 1976). The majority of studies investigating the effect of the operant approach report positive and promising effects: an increase in daily activities, a reduction in medication, a reduction in the amount of time spent lying down because of pain, etc. (Turner and Chapman, 1982). Though this is not the primary purpose of operant interventions, some studies also report a decrease in the subjectively reported pain intensity level (Fordyce et al., 1973).

Although the results of operant behavior therapy programs are encouraging, this does not necessarily mean that therapeutically reduced pain behavior has, in fact, resulted from operant conditioning; therapeutic effects do not prove causation. The question is whether pain behavior really can be learned by operant conditioning. There has been very little fundamental research in this area. An often cited study is that by Block et al. (1980). They demonstrated that chronic pain patients with relatively non-sollicitous spouses reported significantly lower pain levels in a spouse-observing condition (behind a one-way mirror) than in a neutral observer condition. Conversely, patients who reported that their spouses were relatively solicitous in responding to pain behavior reported marginally higher levels of pain in a spouse-observing condition than in a neutral observer condition. Block et al. (1980) demonstrated experimentally that the pain report of the chronic pain patient is related to environmental influences, in this case the presence of the partner. Similar results were obtained by Lousberg, Schmidt and Groenman (1992). In their experiment, back pain patients twice performed a walking-to-tolerance treadmill test: once alone, and once in the presence of their spouse. Spouse solicitousness was measured both from the patient's and from the spouse's perspective. The results obtained, using the spouse solicitous measure from the spouse's perspective, indicated that in the presence of a solicitous spouse, patients performed more poorly, i.e., they had a shorter walking time, and reported more pain than in the absence of the spouse. Analyses of spouse solicitousness, as measured from the patient's perspective, showed no difference in walking time or pain report between the two cue conditions.

More direct experimental evidence indicating that pain report is under operant control is provided by Linton and Göttestam (1985). Healthy subjects received a number of calibrated pain stimuli (using a blood-pressure cuff) and were asked to rate the pain intensity experienced. The conditioning of pain report took place in two ways. Up-conditioning was established by verbally rewarding the subject if pain report increased compared to the previous trial. Down-conditioning of pain report was achieved by rewarding a decrease in the pain score. Both forms of conditioning succeeded. Even where the objective intensity level of the pain

stimulus was decreased, up-conditioning, i.e. an increase in pain report, could be demonstrated.

The studies cited above show that pain behavior is modified by environmental contingencies, and the Linton and Göttestam study suggests that pain behavior is under operant control. It should be noted, however, that the dependent variable in these three studies was pain report. Pain report is usually taken as an indication of the pain experience, but report and experience should not be equated. The main issue is what is learned when pain behavior comes under environmental control. Is it merely the reporting of pain, or is it the experience of pain that is affected? Of course the personal experience of pain cannot be directly measured. If, however, operant conditioning were able not only to affect reported pain, but also psychophysiological parameters that are normally affected by nociception, this would lend considerable support to the theory that operant conditioning influences the very experience of pain.

The purpose of this study is twofold: (1) to replicate the results of Linton's experiment concerning up-conditioning by means of a stable noxious stimulus and (2) to investigate whether other psychophysiological and psychological response systems, related to the pain experience, can also be conditioned. In the present study, pain report was measured using visual analogue scoring. Often, approximations of the "true" pain experience were obtained using the following measures:

1. Skin conductance response. Experimental work by Arntz et al. (1991) has shown that there is a significant correlation (0.43) between stimulus intensity (Ampere) level and skin conductance response.
2. A procedure that is comparable to a magnitude matching procedure, a technique described by Stevens (1956). The subject is presented with a stimulus - in this case, a pain stimulus - of a standard intensity. The subject is then asked to judge the intensity of a new pain stimulus relative to the standard one.

The experiment was similar to that of Linton and Göttestam (1985), but the following modifications were made:

1. A pilot study indicated that more than 50% of healthy subjects rated blood pressure cuff levels of more than 180 mm Hg, (as used by Linton and Göttestam) as not at all painful. Therefore, in the present study, another pain stimulus was used; electrical stimulation.
2. Psychophysiological recordings were made by measuring skin conductance response (SCR) and skin conductance level (SCL).
3. A magnitude matching procedure was added. After the series of pain stimuli, the subject was asked to estimate the intensity of the first and last stimuli. If the pain experience in the experimental group had become stronger at the end

of the experiment as a result of conditioning, the second matching (i.e., from the last trial) would be higher than the first matching, compared to the results from a control group in which no conditioning had occurred.

The following hypotheses were tested:

1. Social rewards for pain report will produce an increase in reported pain intensity.
2. An increase in pain report, occasioned by reward, is accompanied by an increase in electrodermal activity.
3. Rewards for pain report produce a stronger experience of a certain pain intensity level.

3.3 Method

Subjects

The subjects were 44 healthy volunteers who were paid to participate in the study. There were 22 women and 22 men. Forty-one of the subjects were students, two were unemployed individuals and one was working. The average age was 23 (range 18-39 years, s.d.=7.0). Due to technical problems there were no physiological recordings available for the first six subjects.

Materials

The subjects reported pain intensity on visual analogue scales (VAS). These are 10-cm horizontal lines ranging from 0 ("not at all painful") to 100 ("extremely painful").

Apparatus and physiological recording

The electric pain stimuli were delivered via a Neuroton 626, adapted for use with patients. The painful shocks were given via two Beckman Ag-AgCl-electrodes (8 mm diameter; 25 mm between the centers of the electrodes) attached to the subject's ankle, on the side opposite the dominant arm. For the magnitude matching procedure, these electrodes were placed on the non-dominant forearm.

SCR and SCL were measured using a Beckman Skin Conductance Coupler (type 9844) at a constant voltage of 0.5 V. The coupler allowed for a maximum sensitivity of 0.05 micromho. The electrodes were attached by means of adhesive collars, to the medial phalanges of the second and third fingers of the non-dominant hand. Beckman Ag-AgCL-electrodes (8 mm diameter) filled with an isotonic paste were used to measure SCR. The highest deflection within the 10-s stimulation period was measured. SCRs were square root transformed to reduce skewing (cf. Levey, 1980).

Respiration was recorded by means of a Beckman respiration belt fastened around the subject's chest and connected to a Beckman Voltage/Pulse/Pressure Coupler (type 9853A) in order to detect SCRs due to respiratory irregularities.

A personal computer was used to control the protocol.

Procedure

Subjects were matched for sex and age and were assigned either to the control group (N=21) or to the experimental group (N=23). There were two experimenters, one female and one male, who took turns in conducting the experiment, where he or she was physically present, sitting behind the subject. The other experimenter was in an adjacent room, monitoring the physiological recordings. Measures were taken to ensure that both the sex and age of the subject, and control versus experimental conditions were equally distributed / balanced for both experimenters.

The subjects were brought into the laboratory and seated in a chair facing a blank wall. After informed consent had been obtained, the electrodes (SCR and shock electrodes) and respiration belt were fastened onto the subject. The shock level to be used was then determined (calibration phase). The subjects were told that the intensity level would be slowly increased, starting with a zero intensity level. The subjects were asked to say "stop" as soon as they thought the intensity level could be rated at 50 mm VAS, which was defined as being a level that was perceived by the subject as painful but far from intolerable. The intensity level was raised (about 0.1 mA per second) until the subject indicated that the level was equal to 50 mm VAS. It was stressed that the purpose of the experiment was not to investigate how much pain the subject could bear.

The subjects were instructed to concentrate on their ankle and to fill in a separate VAS after each stimulus. They were asked not to speak to the experimenter during the experiment. Furthermore, the subjects were told that the purpose of the experiment was to investigate whether they could detect small differences in stimulus intensity level. In order to allow the use of the entire range of the VAS scale, the subjects were told that these variations in shock might be experienced quite differently by each subject.

Each subject, both in the control and in the experimental group, received 12 unannounced shocks, **all equal** to the calibrated level. The shocks lasted 10 s and were given at random intervals (mean 30s, range 25-35s).

Control condition

Subjects in the control condition received neutral feedback (i.e., the experimenter said "thank you") after each pain rating.

Experimental condition

The first four trials were considered as the baseline phase. After these four pain ratings, the experimenter gave a neutral feedback - "thank you" - as in the control group. The conditioning phase started after trial 5 (i.e., **after** the subject had

delivered the VAS scale from the fifth trial) and continued up to the end of the experiment (trial 12). Up-conditioning of pain report took place by verbally rewarding increases in pain scores and "punishing" decreases in pain scores. If the VAS pain score was greater than or equal to the previous VAS score, the response was rewarded with such positive statements as "That's good", "Nice", "I'm satisfied with this", etc. If, on the other hand, a VAS score was lower than the previous one, the response was punished by such statements as "Hmm, that's not good", "Concentrate more on your ankle", etc. Thus, the decision whether to positively reinforce or to punish a response was made on the basis of the VAS score that had just been given, compared to the VAS score from the previous trial, with the exception of trial 5. The VAS score from this trial was compared to the mean VAS score from the baseline (trials 1-4). As was the case in Linton and Göttestam's (1985) study, these verbal statements were made by the experimenter in the form of subtle 'thinking out loud' comments.

After trial 12, the shock electrodes were placed on the subject's non-dominant arm, halfway along the inner side of the forearm. The instructions regarding the magnitude matching procedure were as follows. Subjects were asked to remember the intensity of the first trial (of the series of 12). They were asked, as precisely as possible, to re-set the intensity level so that it was equal to the experience of the first trial. The procedure for setting this intensity level was identical to that of the calibration phase. After this matching, subjects were asked to remember the intensity of the last trial of the experiment. Again, they were asked to set the intensity level as precisely as possible. All information that could possibly influence the two matchings was withheld from the subjects. No information was given regarding the calibration level and the subjects were seated with their backs to the Neuroton. Finally, another possible influencing factor was ruled out. It could be hypothesized that subjects based their matchings on a comparison of the time taken in the calibration phase, before they reached their estimated level. Therefore, in both matching trials, the experimenter kept the stimulus level at 0.0 mA. for 10 s before actually increasing the level. The speed at which the Ampere level was increased was the same as that in the calibration phase. The whole experimental session lasted approximately 45 minutes.

3.4 Results

Before testing the hypotheses, some analyses were carried out to examine possible initial group differences. The mean Ampere level in the control group was 2.9 mA.; that in the experimental group, 4.5 mA. This difference was significant ($t(42)=3.15$, $p=0.004$) and was taken into account by using the Ampere level as a co-variate.

A MANCOVA was carried out on the first five trials. It should be noted that both the VAS score and the SCR from the fifth trial could not be influenced by conditioning. Neither the multivariate test nor the univariate tests revealed

significant differences between the two groups (p always > 0.13) with regard to the VAS scores. The multivariate effect for the first five trials of the SCRs was not significant either ($F(1,32)=0.08$, $p=0.78$). However, the univariate test for the SCR on the first trial was significant ($t(35)=2.16$, $p=0.04$). This effect was not due to outliers. The difference in Ampere level between the two groups most likely accounts for this initial stronger response in the experimental group. It was therefore decided to take the mean score of trials 2 - 5 as the reference point for SCRs. In order to carry out the subsequent analyses in a consistent manner (and also for VAS scores), the mean score from trials 2 - 5 served as the reference point. Furthermore, because of matching, the control group and the experimental group did not differ either in sex or age of the subject, or in sex of the experimenter. The following steps were taken to test whether conditioning had taken place. From trial 6 onwards, difference scores were computed for each subject, with the mean score from trials 2 - 5 as a reference. For example, the difference score for trial 10 was the VAS score from trial 10 minus the mean VAS score from trials 2 - 5.

VAS pain report

As can be seen in figure 3.1, the VAS pain scores of the control group show a habituation curve. In the experimental group the VAS pain report scores show a clear increase during trial 6, the first trial in which an increase in pain report could be expected. This increase continues through the last (12th) trial.

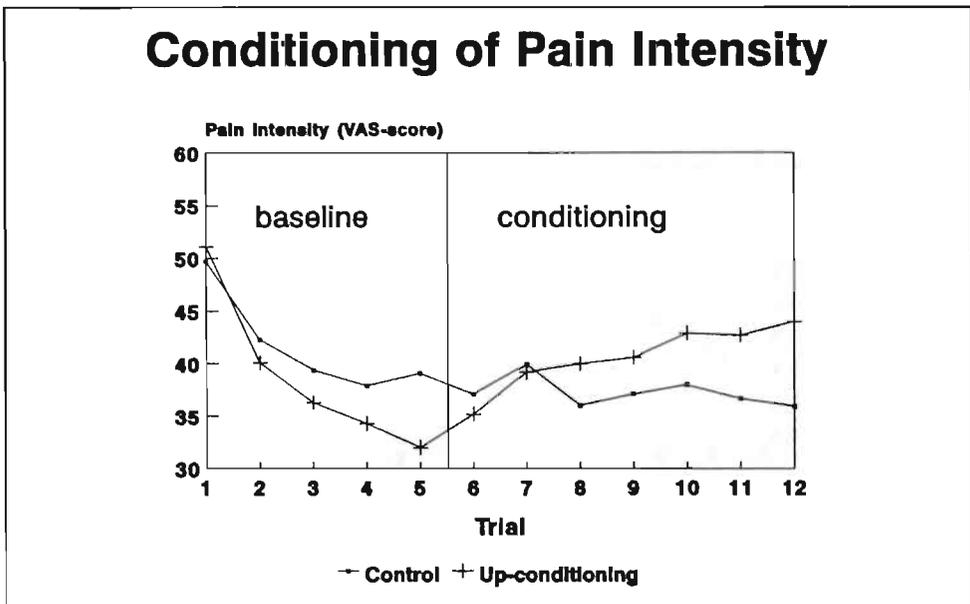


Figure 3.1 Conditioning curves of Pain Intensity.

Multivariate repeated analyses of co-variance were carried out to determine whether the response patterns differed in the two conditions. Difference scores from trials 6 - 12 were used as the within factor, condition as the between factor, and Ampere level as the co-variate. The multivariate effect was significant ($F(1,41)=5.35, p=0.01$). When the analysis was carried out on the last three trials (10, 11 and 12), as in the method used by Linton, the difference between the two patterns was also significant ($F(1,41)=5.89, p=0.01$).

We investigated in a post hoc analysis whether the effect of a reward on the subsequent stimulus was equal to that of a punishment. After a punishment (resulting from a decrease in VAS score), the next VAS score was 4.71 mm higher. After a reward, the next VAS score was 1.71 mm lower. This difference was significant ($t(262)=4.51, p < 0.001$). Thus, on the average, a punishment led to clear increases in the VAS score, whereas a reward resulted in a slight decrease. The increase in VAS scores can be explained by the punishment of decreasing pain intensity reports.

Skin conductance response

The SCRs are depicted in Figure 3.2. The control group is marked by a stable curve after trial 6. The magnitude of the responses tends to decrease, indicating habituation. In the conditioning phase, the SCR's from the experimental group

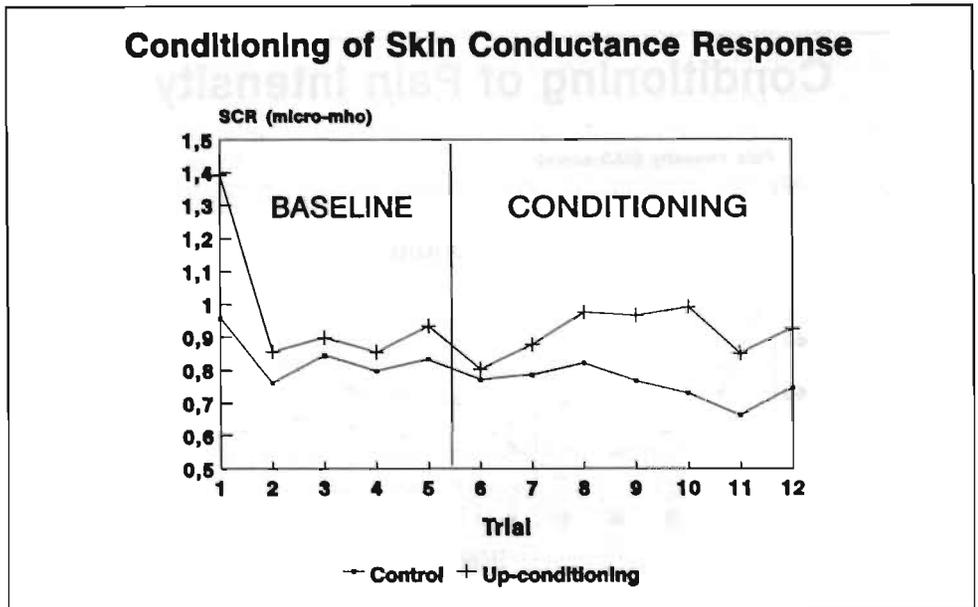


Figure 3.2 Conditioning curves of Skin Conductance Response.

show an increase after the sixth trial. The multivariate test, taking the difference scores from trial 6 through trial 12 together, was marginally significant ($F(1,43)=2.01, p=0.08$). An analysis of the last three trials, however, revealed a significant difference ($F(1,34)=4.29, p=0.02$).

Apart from the SCR, analyses were carried out on the SCL data. A test was conducted to see whether there were different SCL patterns during the conditioning phase between the two groups. No significant difference was found ($F(1,34)=0.94, p=0.34$).

Magnitude matching

Since the mean Ampere level of the experimental group was significantly higher than that of the control group, the decision was made to create matched pairs with respect to the Ampere level. A pair was considered as matched if the difference between the Ampere levels was less than 0.1 mA. Using this criterion, eight pairs could be created. The mean Ampere level for the eight subjects in the experimental group was 2.87 mA.; that for the control group was 2.8 mA. No significant differences in Ampere level were found as a result of the matching procedure ($t(14)=0.06, p=0.95$).

The interaction effect, as illustrated in Figure 3.3, was marginally significant ($F(1,14)=2.40, p=0.08$). Compared to the control group, subjects in the experimental group experienced the last shock as more painful than the first one.

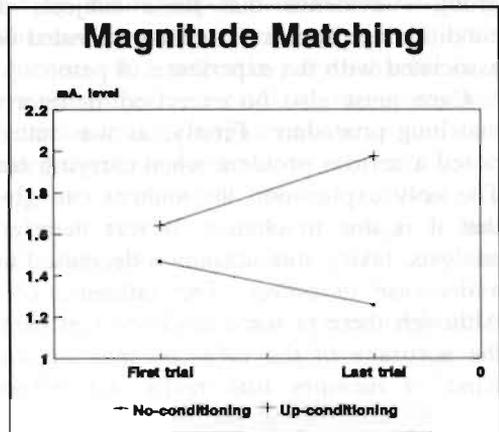


Figure 3.3 Magnitude matchings for control- and up-conditioning group.

3.5 Discussion

The first aim of this study was to replicate Linton and Göttestam's findings (1985) on the up-conditioning of pain report with a stable, noxious stimulus. Although a different pain stimulus was used in this study, the present results show a clear resemblance to those reported by Linton and Göttestam. From trial 6 onwards, the first trial in which the predicted effect could be expected, an increase in the VAS pain score was found in the experimental group. The control group curve represents a habituation pattern comparable to those found in the control groups of several previous studies (Arntz and Hout van den, 1988; Arntz and Lousberg, 1990). It can be concluded that the attempt at replication was successful.

The second part of the study dealt with the question of whether operant conditioning can also influence psychophysiological and psychological response systems. With respect to the SCRs, the hypothesized pattern was found. Although the increase in the SCR was not as immediate as that in the VAS pain score, the effect was obvious. Compared to the control group, the experimental group responded with greater physiological activity to the stimuli in the conditioning phase. It has to be noted, however, that the electrodermal response is known to be a measure associated with several psychophysiological processes (e.g., anxiety). It could be argued that the method of conditioning, including punishment for decreasing scores, could also result in the conditioning of anxiety. In this case, higher SCRs to the electric shock could be explained as an effect of an increase in anxiety and not merely as an increase in the pain experience. An increase in anxiety level correlates positively with a heightened SCL. Though the analysis on the SCL showed no increase in the experimental group compared to the control group - evidence that these subjects did not become more anxious by the conditioning process - it cannot be ruled out that the increase in SCR is exclusively associated with the experience of pain.

Care must also be exercised in interpreting the findings from the magnitude matching procedure. Firstly, as was stated earlier, the difference in Ampere level posed a serious problem when carrying out an analysis on all the subjects together. The only explanation the authors can give for this difference in Ampere level is that it is due to chance. It was therefore decided to carry out a matched pair analysis, taking into account a decreased number of subjects and, as a consequence, a decrease in power. The influence of memory bias must also be mentioned. Although there is some evidence that pain recall is accurate (Hunter et al., 1979), the accuracy of the subjects' matchings in this study is unknown. On the other hand, if memory bias really did influence the magnitude matching, this bias probably affected the matchings of both groups to the same extent.

Another issue pertaining to the validity of the present results, is the role of 'social desirability'. The effects of the intervention were clear as far as the results of the VAS pain report were concerned. It may be argued that the change in the reporting of pain merely reflects a 'social desirability' effect, i.e., subjects may have guessed the true purpose of the experiment and they may have reported their pain according to this assumed purpose. In response to such a claim, it could be pointed out that the social desirability effect, documented in the field of social psychology, can easily be translated in terms of operant conditioning, while the other way around, the effects of operant conditioning can easily be conceptualized as 'just' social desirability. The SCR data is of particular interest in this context. If reported pain reflects the pain experience, and if it is the pain experience that is subject to operant conditioning, we would expect conditioning effects not only on the VAS but also on the SCR. This is precisely what was observed. If changes in reported pain reflect social desirability, independent of any changes in the 'pain experience', one would expect VAS scores to change, but not the SCR. There is no

a priori reason to believe that giving socially desirable responses is accompanied by much electrodermal activity.

The discrepancy between the effect of a reward and a punishment is interesting. Whereas punishing the subject for a decrease in pain report resulted in a clear increase in pain report, a reward for an increase in pain report only showed a slight decrease in pain report. Translated into clinical practice this implies that 'punishment' of healthy behavior, such as complaining less or participating more in activities, is perhaps a more important factor in developing and maintaining pain behavior than a 'reward' for pain behavior. The effect produced by an irritated spouse who observes the patient's pain behavior after the patient has become involved in an activity that both the patient and spouse know will raise the patient's pain level, is perhaps more harmful for the patient's pain behavior than the situation in which the patient's pain behavior is followed by such consequences as receiving attention.

If these findings were to be supported by future experimental research, they could have therapeutic implications. Whereas nowadays the emphasis is often placed on identifying the positive consequences of a patient's pain behavior in a behavioral analysis, in order to serve as a guideline for treatment, more attention may have to be paid to identifying the negative consequences of pain behavior.

In conclusion, this study has demonstrated that pain report in healthy subjects is influenced by operant conditioning and, more precisely, by punishment for decreases in pain report. There is evidence that psychophysiological and psychological processes, associated with the pain experience, can also be influenced. Before definite conclusions can be drawn concerning the operant conditioning of the pain experience, the present findings will have to be replicated. Moreover, the results of this experiment have raised several questions that are worth considering in future research, such as: how do conditioning curves from a sample of chronic pain patients compare to those from healthy subjects?

A final suggestion would be to tackle the question of whether the opposite mechanism - down conditioning, as shown by Linton and Götestam (1985) - can be achieved. Such an experiment (which is currently being undertaken) could provide an answer to the question of whether the present results really reflect operant conditioning of the pain experience. If a decrease in VAS scores is accompanied by an increase in SCRs, the possibility of the conditioning of anxiety (see above) must be taken seriously. If, on the other hand, a decrease in VAS scores is accompanied by a decrease in SCRs, this would further support the interpretation of the present results as operant conditioning of the pain experience.

OPERANT CONDITIONING OF PAIN: A FAILED REPLICATION¹

4.1 Summary

Using a modified experimental design from a previous study, it was investigated whether both pain report and pain related psychophysiological and psychological measures (skin conductance response and magnitude matching) could be conditioned operantly. All the subjects received 12 painful shocks of equal intensity. Healthy subjects were assigned to either a control group (N=23), an up-conditioning group (N=27), or a down-conditioning group (N=25). Pain report in the up-conditioning group was conditioned by verbally 'rewarding' reports of pain increase and 'punishing' decreases. An inverse type of conditioning was employed in the down-conditioning group. The pain ratings of subjects in the control group were not reinforced. The attempt at a replication of previous findings, i.e., the up-conditioning of pain report and pain related psychophysiological and psychological measures, failed. Nor could down-conditioning of the pain report be established. These inconsistent results are most probably due to modified punishing responses. The consequences for the results of an earlier study are discussed. Some post-hoc analyses revealed that up-conditioning of the pain report did take place in a subgroup of the up-conditioning group. Another post-hoc finding made it clear that skin conductance responses were more associated with tension than with pain experience. Some suggestions are made for future operant conditioning studies of pain.

4.2 Introduction

Treatment of chronic pain based on the principles of operant learning theory has become popular and widespread - and for good reason. Operant treatment programs usually yield good therapeutic results (e.g., Linton, 1986). It has, however, been stated that the operant reduction of pain behavior does not necessarily indicate that pain behavior is learned in that way (Fordyce et al., 1985; Linton and Göttestam, 1985; Schmidt, 1987). The question of whether pain behavior really can be learned by means of operant conditioning processes has been the subject of increasing

¹Submitted for publication by Lousberg, R., Schmidt, A.J.M. & Groenman, N.H. (1993b).

interest in recent literature. It was found in two studies that, in the presence of a solicitous spouse, pain patients reported higher levels of pain than in a situation where there was no spouse present (Block et al., 1980). There is also evidence that, in the presence of a solicitous spouse, patients exhibit poorer physical performance (e.g., a shorter walking time in a working-to-tolerance test) than in the absence of the spouse (see Chapter 2).

A straightforward test on operant conditioning of pain behavior (in this case, pain report) was carried out by Linton and Götestam (1985). Healthy subjects received a number of calibrated painful stimuli and were asked to rate the pain intensity level experienced on a visual analogue scale (VAS). Results from this study suggested that pain intensity report systematically varied as a function of verbal rewards and punishments (uttered by the experimenter in the form of 'thinking out loud' comments). Probably the most striking result of this study was that in case of a decreasing objective intensity of the pain stimulus, up-conditioning (i.e., an increase in pain report) could be demonstrated.

In Chapter 3 it was investigated whether a) Linton's findings on up-conditioning could be replicated with a stable noxious stimulus and b) increases in pain report are accompanied by heightened psychophysiological responses (skin conductance response (SCR) and a psychological response (a magnitude matching procedure), presumed to be correlates of the pain experience. Forty-four healthy subjects were assigned either to a control group (no conditioning) or an experimental group (up-conditioning). All the subjects received 12 painful electric shocks. The method of up-conditioning of pain report was identical to that used by Linton, i.e., verbally rewarding increases of pain-ratings (as measured by means of a VAS) and punishing decreases. The results showed that pain report could be up-conditioned, thereby replicating Linton's findings. Analyses revealed that the up-conditioning effect could largely be explained by the punishment of decreases in pain report. The hypothesized effects concerning the other measures were also supported, in that subjects in the up-conditioning group showed heightened SCRs during the conditioning phase. They also indicated a relatively (marginally significant) higher magnitude matching with respect to the last pain stimulus². Although it was argued on page 41 that these results demonstrated that the up-conditioning procedure not only resulted in an increase in pain report but also produced heightened pain-

²Both the skin conductance and the magnitude matching procedure were carried out so as to be additionally certain that no conditioning of pain report alone took place. With respect to the magnitude matching procedure, subjects were asked to perform the following task. After the series of trials, they had to remember the intensity of the first stimulus and to calibrate this intensity again, as precisely as possible (without audio and visual feedback). Next, the subject was asked to repeat this procedure but this time for the last stimulus of the series of trials. It was hypothesized that, if conditioning of the pain experience had taken place, the level of the last matching would be higher than that of the first.

related psychological and psychophysiological responses, there are some problems in drawing definite conclusions.

One of the most likely alternative explanations has to do with an attention-effect connected with certain punishing responses in this experiment. Some of the experimenter's punishing 'thinking out loud' comments related to subjects' concentration on their task. Remarks like "Hmm, concentrate on your task more" are supposed to be perceived as a punishment because the importance of concentration was stressed just before the start of the experiment. Alternatively, however, these 'concentration-related' punishments may be seen as instructions to focus attention on the pain. Attention to a pain locus increases experienced pain (Arntz, 1991) and attention to a stimulus increases the SCR. Thus both the increased subjectively reported pain and the electrodermal effect might be (partly) due to increased concentration/attention.

There are also some difficulties with regard to the magnitude matching procedure. Firstly, the accuracy of the subject's matching is unknown and secondly, the significant difference in calibrated Ampere-level between the experimental and the control group prevented optimal statistical analyses.

As a consequence it was decided to carry out a new, similar experiment but with some modifications to overcome the interpretive and methodological problems mentioned above.

1. A down-conditioning group was added to the no-conditioning (control) group and the up-conditioning group. It was hypothesized that in the down-conditioning group the procedure would result in lower VAS-scores than in the control group. Further, if SCRs reflect pain experience, lower SCRs should occur in the down-conditioning group, and higher SCRs in the up-conditioning group. As to the magnitude matching procedure, the matching of the last trial should be lower in the down-conditioning group and higher in the up-conditioning group than in the control group.
2. With respect to a possible attention-effect, the verbal punishments did not explicitly refer to concentration.
3. The accuracy of subjects' magnitude matchings was assessed by asking the subject to perform a simple task. The result was interpreted as an indication of the subject's ability to make matchings.
4. In order to prevent differences in calibrated Ampere-levels between the three conditions, measures were taken when subjects were being assigned to them, to ensure that this variable was equally distributed for the three conditions.
5. The report on the experiment in Chapter 3 mentioned the conditioning of anxiety as an alternative explanation for the results obtained. Although a specific finding (similar SCL curves) appeared to refute this, further investigation of this effect was required. It was therefore decided to evaluate a possible change in subjective tension immediately after the series of trials.

4.3 Method

Subjects

The subjects were 75 healthy volunteers who received payment for their participation in the study. There were 49 women and 26 men. Sixty-nine of the subjects were students, and the other six people were in employment. The average age was 21.5 (range 18-27 years, s.d.=2.0). Due to technical problems physiological recordings were available for only 53 of the subjects.

Materials and physiological recording

The subjects reported pain intensity on 10 cm visual analogue scales (VAS) ranging from 0 ("not at all painful") to 100 ("extremely painful"). At the end of the series of trials the subjects' change in tension was indicated on a VAS-scale ranging from 0 mm ("much more relaxed"), 50 mm (no change in tension) to 100 mm ("much more tense").

Eleven wooden sticks, differing only in length, were used to test the subjects' ability to make accurate estimations. They measured 17, 19, 21, 23, 24.5, 25, 25.5, 27, 29, 31, and 33 cm, the 25 cm stick being regarded as the cue.

The electric pain stimuli were delivered via an ELTRON-D adapted for use with patients. The painful shocks were administered via two plate-electrodes (size 40 mm x 60 mm) wrapped in 60 mm x 75 mm wash-leather sponges dampened with lukewarm water. The electrodes were attached to the subject's ankle, on the side opposite the dominant arm, with 80 mm between the centers of the electrodes. For the magnitude matching procedure, these electrodes were placed on the inner side of the non-dominant forearm, with the same distance between the two electrodes.

SCR and SCL were measured using a Beckman Skin Conductance Coupler (type 9844) at a constant voltage of 0.5 V. The coupler allowed for a maximum sensitivity of 0.05 micromho. Electrodes were attached with adhesive collars to the medial phalanges of the second and third fingers of the non-dominant hand. Beckman Ag-AgCL-electrodes (8-mm diameter) filled with an isotonic paste, were used to measure SCR. The highest deflection within the 10-s stimulation period was measured. SCRs were square root-transformed to reduce skewing (cf. Levey, 1980).

Respiration was recorded by means of a Beckman respiration belt fastened around the subject's chest and connected to a Beckman Voltage/Pulse/Pressure Coupler (type 9853A), in order to detect SCRs due to respiratory irregularities.

A personal computer was used to control the protocol.

Procedure

Subjects were matched for sex and age and were assigned either to the control group (N=23), or to one of two experimental groups: an up-conditioning group (N=27) and a down-conditioning group (N=25). During the experiment a male experimenter was physically present, seated behind the subject. An assistant was in an adjacent room, monitoring the physiological recordings.

The subjects were brought into the laboratory and seated in a chair facing a blank wall. After informed consent had been obtained, the electrodes (SCR and shock electrodes) and respiration belt were fastened onto the subject's body. The shock level was then determined (calibration phase). The subjects were told that the intensity level would be slowly increased, starting with a zero intensity level. The subjects were asked to say "stop" as soon as they thought the intensity level could be rated 50 mm VAS, which was defined a level that the subject perceived as being painful but far from intolerable. The intensity level was increased (at about 1 mA per second) until the subject indicated that the level was equal to 50 mm VAS. It was stressed that the purpose of the experiment was not to investigate how much pain the subject could bear.

After completion of the calibration phase, the subject was asked to hold the 25 cm cue stick and to squeeze it gently in his/her right hand for one minute. The subject was told that the point of this procedure was to check whether electrodermal reactions could be registered. The experimenter left the room for one minute. The experimenter then took the cue stick and placed it among the other wooden sticks on a table in the adjacent room. The sticks were placed vertically (at a distance of 5 cm from one another) in a fixed sequence of length: 25.5, 19, 33, 23, 27, 17, 25, 29, 21, 24.5, 31 cm. The cue stick was thus placed in the 7th position.

The subjects were instructed to concentrate on their task and to fill in the experienced pain intensity level on a separate VAS, as precisely as possible, after each stimulus. They were asked not to speak to the experimenter during the experiment. Furthermore, the subjects were told that the purpose of the experiment was to investigate whether they could detect small differences in stimulus intensity level. In order to allow the entire length of the VAS scale to be used, the subjects were told that they might each experience these variations in shock quite differently.

Each subject, both in the control and in the experimental groups, received 12 unannounced shocks, **all equal** to the calibrated level. The shocks lasted 10 s. The inter-stimulus interval fluctuated randomly between 25 and 35 s. (mean 30 s.).

Experimental group 1: Up-conditioning

The first four trials were considered the baseline phase. After these four pain ratings, the experimenter gave a neutral feedback - "thank you" - as in the control group. The conditioning phase started after trial 5 (i.e., **after** the subject had given the VAS scale from the fifth trial) and continued until the end of the experiment (trial 12). Up-conditioning of pain report was carried out by verbally 'rewarding' increases in pain scores and 'punishing' decreases in pain scores. If the VAS pain score was greater than or equal to the previous VAS score the response was rewarded with such positive statements as "That's good", "Nice", "I'm satisfied with this", etc. If, on the other hand, a VAS score was lower than the previous one, the response was punished by such statements as "Hmm, that's not good", "This estimation is not good", etc. Thus, the decision on whether to positively reinforce or to punish a response was made on the basis of the VAS score that had

just been given, compared to the VAS score from the previous trial - with the exception of trial 5. The VAS score from this trial was compared with the mean VAS score from the baseline (trials 1-4). As was the case in Linton and Götestam's (1985) study, these verbal statements were made by the experimenter in the form of subtle 'thinking out loud' comments.

After trial 12, the shock electrodes were placed on the subject's non-dominant arm, halfway along the inner side of the forearm. The instructions regarding the magnitude matching procedure were as follows. Subjects were asked to remember the intensity of the first trial (from the series of 12) and then to re-set the intensity level, as precisely as possible, so that it was equal to the experience of the first trial. The procedure for setting this intensity level was identical to that of the calibration phase. After this matching, the subjects were asked to remember the intensity of the last trial of the experiment. Again, they were asked to set the intensity level as precisely as possible. All information that could possibly influence the two matchings was withheld from the subject; no information was given regarding the calibration level and the subjects were seated with their backs to the Eltron. To prevent subjects from basing their matchings on the time taken in the calibration phase to reach their estimated level, in both matching trials, the experimenter kept the stimulus level at 0.0 mA. for 10 s before actually increasing the level. The speed at which the Ampere level was increased was the same as that in the calibration phase. After the magnitude matching procedure and removal of the electrodes, subjects were taken to the adjacent room where they were asked to remember the length of the wooden cue stick and to choose the correct one out of the series of 11 sticks (without touching them). Finally, the subjects were debriefed. The whole experimental session lasted approximately 45 minutes.

Experimental group 2: Down-conditioning

Exactly the same procedure was followed as described for the up-conditioning group, except for an inverse type of conditioning. Decreases of pain report were rewarded, whereas increases were punished.

Control condition

The procedure for the control condition was identical to that of the experimental groups with one exception: subjects received neutral feedback (i.e., the experimenter said "thank you") after each pain rating.

4.4 Results

There were no significant differences between the three conditions with regard to the variables sex, age and milli-Ampere level (all P s > 0.17). As a check on initial differences with respect to VAS and SCR, MANOVAs were performed on the first five trials. The multivariate tests showed no significant differences (VAS-scores and SCRs: $p > 0.20$). With regard to the test relating to the subjects' ability to

make accurate estimations, it was decided to consider all choices within a range of 20 to 30 cm as acceptable. One subject chose stick number 2 (length 19 cm.). This subject was omitted from the magnitude matching analyses.

The following steps were taken in order to determine whether conditioning had taken place. From trial 6 onwards, difference scores were computed for each subject, with the mean score from trials 1 - 5 as a reference. For example, the difference score for trial 10 was its VAS score minus the mean VAS score from trials 1 - 5.

VAS pain report

Figure 4.1 depicts the three VAS profiles. The control group scores show a decreasing, habituation-like curve. The relatively sharp increase on trial 6 followed by the sharp decrease on trials 7 and 8 are mainly due to one subject's extreme

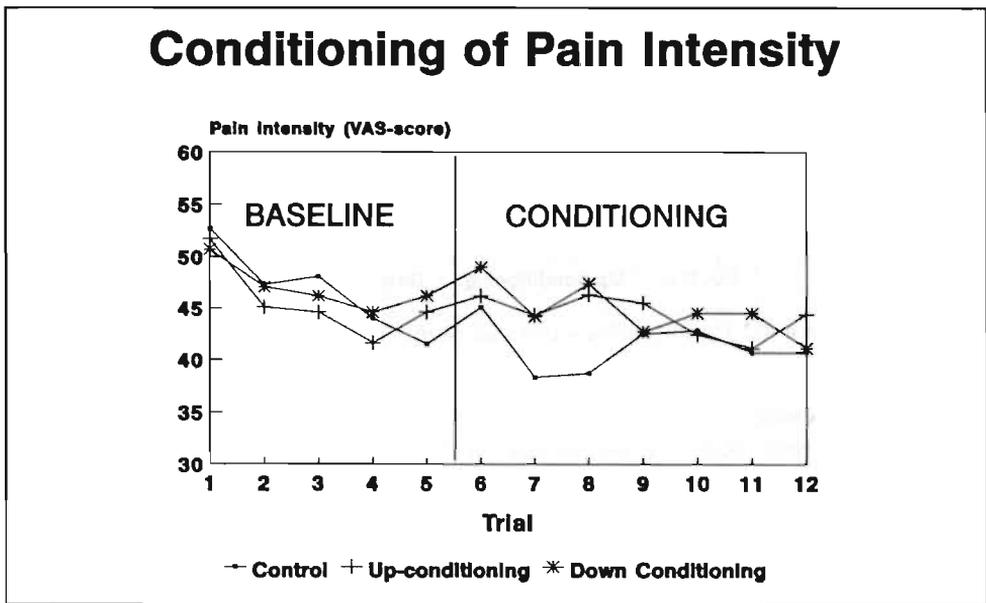


Figure 4.1 Conditioning curves of Pain Intensity.

scores on these specific trials. After trial five the up-conditioning group scores tend to remain at a stable level. After a smoothly habituating baseline, the down-conditioning group scores follow an undulating course of a generally decreasing tendency. Multivariate repeated analyses of variance indicated that there were no different response patterns during the conditioning phase ($F(2,72)=1.33, p=0.27$).

Skin conductance response

As can be seen from figure 4.2, all three SCR curves follow a stable decreasing course. The multivariate effect on the difference scores from trials 6 to 12 was not significant ($F(2,48)=0.09, p=0.92$).

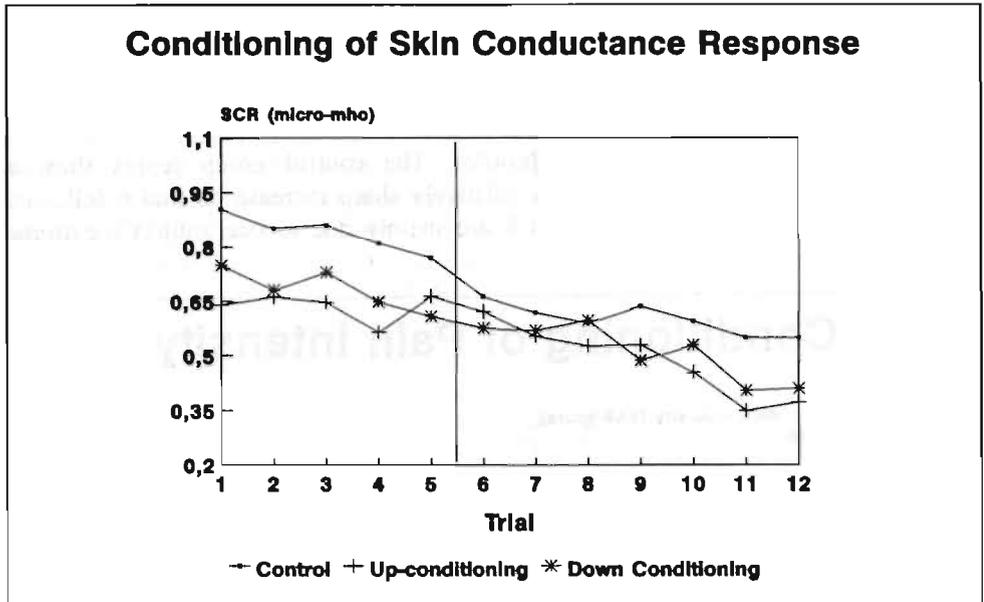


Figure 4.2 Conditioning curves of Skin Conductance Response.

Magnitude matching

Relative difference scores ($(\text{last} - \text{first}) / \text{first}$) were computed in order to compare the magnitude matchings from the first and last trials (see table 4.1).

Table 4.1 Magnitude matching (relative difference scores) in milli-Ampere

Condition	Mean	s.d.
Control	-0.129	0.158
Up	-0.036	0.280
Down	-0.113	0.236

There were no significant differences between the three conditions with regard to these matching differences ($F(2,72)=1.47, p=0.12$), all conditions matching the last trial somewhat lower than the first one. This finding concurs with the habituating VAS profiles.

Finally, an investigation was carried out to determine whether there had been a change in subjective tension during the series of trials. The mean values of the

three conditions were 39.6 mm (control group), 37.1 mm (up-conditioning group) and 37.9 mm (down-conditioning group), indicating that, on the average, subjects became more relaxed during the experiment. There were no significant differences between the conditions ($F(2,72)=0.12, p=0.89$).

4.5 Discussion

One of the purposes of the present experiment was to replicate findings reported in Chapter 3. In the present study some procedural modifications were made in order to be able to draw more definite conclusions about the significance of the results obtained both here and in the previous study, and a new experimental group (down-conditioning) was added. In addition to these modifications the present study differed from the previous one in the following respects: there was one (male) experimenter (instead of one male and one female) and a different pain stimulation apparatus was used (because of problems with the Neuroton)³.

The attempt at replication failed. Up-conditioning of neither pain report nor SCR could be re-established; nor did down-conditioning of the pain report take place. At least two issues need to be deliberated upon: how the results of the previous study should be interpreted⁴ and whether there are reasons the hypothesized effects were not found.

There are profound consequences with respect to the interpretation of the results of the study reported in Chapter 3. It is very likely that the up-conditioning group's VAS pain-intensity curve in that study does not in fact represent an operant conditioning effect. A more plausible explanation of this VAS effect, in terms of an attention shift towards pain, has to be sought. This is probably caused by the concentration-like punishments. The same arguments can be applied for the SCR effect. The content of the verbal reactions in the conditioning procedure seems to be of crucial importance. The failure in replication can most likely be accounted for by the difference in the content of the punishing remarks. Since the Linton and

³It is highly unlikely, however, that these two factors are responsible for the discrepancy between the present and previous results. In the previous experiment, analyses showed no difference between the two experimenters with regard to the degree of 'conditioning' effects. As far as the different pain stimulator is concerned, the subjective experience from the pain stimulation apparatus was highly comparable to that of the previous apparatus (this was tested by the experimenter and colleagues). Moreover, the VAS-curve of the control group in the earlier study was very similar to the present one. As a final argument it can be stated that, irrespective of the type of pain-stimulation, comparable conditioning curves have to be found.

⁴Where a replication fails, investigations should be carried out to see whether mistakes, flaws and/or other problems in the replication study can be found to explain the failure. There are no indications here that the failed replication is due to methodological and/or other factors.

Götestam article does not provide detailed information on this specific point, no conclusions can be made on the implications of the present findings for the results of the Linton and Götestam study (1985).

A further question raised by the results of the present study, is whether, using this experimental design, pain report and pain related psychophysiological parameters can be operantly conditioned. Before deciding the answer, we asked ourselves why (other than for reasons connected with the concentration issue) operant conditioning of both pain report and psychophysiological responses did not take place. One of the suggestions, which mainly stemmed from observations during the experimental sessions, related to the existence of sub-groups. It might be that, in this design, there were differences in the degree to which the subjects could be conditioned⁵. In other words, it may be that the conditioning procedure produced the desired effect relatively easily in some subjects, but not in others. This idea tied in with some laboratory observations made by both the experimenter and the assistant. One of the questions in the debriefing session (with subjects in the conditioning groups) was: "What did you think of the remarks made by the experimenter during the series of trials?". There were some subjects who actually said, "A good experimenter doesn't say things like that because it influences the scoring; I was not bothered by that". In the experimenter's opinion, these subjects displayed a self-assured, critical attitude. On the other hand, the subjects who tended during the experiment to respond in the direction conditioning was aimed at (especially in the up-conditioning group), generally made a somewhat timid, shy and docile impression. The subjects were told in the debriefing session that all stimuli were of equal intensity, and yet some of the subjects of the type just mentioned said that they "were happy with the experimenter's encouraging and positive remarks and that they really felt more pain".

With regard to the conditioning groups' VAS curves, this means that some subjects' conditioning effects might be averaged out. Analyses were carried out to investigate these *post-hoc* statements, applying the following notion: It is known that, when submitted to a series of painful shocks (without any attempt at conditioning), some subjects habituate, some report unchanged scores and others sensitize (in fact, this was found in the control group in Linton's study). In order to perform the *post-hoc* analyses the two experimental groups and the control group were split up into three sub-groups. The first sub-group was formed on the basis of more or less habituating scores (the 'habitulators'); the second, on the basis of stable, unchanged scores (the 'stabilizers'); and the third, on the basis of increases

⁵In the authors' opinion, 'degree of conditionability' is equal in everyone when the rewards and punishments are associated with biologically significant stimuli. It may be that verbal approval or disapproval of behavior have less potency in a conditioning procedure, and therefore cause differences in conditionability. Possibly money would have a more potent effect in studies of this kind.

in VAS-pain scores (the 'sensitizers')⁶. If operant conditioning succeeded in a sub-group of subjects, then, on the basis of the just mentioned creation of sub-groups, one would expect

1. the mean pain rating score of the 'sensitizing' sub-group of the up-conditioning group to be higher than those of the 'sensitizers' in the other two groups;
2. the mean pain rating score of the 'habituating' sub-group of the down-conditioning group to be lower than those of the habituators in the other two groups.

Similar hypotheses can be formulated with respect to the effects relating the SCR and magnitude matching data.

The results of the post-hoc analyses are reported here in summarized form (for a detailed discussion see Chapter 9). There are indications that up-conditioning of the VAS pain report took place in a sub-group. These VAS results were not accompanied by significant effects on the magnitude matching or SCR. Moreover, subsequent analyses (including of the tension-change scores) provided convincing evidence that, at least in this experimental design, the SCRs have to be interpreted as mainly reflecting states of arousal associated with tension. The correlation between SCR and VAS-pain intensity did not reach significance after correction for the effect of tension-change scores. In contrast to the conditioning effect in the up-conditioning sensitizing sub-group, the predicted stronger habituating curve in the down-conditioning habituator-sub-group was not found. It has to be concluded, therefore, that there is insufficient evidence to show a down-conditioning effect in this specific sub-group.

A last post-hoc finding relates to the effects of rewards and punishments on VAS-scores. The increasing VAS-scores of the sensitizers in the up-conditioning group were principally determined by the effects of punishments. An asymmetric effect was observed in the habituator sub-group of the down-conditioning group. The decreasing VAS-scores of this group could be explained chiefly as a function of reactions to punishments. In other words, the VAS score patterns of the subjects who responded in the predicted direction (whether increases or decreases!), are principally determined by punishing remarks. This specific observation provides further evidence that the VAS-curves do not reflect an effect of conditioning of tension.

As stated above, the post-hoc notion that conditioning might have taken place in a sub-group was tested. However, since the data-set used for the analyses of the

⁶Re the criterion used to define subgroups: the 33% split was decided upon because of the overall relatively small sample size. In actual fact, the lowest 33% of the decreasing VAS profiles from each experimental condition were assigned to the 'habituating' sub-groups and the highest 33% of the increasing VAS profiles were assigned to the 'sensitizing' groups sub-groups. The remaining 33% were assigned to the 'stabilizing' groups sub-groups.

post-hoc hypotheses was the same as that used for testing the original hypotheses, and given the small sample sizes of the sub-groups, the results of these analyses have to be viewed with extreme caution. They may serve as guidelines or means of generating further hypotheses. One post-hoc finding, however, warrants further research. The SCR-effect, as observed in the present study, is probably less related to pain experience than presumed. An interpretation in terms of an 'arousal-indicator' (feeling tense) is probably more likely. This finding serves as an argument for the use of other psychophysiological variables (such as the evoked potential P300 signal; e.g., Wright and Davies, 1989; for an overview see Chapman et al., 1979) which are believed to be associated with the pain experience.

■ In conclusion, the results previously reported in the Lousberg et al. study (1993a) suggesting evidence for the operant conditioning of pain-report and pain related psychophysiological and psychological variables have to be re-interpreted in terms of attention/tension effects. Based on post-hoc analyses, up-conditioning of pain report may have taken place in a sub-group of subjects. Evidence that this effect is associated with psychophysiological and psychological pain related parameters is, however, marginal. In addition, the post-hoc analyses suggested that 'successfully conditioned subjects' in particular, responded to punishing remarks. No down-conditioning of the pain report could be demonstrated in the present study. As a final conclusion, it would seem somewhat premature at the moment to say that, using the present design, it is impossible to condition the experience of pain in an operant way.

CHAPTER 5

PSYCHOMETRIC PROPERTIES OF THE MULTIDIMENSIONAL PAIN INVENTORY - DUTCH LANGUAGE VERSION (MPI-DLV)¹

5.1 Summary

The purpose of this study was to investigate the psychometric properties of a Dutch translation of the Multidimensional Pain Inventory, MPI-DLV. Data was available on 733 chronic pain patients. There were three issues of special interest. The first one related to the comparability between the MPI-DLV and the American and German MPI versions with regard to the psychometric aspects. The second dealt with the construct validity of the MPI-DLV scale 'General Activity'. It was predicted that patients with high scores on this scale would be in better physical condition, as measured on a working-to-tolerance bicycle ergometer test. In relation to the third issue, attention was given to the factor-invariance between fibromyalgia patients and back pain patients. From the results obtained it was concluded that 1) results pertaining to the factorial structure of MPI-DLV scales, and to reliability and validity, are similar to those from the American and German versions; 2) patients with high scores on the 'General Activity' scale are in better physical condition and 3) MPI-DLVs of fibromyalgia and back pain patients do have similar factor-structures. Evidence was also obtained that the MPI-DLV is sensitive to treatment changes. Applications of the MPI-DLV are discussed.

5.2 Introduction

The self-report questionnaire is one of the instruments often used in the assessment of chronic pain. To date, several questionnaires have been published, each emphasizing different aspects of chronic pain, and each stemming from different theoretical perspectives. For example, the McGill Pain Questionnaire, MPQ, was developed from the perspective of the 'gate control' theory of chronic pain (Melzack, 1975). Another example is the Cognitive Error Questionnaire, CEQ, developed from a cognitively oriented view of pain (Lefebvre, 1981).

In recent years it has been argued that the assessment of chronic pain has to be made within a multidimensional / multiaxial framework (e.g., Duncan et al., 1978;

¹Submitted for publication by Lousberg, R., Groenman, N.H., Schmidt, A.J.M., Amtz, A. & Winter, F.A.M. (1993c).

Feuerstein et al., 1987; Naliboff et al., 1985; Syrjala and Chapman, 1984). According to Turk and Rudy (1986), this implies that pain-relevant information has to be gathered on three axes, i.e., Multiaxial Assessment of Pain, MAP-axes. Somatic-medical aspects are measured on the first axis. These would include, for example, information resulting from neurological examination, examination of muscular function (tone, mass and strength), assessment of spinal mobility, etc. Rudy et al. (1990) described a system which allows for the quantification of biomedical findings on chronic pain patients. On the second MAP-axis - the 'psycho-social axis'- pain-relevant information is obtained on aspects such as subjectively experienced pain-intensity, interference in daily life occasioned by the pain, perceived (pain) control, etc. The third axis pinpoints the quantification of pain behavior such as taking medication, activity level, the responses to pain behavior by the patient's partner, etc.

Using this MAP-reference framework as a basis and applying cognitive-behavioral notions to chronic pain, Kerns et al. (1985) developed the (West Haven Yale) Multidimensional Pain Inventory, (WHY)MPI, which measures several of the pain-relevant aspects on MAP-axes two and three. A central theme in cognitive-behavioral formulations of chronic pain is the emphasis which is placed on the subjectively experienced distress caused by pain, and its consequences for various aspects of the patient's life. The MPI contains 61 items (52 supplemented by 9 experimental, unscaled items), divided into three parts. MPI-part one, which measures pain-relevant psychosocial aspects (MAP-axis 2), is made up of 5 scales: Pain Severity (PS), Interference with the daily life due to pain (I), perceived Life Control (LC), Affective Distress (AD) and social Support (S). In the MPI-parts two and three, some behavioral aspects (MAP-axis 3) are measured. In MPI-part two, the responses of the patient's partner to the pain (as perceived by the patient himself) are mapped out, resulting in three scales: Punishing Responses (PR), Solicitous Responses (SR) and Distracting Responses (DR). In MPI-part three, which is composed of 4 scales, an inventory is made of the frequency of common daily activities labelled as Household Chores (HC), Outdoor Work (OW), Activities Away From Home (AAFH) and Social Activities (SA), which together form the General Activity (GA) level². The psychometric qualities of the MPI have been shown to be sound (Kerns et al., 1985) as is illustrated by the excellent results obtained from reliability and validity analyses. Furthermore, evidence is reported that the MPI is sensitive to treatment changes.

Recently, Flor et al. (1990b) carried out confirmatory factor analyses of a German version of the MPI (MPI-D). Apart from some small changes, the factor structure of parts one and two could be replicated. However, the structure of part three was different in the German version. Instead of there being four factors, it was found that the MPI-D part three consisted of three factors (Household Chores,

²The three MAP-axes should not be confused with the three MPI-parts. MPI-part one measures some psycho-social aspects from MAP-axis two. In MPI-parts two and three, information on pain behavior is obtained which is related to MAP-axis three.

Activities Away From Home and a combined factor Social and Recreational Activities). Results obtained from reliability and validity analyses were, however, comparable to those of the American version.

The present article reports on:

- 1.0) The development of a Dutch translation of the MPI, attempting to replicate the MPI scales configurations, but with three additional points to which special attention has been paid.
- 1.1) The first concerns the assessment of the validity of the MPI-scales. Validity assessment of constructs from self-report questionnaires is often carried out by computing correlations with similar constructs as measured in other, validated, questionnaires. The administration of the other questionnaires usually takes place at the same time. Hence the word 'concurrent' is often used for this kind of validity (Nunnally, 1978). In addition to this concurrent method of validation, it was decided to conduct an experiment to investigate the construct validity of the General Activity scale. Assuming that the performance level as measured on a working-to-tolerance bicycle ergometer test, can be conceived as an indicator of physical condition, it was hypothesized that scores on the MPI General Activity scale would have a positive correlation with the bicycle performance (in this case maximum Watt) levels.
- 1.2) The second point of interest related to the invariance of the MPI-factor structure between sub-samples of pain patients. The patient samples of the above-mentioned Turk and Flor studies are heterogenous with respect to location/diagnosis of the pain. It is implicitly assumed that the underlying construct being measured is the same for patient sub-groups as distinguished by pain locus. Since it seemed important to verify this assumption, it was decided to compare the factor structure of two sub-groups (only two, because they were large enough for the necessary analysis to be carried out): patients with the diagnosis fibromyalgia and back pain patients.
- 1.3) The third issue related to the comparability between the MPI-DLV and the American and German MPI versions on several psychometric aspects, especially on the inter-scale correlations.

5.3 Method

Subjects

Data was available on 733 pain patients who completed the MPI-DLV. Three patient sources were drawn upon:

- 1) 444 patients at the rehabilitation center 't Roessingh in Enschede (110 men, 334 women). These patients completed the MPI-DLV during the assessment phase. They were then offered an intermittent inpatient cognitive-behavioral treatment program. (For a detailed description of this treatment modality, see

- phase. They were then offered an intermittent inpatient cognitive-behavioral treatment program. (For a detailed description of this treatment modality, see Winter, 1992). The largest group of these patients (65.2%) had the diagnosis fibromyalgia; 16.6% were diagnosed as back pain patients.
- 2) 142 fibromyalgia patients (all women) from the Rheumatology Department of the University Hospital Maastricht (Santen-Hoeufft van, 1992) who were participating in a randomized controlled trial. After a psychological assessment phase, during which the MPI-DLV was completed, these patients were offered an outpatient treatment program.
 - 3) 147 patients from the Department of Medical Psychology AZM (52 men, 95 women). These patients were consecutive referrals from the pain clinic of the University Hospital Maastricht.

Altogether there were 162 men and 571 women. The mean age was 42.5 years (s.d.=9.7). The mean duration of pain was 10.5 years (s.d.=9.3). 58.8% of the patients had the diagnosis fibromyalgia, 18.6% were low back pain patients. A small number of patients reported having pain in the head, face and mouth region (3.2%), cervical region (1.5%), upper shoulders and upper limbs (2.1%). In 107 patients (14.7%) the pain was not located in one specific area of the body, nor could a specific diagnosis be made. 38.5 % of the patients were receiving a disability allowance.

The 61 items of the Multidimensional Pain Inventory were translated into Dutch by the first two authors. The contents and translation quality were assessed by a translator. This resulted in the Multidimensional Pain Inventory - Dutch Language Version (MPI-DLV).

Statistical analyses

With regard to aim 1.0 above, confirmatory factor analyses were carried out using the following statistical software programs: PRELIS 1.13, LISREL 7.16 (Jöreskog and Sörbom, 1988) and Simultaneous Components Analyses SCA (Kiers and Berge ten, 1989). PRELIS was used to compute the correlation matrices of each part of the MPI-DLV. These served as input matrices for the LISREL analyses. The fit between the hypothesized model and the estimated model was investigated by considering the following LISREL output aspects in particular: A) The goodness of fit index - a measure which indicates how well the items fit into the scales. B) A chi-square test - a measure which tests for significance between the observed (input) co-variance matrix and the fitted co-variance matrix using the LISREL model. When the sample size is large, as was the case here, small differences between the observed and the estimated co-variance matrices become significant. A corrected chi-square value (chi-square divided by the number of degrees of freedom; March and Hocevar, 1985) was therefore calculated and interpreted. Corrected chi-square values which are lower than 5, are assumed to indicate a reasonable fit. C) The significance of T-tests for the Lambda estimates. Relationships between scales (as is indicated in the estimated phi-matrix) and the fitted

Two kinds of analyses were carried out using the software program SCA. One option on the program allows Multiple Group Method analyses to be carried out (see also Nunnally, 1978). Confirmatory evidence can be obtained from this analysis. In short, the structure of a theoretical model (in this study the scale configurations of the MPI) is represented by means of a weight matrix (the weights of all items are either 0, 1 or -1). The amount of variance accounted for by the scales (components) is of particular interest. If these components do account for a great part of the variance, this set of components summarizes the information well. The amount of variance explained by this hypothesized structure was then compared to the amount of variance explained by a separate Principal Component Analysis (PCA). We also examined 1) the factor-loading matrices to detect items which had high correlations on more than one factor and 2) the factor-correlation matrix.

The program was also used to examine the factor-invariance issue. Separate PCAs were carried out on two patient sub-samples (fibromyalgia and back pain patients). The degree of similarity between the two factor patterns was assessed using Tucker's phi coefficients (Tucker, 1951).

Reliability assessment

The internal consistency of the scales was measured by computing Cronbach's alpha coefficients (Cronbach, 1951). In order to assess test-retest stability, the MPI-DLV was completed twice by a sub-group of 79 patients who had participated in the Roessingh cognitive-behavioral intermittent inpatient treatment program. The MPI-DLVs were mailed to these 79 patients, with a two-week interval between mailings, one year after the patients had finished the treatment program. If, after a while, the lists had not been returned, a telephone call was made to the patient. The final response was 75% (59 patients). Of the remaining patients, some could not be traced because they had moved house and two had died.

Scale validity

Correlation coefficients were computed between the MPI-DLV scales and scales from the following questionnaires validated in the Netherlands:

1. The Dutch translation (Arrindel and Ettema, 1986) of the Symptom Check List 90, SCL-90 (Derogatis, 1983). In essence this list measures factors comparable to those in the English SCL-90 version. It was expected that the factors 'depression' and 'anxiety' would show a positive correlation to the MPI-DLV Affective Distress scale.
2. An interactional problem solving questionnaire, IPOV (Lange, 1983). This list investigates the degree to which the patient is able to reach a satisfactory solution to problems with his partner. It was expected that positive correlations with the scales Support, Solicitous and Distracting Responses, and negative correlations with the Punishing Responses scale would be found.

3. The Dutch translation (Kloot van der and Vertommen, 1989) of the McGill Pain Questionnaire (Melzack, 1975). The total pain rating index was expected to correlate positively with the Pain Severity scale.
4. The Pain Cognition Questionnaire, PCL (Vlaeyen et al., 1989a). This questionnaire measures 5 different pain-cognitions. It was hypothesized that the NZ-scale (Negative Self-efficacy) and the CA-scale (Catastrophizing) would correlate positively with the Interference scale and negatively with the Life Control scale.
5. The Utrecht Coping List, UCL (Schreurs et al., 1988). This list assesses several coping strategies related to general life problems. It was hypothesized that there would be a positive correlation between the 'depressive-reaction pattern' scale and the Affective Distress scale.

Validation data was available on a sub-sample of patients from the Department of Medical Psychology of University Hospital Maastricht. Since the composition of the questionnaire packages was not identical for each patient, the correlations between the MPI-DLV scales and the scales from the other questionnaires were based on different numbers (all questionnaires $N > 80$ except for the UCL which was completed by 13 patients).

With regard to the construct validity of MPI-DLV part three, data was available on 137 female patients (from the fibromyalgia project). A bicycle ergometer (type Jaeger, Würzburg Germany) was used. A working-to-tolerance test was carried out on each patient. The intensity-level was raised by 10 Watts. at minute intervals. The patients' maximum performance levels (expressed in Watts) were noted.

Change sensitivity to treatment

The same patient sample used to assess test-retest stability was employed to determine the MPI-DLV's sensitivity to change. Pre- and post-treatment and (one year) follow-up MPI-DLV and SCL-90 data was available on 55 patients³.

5.4 Results

A large number of items in all three MPI-DLV parts showed a skewed distribution. In order to carry out LISREL analyses, all items were transformed into dichot-

³Multivariate repeated analyses of variance were required to test the MPI-DLV change sensitivity. In this type of analyses a case is left out if one of the variables is missing at one of the three points in time. There was incomplete data in four cases, so the analyses were carried out on the basis of 55, instead of 59 patients.

omous items using the median-split criterion⁴. The results are presented in the order of the aims listed above.

1.0 Development of the MPI-DLV: Confirmatory factor analyses

MPI-DLV part one

There was a large proportion of missing data (60%) on one question (Item 19) on the Interference scale relating to work issues. This item was therefore omitted from the confirmatory analyses.

The goodness of fit index was reasonable (0.934). The adjusted chi-square (2.45) was lower than 5. The SCA analysis showed a fairly high percentage of explained variance (60.97% vs. 62.91% from the PCA). These results, therefore, suggested a confirmation of the structure of part one.

A more detailed examination of the results revealed that the correlations between the scales Pain Severity (PS) and Interference (I) and also between the scales Life Control (LC) and Affective Distress (AD) were relatively high (0.58 and -0.54). However, correlations corrected for attenuation between these two pairs of scales (0.69 and 0.81 respectively) indicated that the scales accounted for substantial unique parts of variance. Because of this statistical argument, and also for theoretical reasons (see the discussion section), it was decided to retain a 5-factor model.

Regarding the inter-correlation between the factors PS and I, the factor loading matrix indicated that the Interference scale item "In general, how much does your pain interfere with your day-to-day activities" (item 2) had high loadings on both scales (PS-scale: 0.648; I-scale: 0.686). Comparable findings were observed on item 16 ("How much suffering do you experience because of your pain?") on the PS-scale. The loading on the PS-scale was 0.648, on the I-scale 0.619. In addition, a reliability analysis revealed that if item 16 were omitted from the PS factor, the Cronbach's alpha coefficient increased remarkably (from 0.74 to 0.81). With reference to the correlation between the scales LC and AD, all items had high loadings on both scales. A reliability analysis of the LC-scale made it clear that item 22 ("How much control do you feel that you have over your pain?") had a strong negative effect on the internal consistency of this scale.

Consequently, new LISREL and SCA analyses were carried out to test a modified model, i.e. the original structure of MPI-DLV part one with three items (2, 16 and 22) omitted. The resulting goodness of fit index was fair (0.953). The adjusted chi-square value (2.14) was somewhat lower than the previous one. SCA analyses of the modified model showed that it explained a high percentage of variance (63.38%) compared to that accounted for by a 5-factor principal component analysis (64.42%). The exclusion of items 2 and 16 resulted in a clear

⁴Post hoc, the same confirmatory analyses were carried out on the non-transformed data. The results obtained from the analyses of transformed and non-transformed data showed such marginal differences that the solutions could be regarded as identical.

decrease in the correlation between the Pain Severity and Interference scales ($r=0.43$ vs. 0.58). These modifications, therefore, improved the results substantially. The 20 items retained, and their factor loadings, are shown in table 5.1.

Table 5.1: Item composition of MPI-DLV part 1 with factor loadings

Scale-names and items	Factor loading
Pain Severity	
- Rate the level of your pain at the present moment.	0.92
- On the average, how severe has your pain been during the <i>last week</i> ?	0.92
Interference	
- Since the time your pain began, how much has your pain changed your ability to work?	0.66
- How much has your pain changed the amount of satisfaction or enjoyment you get from participating in social and recreational activities?	0.76
- How much has your pain changed your ability to participate in recreational and other social activities?	0.80
- How much do you limit your activities in order to keep your pain from getting worse?	0.58
- How much has your pain changed the amount of satisfaction or enjoyment you get from family-related activities?	0.78
- How much has your pain changed your relationship with your spouse, family, or significant other?	0.60
- How much has your pain changed your ability to do household chores?	0.60
- How much has your pain interfered with your ability to plan abilities?	0.73
- How much has your pain changed or interfered your friendships with people other than your family?	0.68
Life Control	
- During the <i>past week</i> how much control do you feel that you have had over your life?	0.81
- During the <i>past week</i> how much control do you feel that you've been able to deal with your problems?	0.87
- During the <i>past week</i> how successful were you in coping with stressful situations in your life?	0.83

Table 5.1 (Continued)

Scale-names and items	Factor loading
Affective Distress	
- Rate the overall mood during the <i>past week</i> .	0.81
- During the <i>past week</i> how irritable have you been?	0.84
- During the <i>past week</i> how tense or anxious have you been?	0.80
Support	
- How supportive or helpful is your spouse (significant other) to you in relation to your pain?	0.87
- How worried is your spouse (significant other) about you because of your pain?	0.90
- How attentive is your spouse (significant other) to you because of your pain?	0.90

MPI-DLV part two

From the LISREL analysis it became clear that the items fitted very well into the scales (goodness of fit index = 0.950). The corrected chi-square value was 2.54. The SCA results showed that the hypothesized model explained 59.97% of the variance, compared to the 60.59% that a three factor PCA accounted for. Reliability analyses showed that the items "Ignores me" and "Reads to me" should be deleted from the Punishing Responses (PR) scale and Distracting Responses (DR) scale respectively. The adjusted model (the original one with these two items left out) showed slightly different results: the goodness of fit index was 0.967, the corrected chi-square was 2.61, and the percentage variance accounted for was 60.47%. The correlations between the three factors were low to moderate (see table 5.4). Table 5.2 presents the 12 items and their factor loadings.

MPI-DLV part three

Both LISREL and SCA analyses revealed problems with the hypothesized structure of part three. This was due to a high correlation between the factors Activities Away From Home and Social Activities. A PCA was therefore carried out. The scree criterion suggested a 3-factor solution in which the factors Household Chores and Outdoor Work were retained, and the items of the factors Social Activities (SA) and Activities Away From Home (AAFH) were merged into the SA/AAFH factor. A LISREL analysis to test this new structure produced good results. The goodness of fit index was 0.932 and the corrected chi-square was 2.53. The percentage of variance explained by this model was 50.81% compared to 51.37%

explained by a three factor PCA. The correlations between the 3 factors were moderate (see table 5.4). The mean score of these three factors taken together forms the General Activity (GA) level. The 18 items and their factor loadings are shown in table 5.3.

Table 5.2: Item composition of MPI-DLV part 2 with factor loadings

Scale-names and items	Factor loading
Punishing Responses	
- Expresses irritation at me.	0.89
- Expresses frustration at me.	0.88
- Expresses anger at me.	0.83
Solicitous Responses	
- Asks me what he/she can do to help.	0.74
- Takes over my jobs or duties.	0.78
- Tries to get me to rest.	0.72
- Gets me some pain medications.	0.51
- Gets me something to eat or drink.	0.75
- Turns on the T.V. to take my mind off my pain.	0.64
Distracting Responses	
- Talks to me about something else to take my mind off the pain.	0.73
- Tries to involve me in some activity.	0.84
- Encourages me to work on a hobby.	0.81

Reliability and scale inter-correlations

The internal consistency of the MPI-DLV scales was assessed by computing Cronbach's alpha coefficients. The test-retest stability of the MPI-DLV scales was assessed using Pearson correlation coefficients. The results, shown in table 5.4, indicate that the internal consistency of the scales was good (alpha's ranging from 0.74 to 0.94). The stability coefficients were also satisfactory.

Validity of MPI-DLV sub-scales

Part one: The PS-scale correlated significantly with the MPQ total pain rating index ($r=0.334$, $p=0.003$). A significant correlation ($r=0.606$, $p<0.001$) was also found between the I-scale and negative self-efficacy (PCL-NZ). The LC-scale had strong negative correlations with the PCL-NZ ($r=-0.661$, $p<0.001$) and the PCL-CA (catastrophizing) scale ($r=-0.577$, $p<0.001$). There were significant correlations between the AD-scale and the SCL-90 depression scale, the SCL-90

anxiety scale and the UCL depressive reaction pattern scale ($r=0.633$; $r=0.587$; $r=0.553$; all $P_s < 0.025$). The Support (S) scale correlated significantly with the IPOV (interpersonal problem solving) total score ($r=0.46$ $p < 0.001$).

Part two: The PR-scale showed a negative correlation with the IPOV total score ($r=-0.60$, $p < 0.001$); The SR and DR scales showed positive relationships with the IPOV total score ($r=0.38$; $r=0.31$; $P_s < 0.002$).

The hypotheses on the relationships between the MPI-DLV scales and scales from other questionnaires could therefore be confirmed.

Table 5.3: Item composition of MPI-DLV part 3 with factor loadings

Scale-names and items	Factor loading
Household Chores	
- Wash dishes.	0.67
- Go grocery shopping.	0.59
- Help with the house cleaning.	0.80
- Prepare a meal.	0.82
- Do the laundry.	0.83
Outdoor Work	
- Mow the lawn.	0.75
- Work in the garden.	0.73
- Work on the car.	0.74
- Wash the car.	0.77
- Work on a needed household repair.	0.69
Social Activities / Activities Away From Home	
- Go out to eat.	0.62
- Play cards or other games.	0.43
- Go to a movie.	0.48
- Visit friends.	0.72
- Take a ride in a car or bus.	0.66
- Visit relatives.	0.69
- Take a trip.	0.83
- Go to a park or beach.	0.63

Table 5.4: Reliability of the MPI-DLV scales

Scale-names	Cronbach's alpha	Test-retest correlation
Pain Severity	0.81	0.81
Interference	0.86	0.88
Life Control	0.78	0.78
Affective Distress	0.74	0.74
Support	0.87	0.88
Punishing Responses	0.76	0.84
Solicitous Responses	0.83	0.78
Distracting Responses	0.82	0.72
Household Chores	0.94	0.81
Outdoor work	0.85	0.79
Social Activities / Activities Away From Home	0.93	0.79
General Activity	0.87	0.78

Sensitivity to change

Pre- and post-treatment and one year follow-up data, which could be used to assess change sensitivity, was available on 55 patients who had participated in the above-mentioned intermittent inpatient cognitive-behavioral treatment protocol (Winter, 1992). Changes on a validated questionnaire (the SCL-90) showed the treatment to have been effective. There were significant decreases on the scales anxiety, agoraphobia, depression, somatic complaints and sleep problems (all P s < 0.05). The SCL-90 total score, the psycho-neuroticism score, also showed a significant decrease ($p < 0.01$).

Table 5.5 presents the mean MPI-DLV scale scores and standard deviations at the three points of measurement: pre- and post-treatment and one year follow-up. Multivariate repeated analyses of variance showed a highly significant effect (multivariate $F(18,37)=3.68$, $p < 0.001$) in the predicted direction. The significant decrease observed on the Pain Severity scale between pre- and post-treatment is interesting and worth noting. These changes concur with the changes found on the SCL-90. It can therefore be concluded from this data that the MPI-DLV is sensitive to treatment changes.

1.1 Construct validity of the General Activity scale

The three scales on MPI-DLV part three form the General Activity scale. Scores on this scale showed a significant positive correlation with the maximum Watt level

as measured at the end of the bicycle ergometer test ($N=137$, $r=0.271$, $p=0.001$). Thus, patients who report being frequently engaged in daily activities such as household chores, social activities, etc., do reach a higher performance level on the bicycle ergometer task.

Table 5.5 Means and standard deviations on the MPI-DLV scales at pre-treatment, post-treatment and one year follow-up

	Pre		Post		Follow-up	
	mean	s.d.	mean	s.d.	mean	s.d.
Pain Intensity	4.50	0.88	4.12	1.12	4.32	1.04
Interference	4.52	0.81	3.82	1.21	3.72	0.96
Life Control	3.12	1.29	3.93	1.29	3.86	1.25
Affective Distress	3.24	1.26	2.82	1.30	2.72	1.43
Support	4.78	1.28	4.34	1.32	4.25	1.36
Punishing Responses	1.15	1.16	0.94	1.10	1.04	1.23
Solicitous Responses	3.28	1.08	2.84	1.11	2.55	0.99
Distracting Responses	2.50	1.33	2.49	1.29	1.87	1.14
General Activity	2.61	0.73	2.59	0.63	2.69	0.83

1.2 Back pain and fibromyalgia: a comparison of the factor structure

Two data sets were compared: the first consisted of data on 405 fibromyalgia patients; the second, data on 96 back pain patients. The final factor structure of the three MPI-DLV parts, as just described, formed the basis of the computations. Tucker's congruency-coefficient phi was computed for each scale, thus using the factor-loadings. All phi values were very high (range: 0.966 - 0.989). Also, the correlations between the factors on the two data sets were, to a very large extent, comparable. Thus, it can be stated that the factors on the two data sets would appear to have the same meanings (cf. Berge ten, 1986).

1.3 Scale inter-correlations between the (WHY)MPI, MPI-D and MPI-DLV

Table 5.6 shows the correlations between the MPI-scales on the (WHY)MPI (Kerns et al., 1985), the MPI-D (Flor et al., 1990b) and the MPI-DLV. Analyses to test for statistical differences between pairs of correlations (see McNemar, 1969) were not carried out because of the different sample sizes on which the correlations were based. The parallel between the American, German and Dutch inter-correlations is notable. In 16 cases, the Dutch inter-scale correlation is situated between the American and German ones. The MPI, MPI-D and MPI-DLV show relatively high inter-scale correlations on three pairs of scales (PS-I, LC-AD and S-SR).

Table 5.6 Inter-scale correlations from the American, German and Dutch MPI versions

		I	LC	AD	S	PR	SR	DR	GA
PS	a	0.58	-0.16	0.34	0.05	0.03	0.31	0.05	-0.17
	b	0.68	-0.13	0.33	0.29	0.11	0.17	0.31	-0.05
	c	0.43	-0.19	0.34	0.17	0.08	0.20	0.10	-0.12
I	a		-0.15	0.26	0.58	0.00	0.34	0.10	-0.22
	b		-0.18	0.44	0.34	0.25	0.24	0.34	-0.07
	c		-0.19	0.42	0.23	0.20	0.22	0.14	-0.27
LC	a			-0.52	0.06	-0.14	-0.08	0.11	0.19
	b			-0.52	0.23	-0.24	0.17	0.05	0.09
	c			-0.57	0.09	-0.25	0.02	0.08	0.23
AD	a				-0.03	0.20	0.04	-0.01	-0.10
	b				0.06	0.28	0.01	0.16	0.02
	c				-0.05	0.29	0.08	0.03	-0.19
S	a					-0.38	0.56	0.42	-0.12
	b					-0.23	0.66	0.49	0.06
	c					-0.28	0.67	0.46	-0.08
PR	a						0.04	-0.01	-0.08
	b						-0.29	-0.02	0.05
	c						-0.13	-0.16	-0.16
SR	a							0.49	-0.18
	b							0.40	0.09
	c							0.51	0.17
DR	a								0.08
	b								0.18
	c								0.22

a) American version; b) German version; c) Dutch version.

PS: Pain Severity; I: Interference; LC: Life Control; AD: Affective Distress; S: Support; PR: Punishing Responses; SR: Solicitous Responses; DR: Distracting Responses; GA: General Activity.

5.5 Discussion

Broadly speaking, the present results show that the MPI-DLV has good psychometric properties. In addition, the pattern of scale inter-correlations is highly comparable to those from the American and German MPI-versions. The analyses of the MPI-DLV, especially with regard to the item composition of the scales, revealed some interesting points.

The first point relates to the structure of MPI-DLV part one. As already stated in the section on results, there were statistical as well as theoretical reasons for keeping a five-factor structure. The similarity between the present confirmatory analyses and those described by Flor et al. (1990b) is striking. In their analyses of

the German version, high inter-correlations were found between exactly the same scales: Pain Severity vs. Interference on the one hand and Life Control vs. Affective Distress on the other. The arguments they gave in favor of a five-factor solution of MPI part one, can also be applied to the present study. The first one relates to the clinical diagnostic process. It is a good idea, in this process, to differentiate between the concepts Pain Severity and Interference. Although it is frequently observed that a person experiencing high levels of pain feels very hampered, this does not necessarily hold true for everyone. Indeed it is well known from clinical experience that some chronic pain patients report high levels of pain, but do not seem to be too greatly handicapped by it. If the items on the two scales were merged into one Pain/Interference scale, an inverse relationship between these concepts (which is not expected to occur frequently) would not emerge and valuable diagnostic information would be missed. Similarly, it seems to be important to distinguish between affective (Affective Distress) and cognitive (Life Control) related information.

The second argument which favors the 5-factor structure of part one concerns the empirically derived patient clusters as described in a series of articles by Turk and his colleagues (for an overview see Turk, 1990). Because of the meaningfulness of the cluster types, as proved by validation data (Turk and Rudy, 1988), and the fact that the classifications hold true with diverse patient samples (Turk and Rudy, 1990), an attempt at replication of these findings would seem worthwhile. In order to compare the characteristics of the patient clusters derived from different translations of the MPI, it would be advisable to keep the same number of factors.

The reliability analyses of the part one scales suggested that three items should be removed. For two of these items the decision was also made on theoretical grounds. The item "How much suffering do you experience because of your pain?" from the Pain Severity scale contains, in the Dutch version at least, a more or less evaluative and affective connotation of pain. It is therefore understandable that this item showed high loadings on the Interference scale and, albeit to a somewhat lesser degree, on the Affective Distress scale. The item "How much control do you feel that you have over your pain?", as part of the Life Control scale, asks about the pain coping-strategies a person is able to apply successfully. However, control over 'life-problems' - such as stress, conflicts, etc. - and control over pain, are not necessarily dependent abilities.

The item "In general, how much does your pain interfere with your day-to-day activities?", was taken out because of very high loadings on both the Interference and Pain Severity factor. An explanation based on conceptual grounds could not be found. It could however be the sequence of the items on the list that caused some 'interference'. This particular item (the second one in the list) immediately follows an item from the Pain Severity scale: "Rate the level of pain at the present moment". It could be, therefore, that there was a 'carry-over effect' of item-response.

With respect to the high inter-correlation between the Life Control and Affective Distress scales, it was found that there were no single items that caused extremely high loadings on both factors, which would then have caused the high correlation. Also the removal of the pain control item had no effect on the scale inter-correlation.

Confirmatory analyses of the structure of parts two and three were marked by fewer problems. As indicated by the reliability analyses, two items were deleted in part two. As was also found by Flor et al. (1990b), the item "Ignores me" did not fit well into the Punishing Responses scale. The item "Reads to me" was characterized by a strongly skewed frequency distribution. Almost every patient indicated that their partner did not read to them as a response to their pain. The problem in part three was straightforward. The factors Activities Away From Home and Social Activities were highly inter-correlated. Consequently, the items on these two scales were merged into one combined 'Social Activities and Activities Away From Home' factor. The final three-factor structure of part three had psychometrically sound properties: moderate scale inter-correlations and good internal consistencies of the scales.

The three scales in part three are together labelled the 'General Activity' level. Patients with high scores on this construct reached higher Watt levels on the working-to-tolerance bicycle ergometer test. This information provides the clinician with a more concrete/practical basis for the interpretation of constructs; in this case, the patient's behavior outside the test situation.

The present chapter provides evidence to support the use of the MPI-DLV as a research instrument. More specifically, it was found that the MPI-DLV, as an instrument of measurement, is sensitive to changes brought about by treatment. With regard to the invariance between factor structures of fibromyalgia and back pain patients, the results clearly showed that the factors could be interpreted equally. In this study comparisons of factorial invariance between other 'pain-location' groups could not be investigated, due to the small number of cases. This question could be answered by future research.

With respect to the comparability of the American, German and Dutch MPI-versions, it can be concluded that the psychometric properties of the lists are roughly the same. That is, the scales in these three MPI versions consist broadly of the same items and the correlations between the scales display a strong degree of similarity. Further, the resemblance between the versions on reliability estimates and validity indicators is striking. The inconsistent findings, especially on the structure of part three (the German and Dutch versions consist of three factors, whereas the American version has four factors), are considered to be differences of minor importance. That is, the scales in part three are almost always taken together as the General Activity level and the deletion of some items in parts one and two can be considered as a sharpening of the constructs.

The main aim of the present study was to develop a psychometrically sound version of a Dutch translation of the MPI. This goal has been achieved. The MPI-

DLV can prove his value as an instrument for producing reliable, valid information for diagnostic purposes and for therapy-outcome studies. A third possible application of the Dutch version of the MPI, the ability to distinguish patient subgroups, has yet to be investigated.

MPI-DLV PAIN PATIENT CLUSTERS: AN ATTEMPT AT REPLICATION¹

6.1 Summary

The present article reports on the development of pain patient sub-groups based on a Dutch translation of the Multidimensional Pain Inventory (MPI-DLV), thus attempting to replicate the American MPI profile configurations of three MPI patient sub-groups. The total sample of chronic pain patients (N=663) was split up into two sub-samples (N=331 and N=332). The first sub-sample was used for the replication analyses. Although replication of the MPI profile configurations was achieved, subsequent validation analyses produced contradictory results. A less restrictive analysis was therefore carried out. Results indicated that four patient clusters could be identified. Because three of these patients types had roughly the same profile characteristics as the three types in the American solution, the American MPI clusters names were retained: Dysfunctional, Interpersonally Distressed and Adaptive Coper. In addition, a fourth cluster, designated 'Average' was distinguished. This patient type possesses characteristics of the other three clusters. Evidence is reported on the validity and distinctiveness of the four MPI-DLV patient types. A cross-validation was successfully carried out, based on the second sub-sample. The (dis)similarity with the American findings is discussed. Finally, the importance of experimental research on the MPI-DLV patient types is stressed.

6.2 Introduction

In recent years there has been growing interest in the development of pain patient sub-groups. Sophisticated multivariate statistical techniques have been used in a considerable number of studies. The reason for the development of patient sub-groups is that, if each sub-group can be offered a tailored treatment program, this should lead to better treatment results. Several instruments, in most cases questionnaires, have been used to this end. Using the Minnesota Multiphasic Personality Inventory (MMPI), Bradley et al. (1978) found that four female and three male patient sub-groups - i.e., clusters - could be distinguished. These findings have

¹Submitted for publication by Lousberg, R., Groenman, N.H., Schmidt, A.J.M., Schouten, E.G.W. & Meijer, S. (1993d).

been successfully replicated in several other studies (e.g., Armentrout et al., 1982; Bernstein et al., 1983; Hart, 1984; McGill et al., 1983; Prokop et al., 1980). Other studies have used the Symptom Checklist-90 (SCL-90) to identify sub-groups of pain patients (Butterworth and Deardorff, 1987; Groenman et al., 1993; Jamison et al., 1988; Schwartz and DeGroot, 1983). Based on data derived from a standardized observation protocol assessing pain behavior in low back pain patients, Keefe et al. (1990) reported that four sub-groups could be distinguished. Replication of these four sub-groups was achieved in a second independent sample.

There are, however, some concerns with regard to the aforementioned studies (see also Turk, 1990). The psychometric qualities of the MMPI, as also the interpretation of elevations on clinical scales, have been criticized (by e.g., Watson, 1982). Perhaps the most important argument against the MMPI typology is that the predictive validity of the MMPI-sub-groups is questionable, since the relationship between them and treatment outcome has been demonstrated in only one study (McCreary, 1985) and other studies have failed to demonstrate it (Guck et al., 1988; McGill et al., 1983; Moore et al., 1986).

Furthermore, as is argued by Turk (1990), both the MMPI and SCL-90 studies and the Keefe et al. (1990) study, focus on single factors such as psychopathology and pain behavior. Emphasizing the multidimensional nature of chronic pain, Turk and Rudy used the Multidimensional Pain Inventory (MPI) (Kerns et al., 1985) to classify a group of pain patients who were heterogeneous as to pain-locus and diagnosis (Turk and Rudy, 1988). They found three patient clusters which they designated 'Dysfunctional', 'Interpersonally Distressed' and 'Adaptive Copers'. The Dysfunctional patient is characterized by the experience of a high level of pain-intensity and psychological distress, a great deal of support from his environment and a relatively low activity level. The profile of the Interpersonally Distressed patient, on the other hand, is marked by somewhat lower levels of pain-intensity and psychological distress. These patients report that their partners / significant others are not very supportive of their pain problems. In contrast to the other two patient types, Adaptive Copers report relatively low levels of pain-intensity and psychological distress. They also seem to be the most active. Investigations into the validity of this cluster analysis have produced excellent results. Moreover, it has been possible to rule out potentially confounding elements such as demographic variables and the degree of medical-physical pathology, in that the differences between the clusters could not be explained by these variables. Using multivariate generalized squared distances and Bayesian posterior probabilities (Tatsuoka, 1988) the accuracy of the classification of patients with unknown cluster profiles was assessed (Turk and Rudy, 1988). Only 5 pain patients out of a sample of 100 could not be assigned to one of the three clusters.

Recently, psychometric analyses of the Dutch version of the MPI (MPI-DLV) demonstrated that, apart from some small differences, the factor structure of the MPI had been replicated (Lousberg et al., 1993c). Results of reliability and validity analyses also concurred with the findings on the MPI (Kerns et al., 1985).

The purpose of the present study was to investigate whether the same clusters, as reported by Turk and Rudy (1988), could be distinguished in the Netherlands. Data from other questionnaires was used to investigate the validity of the clusters. In addition, an experiment was carried out: a bicycle ergometer working-to-tolerance test. If replication of the three MPI-clusters had been achieved, Adaptive Copers would exhibit the highest performance level on the bicycle ergometer test.

6.3 Method

Subjects

Data was available on 663 pain patients who had completed the MPI-DLV. For this purpose three sources were used:

- 1) 397 patients at the rehabilitation center 't Roessingh in Enschede. These patients completed the MPI-DLV during the assessment phase. They were then offered a 'semi-clinical' cognitive-behavioral treatment program. (For a detailed description of this 'semi-clinical' treatment modality, see Winter, 1992). The largest group of these patients (66.5%) had the diagnosis fibromyalgia.
- 2) 125 fibromyalgia patients (all women) from the Rheumatology Department of the University Hospital Maastricht (Santen-Hoeufft van, 1992) who were participating in a randomized controlled trial. After a psychological assessment phase, during which the MPI-DLV was completed, an outpatient treatment program was offered to these patients.
- 3) 141 patients from the Department of Medical Psychology of the University Hospital Maastricht. These patients were consecutive referrals from the pain clinic at the same hospital.

Altogether, there were 152 men and 511 women. The mean age was 42.35 years (s.d.=9.5). The mean duration of pain was 11.1 years (s.d.=9.7). 58.7% of the patients had the diagnosis fibromyalgia, 18.2% were low back pain patients. A minor number of patients reported having pain in the head, face and mouth region (3.3%), cervical region (1.5%), upper shoulders and upper limbs (2.1%). In 92 patients (13.9%) the pain was not located in one specific part of the body, nor could a specific diagnosis could made.

Measures

Data from three other questionnaires was used for validation of the clustering solution:

1. The Dutch translation (Kloot van der and Vertommen, 1989) of the McGill Pain Questionnaire (Melzack, 1975). The total pain rating index was used.

2. The Dutch translation (Arrindel and Ettema, 1986) of the SCL-90 (Derogatis, 1983).
3. An interactional problem solving questionnaire, IPOV (Lange, 1983) which investigates the degree to which the patient is able adequately to resolve problems with his partner. This list also measures marital satisfaction.

In addition, bicycle ergometer data was available on 109 female patients (from the fibromyalgia project). A bicycle ergometer (type Jaeger, Würzburg Germany) was used. A working-to-tolerance test was carried out on each patient. The intensity level was increased by 10 Watts at minute intervals. The patients' maximum performance levels (expressed in Watts) were noted.

If a replication of the three cluster-types were achieved, the following hypotheses would hold true:

1. Patients classified as 'Dysfunctional' would have higher scores than patients in the other clusters on the depression sub-scale of the SCL-90, and the MPQ total pain rating index.
2. Patients classified as 'Interpersonally Distressed' would be characterized by the lowest marital satisfaction score on the IPOV.
3. 'Adaptive Copers' would have the lowest scores on the SCL-90 depression sub-scale, and the MPQ total pain rating index. These patients would, in addition, achieve a higher performance level than patients in the other clusters on the bicycle ergometer test².

Due to differences in the content of the psychological test batteries, the number of available tests varies. As a consequence, univariate tests were carried out. The number of cases on which data was available was too small for multivariate tests to be carried out.

6.4 Results

The total sample ($n=663$) was randomly divided into two sub-samples ($N=331$ and $N=332$). There were no differences between the two samples in terms of sex, age and duration of pain (all P s >0.12). There was a marginally significant difference between the two samples with regard to primary IASP pain location ($\chi^2=17.3$, $d.f.=10$, $p=0.07$). A 'confirmatory' K-means cluster-analysis was carried out on the first sub-sample. 'Confirmatory', because a solution with three clusters was requested and the mean MPI T-scores³ of Turk and Rudy's (1988, study 1) three

²These hypotheses are tested one-tailed.

³A T-distribution is marked by a mean of 50 and a standard deviation of 10.

clusters were specified in advance. The second sub-sample was used to investigate the accuracy of patient classification, based on the results from sub-sample one. For the sake of convenience, the T-score cluster means and standard deviations on the MPI scales from the Turk and Rudy study 1 (N=122) are given in Table 6.1. Table 6.2 shows the cluster centers (mean T-scores and standard deviations) of the Dutch MPI K-means clustering. As can be seen, the Dutch profiles are almost identical to those reported by Turk and Rudy. There are some small differences in T-scores (e.g., the Adaptive Coper from Turk's solution has a lower T-score on the Pain Severity scale than the Adaptive Coper from the Dutch solution) but the essential features of the three profiles are present. That is, the Dysfunctional type of patient has the highest score on the scales Pain Severity, Interference and Affective Distress and the lowest score on the scales Life Control and General Activity. The Interpersonally Distressed patient type has the highest score on the Punishing Responses scale and the lowest score on the Support, Solicitous Responses and Distracting Responses scales.

Table 6.1 T-score cluster means and standard deviations on MPI scales (Turk and Rudy, 1988)

	Cluster 1		Cluster 2		Cluster 3	
	mean	s.d.	mean	s.d.	mean	s.d.
Pain Intensity	56.03	7.51	50.17	7.46	41.13	8.73
Interference	55.66	7.98	50.74	7.43	41.13	8.46
Life Control	45.47	9.30	48.92	9.08	57.55	7.18
Affective Distress	55.32	8.95	49.98	8.53	42.32	7.59
Support	53.99	7.88	40.61	9.51	53.09	7.14
Punishing Responses	47.77	7.76	57.48	11.94	46.16	6.83
Solicitous Responses	54.72	7.77	40.12	6.72	52.59	8.98
Distracting Responses	53.03	10.13	42.16	6.99	53.04	8.20
General Activity	44.89	9.68	52.73	7.97	54.79	8.93

Cluster 1: Dysfunctional patients (N=52, 42.6%)

Cluster 2: Interpersonally Distressed patients (N=36, 29.5%)

Cluster 3: Adaptive Coper patients (N=34, 27.9%)

The Adaptive Coper, reports the lowest levels of Pain Severity, Interference and Affective Distress, and the highest levels of Life Control and General Activity. It is also worth noting that the relative frequency of each cluster type concurs with the relative frequencies of the American patient types.

The classification accuracy was then investigated by comparing the K-means cluster analysis profile assignment with the cluster assignment determined by a classification model as outlined by Tatsuoka (1980). A (software) classification program was written for the purposes of assigning individual patients to one of the three clusters. Using the nine MPI-DLV scale scores, multivariate generalized

squared distances and Bayesian posterior probabilities were computed for each cluster. A classification was considered as correct if the cluster with the highest posterior probability accorded with the corresponding K-means cluster. If the cluster with the highest posterior probability did not accord with the corresponding K-means cluster, the case was considered a misclassification.

Table 6.2 T-score cluster means and standard deviations on MPI-DLV scales (with initially specified cluster centers)

	Cluster 1		Cluster 2		Cluster 3	
	mean	s.d.	mean	s.d.	mean	s.d.
Pain Intensity	54.98	8.10	47.76	8.96	45.43	10.45
Interference	55.25	7.83	48.70	8.99	44.08	9.99
Life Control	45.37	9.40	48.30	8.46	58.14	6.90
Affective Distress	55.69	8.23	51.47	8.04	40.63	6.86
Support	54.98	5.89	39.26	8.44	54.20	7.04
Punishing Responses	49.98	9.58	54.96	11.43	44.92	5.49
Solicitous Responses	55.23	8.07	40.86	7.10	52.19	8.46
Distracting Responses	53.59	8.79	42.74	7.94	52.53	9.61
General Activity	46.79	9.58	50.90	7.92	53.50	11.15

Cluster 1: Dysfunctional patients (N=134, 40.5%)

Cluster 2: Interpersonally Distressed patients (N=100, 30.2%)

Cluster 3: Adaptive Copers patients (N=97, 29.3%)

As in the method described by Rudy (1989), the classification procedure could have three results:

1. A patient was assigned to one of the three MPI-DLV cluster prototypes if the posterior probability belonging to a cluster was a) at least 0.66 (2 times the random chance of being allocated to one of the clusters) and b) the chi-square test of their generalized squared distance from the cluster centroid was non-significant at the 0.01 level.
2. A case having three posterior probabilities lower than 0.66 (and at least one non-significant chi-square test) was designated 'Hybrid'. This patient profile embodies characteristics of more than one cluster.
3. A case was designated 'Anomalous' if the distance to all three cluster centers was significant at the 0.01 level. That is, "the patient's set of MPI scores does not fall within the 99% confidence intervals for any of the patient prototypes" (Rudy, 1989).

For the sake of clarity, an Anomalous or a Hybrid classification were considered as correct classifications if the cluster with the highest posterior probability accorded with the K-means cluster assignment.

Ninety-four percent of the cases were classified correctly. The mean posterior probability of patients assigned to the Dysfunctional, Interpersonally Distressed and Adaptive Copers clusters was 0.84 (N=77), 0.86 (N=66) and 0.82 (N=63) respectively. Two cases were classified as anomalous. One hundred and three cases (31.1%) were classified as Hybrid. The percentage of patients classified as Hybrid is well in excess of those reported by Rudy (1989) (4-5%). Twenty cases (6.0%) of the total sample were misclassified.

The next step in this cluster replication process was the validation of the MPI-DLV cluster prototypes using data not used to generate the clustering solution. Using the required contrasts, analyses of variance were carried out to test the hypotheses stated above (the Anomalous, Hybrid and misclassified patients were excluded from the analyses).

Pain severity: In accordance with the hypothesis, the Adaptive Copers had the lowest MPQ total pain rating index ($t=2.04$, $p=0.02$). However, the Dysfunctional patient type had a lower, though not significant ($t=0.20$, $p=0.87$) total pain rating index than the Interpersonally Distressed patient type. Thus, the hypothesis that the Dysfunctional patient is characterized by the highest MPQ total pain rating index could not be confirmed.

Mood: Using the SCL-90 depression score, Adaptive Copers were significantly less depressed than the other two patient types ($t=3.89$, $p<0.001$). In a comparison between the Dysfunctional and the Interpersonally Distressed type, a relationship contrary to expectation again emerged, in that the Interpersonally Distressed type tended to have a somewhat higher depression score than the Dysfunctional type ($t=1.18$, $p=0.122$).

Support: As was hypothesized, patients classified as Interpersonally Distressed had a significantly lower marital satisfaction score than the two other patient types ($t=6.73$, $p<0.001$). However, the Adaptive Copers had a marginally significant higher score than the Dysfunctionals ($t=1.45$, $p<0.08$), a finding which was not in line with cluster T-scores on the support scale (no difference between the Adaptive Copers and Dysfunctional subtype was expected on the Support scale).

These validation analyses revealed that the hypotheses regarding the Adaptive Copers subtype could be confirmed. There were, however, major validation problems regarding the other two types.

In summary, the configurations of the three MPI types could be replicated using the K-means cluster analysis. However, the interpretation and clinical usefulness of the patient clusters are very problematic since the validation analyses were not in accordance with cluster differences on several MPI-DLV-scales.

It was decided to carry out another, this time exploratory, K-means clustering analysis with three clusters requested but without pre-specified cluster centers. The results are shown in Table 6.3.

Cluster 1 patients were shown to have the highest level of pain severity and the pain interferes with their life to a great extent. This patient sub-type is also typified by a high level of support and high levels of solicitous and distracting responses from their partners.

Cluster 2 patients were shown to have the highest level of interpersonal distress, clearly indicated by the highest score on the Punishing Responses scale and the low scores on the Support and Solicitous Response scales. These patients experience high levels of interference with their daily life, as well as affective distress. Further, they appear to have the lowest level of life control.

Table 6.3 T-score cluster means and standard deviations on MPI-DLV scales (without initially specified cluster centers)

	Cluster 1		Cluster 2		Cluster 3	
	mean	s.d.	mean	s.d.	mean	s.d.
Pain Intensity	52.33	8.66	49.45	9.31	40.99	11.57
Interference	50.94	9.32	51.30	8.48	43.28	12.89
Life Control	49.48	9.66	46.25	8.91	59.81	6.92
Affective Distress	50.97	9.69	53.49	8.34	38.76	5.90
Support	54.24	6.97	40.20	9.76	51.43	7.77
Punishing Responses	46.51	6.80	60.23	10.36	44.45	4.95
Solicitous Responses	54.32	8.67	43.06	9.11	45.30	6.64
Distracting Responses	54.34	8.30	43.08	8.38	45.18	9.81
General Activity	49.93	10.44	49.24	8.20	51.86	11.26

Cluster 1: (N=195, 59%)

Cluster 2: (N=91, 27%)

Cluster 3: (N=45, 14%)

The cluster 3 patients possess a MPI-DLV sub-scale score configuration which is marked by a relatively low level of pain severity and interference with daily life. They are far less psychologically distressed (a low Affective Distress scale in combination with a high Life Control scale), and are somewhat more active than the other two patient types.

The three clusters obtained using this clustering solution were, then, highly comparable to those reported by Turk and Rudy (1988). It was decided to keep the same nomenclature for the patient clusters: cluster 1 patients will be designated 'Dysfunctional'; cluster 2 patients, 'Interpersonally Distressed'; and cluster 3 patients, 'Adaptive Coper'.

Classification accuracy was assessed by using an adjusted version of the classification program described above. Only 4 cases (1.2%) of the total sample were misclassified. The mean posterior probability of the MPI-DLV prototypes

Dysfunctional, Interpersonally Distressed and Adaptive Coper, was 0.86 (N=117, 35.3%), 0.89 (N=76, 23.0%) and 0.81 (N=40, 12.1%) respectively. Six cases (1.8%) were classified as Anomalous. The remaining 88 cases (26.6%) were classified as 'Hybrid'. The average posterior probability (the highest probability was taken) of these Hybrid patients was 0.55.

Validation of the clustering

Analyses identical to those of the previous validation process were carried out, again using variables which were not used to generate the K-means cluster solution. Based on the new configuration of the MPI-DLV scales, the following hypotheses were formulated. It was expected that the Interpersonally Distressed patients would be more depressed than the other two subtypes and would have a lower level of marital satisfaction. The Dysfunctional type was expected to have a higher score on the MPQ total pain rating index, and the Adaptive Coper was expected to have the lowest levels on the MPQ total pain rating index scale and the SCL-90 depression sub-scale. The Anomalous, Hybrid and misclassified patients were left out of these analyses.

Pain severity: The Adaptive Coper had a significantly lower total pain rating index than the Dysfunctional ($t=-1.82$, $p=0.04$) and Interpersonally Distressed type ($t=2.91$, $p=0.003$). There was no significant difference between the Dysfunctional and Interpersonally Distressed type.

Mood: Interpersonally Distressed patients had a significantly higher depression score than Dysfunctional patients ($t=2.60$, $p=0.005$). Dysfunctional patients had a significantly higher depression score than Adaptive Coper patients ($t=2.51$, $p=0.007$).

Support: The Interpersonally Distressed patient was shown to have a significantly ($t=6.36$, $p<0.001$) lower marital satisfaction level than the other two patient types, and the Dysfunctional type was shown to have a somewhat (though not significantly) higher score on this scale than the Adaptive Coper type. This is in accordance with the patient profile distribution on the MPI-DLV Support scale.

All but one validation hypotheses could therefore be confirmed. It was concluded that, on the basis of the results so far obtained, the three patient clusters have proved to be unique and distinct.

In discussions relating to the high percentage (26.6%) of patients classified as Hybrid, an investigation was suggested into whether these patients could be viewed as a fourth, unique, cluster (Lousberg, 1992). Firstly, the mean MPI-DLV T-score profile of the Hybrids was computed. Then, using the MPI-DLV scale scores and contrasting the Hybrid with the other three profiles, a multivariate analysis of variance was carried out. The result strongly supported the distinctiveness of the Hybrid profile ($F(\text{Hot})=3.20$, $p<0.0001$). Table 6.4 shows the mean MPI-DLV T-score profile of patients classified as Hybrid.

In general, the MPI-DLV T-score configuration of the Hybrid type is situated somewhere between the other types, except on two scales: the Interference scale

has a high T-score and the General Activity scale a low T-score. Since the T-scores of the Hybrid profile all vary close around 50 (the mean of the T-distribution) and taking into account the other three profile configurations, the term Hybrid was redesignated 'Average'. Figure 6.1 depicts the mean MPI-DLV T-score profiles of the four clusters. In validating the Average type, scores on the MPQ total pain rating index, SCL-90 depression scale and IPOV marital satisfaction score were compared with those from the other clusters.

Table 6.4 T-score means and standard deviations on MPI-DLV scales of the Average (Hybrid) cluster

	Mean	s.d.
Pain Intensity	49.71	7.83
Interference	51.38	8.13
Life Control	49.89	9.55
Affective Distress	50.19	9.29
Support	49.39	7.51
Punishing Responses	46.93	6.66
Solicitous Responses	49.05	7.53
Distracting Responses	47.97	8.21
General Activity	48.80	8.66

The MPQ total pain rating index of the Averages was significantly higher than that of the Adaptive Copers ($t=2.44$, $p=0.02$). There were no significant differences between them and the Dysfunctional and Interpersonally Distressed types (all $P_s > 0.51$). The SCL-90 depression score was significantly higher than that of the Adaptive Coper type ($t=3.03$, $p=0.002$), but not different from the other two types ($p > 0.23$). The IPOV marital satisfaction scale was significantly higher than that of the Interpersonally Distressed type ($t=4.10$, $p=0.0001$) and significantly lower than the Dysfunctional type ($t=-2.33$, $p=0.022$). This data gave further support to the uniqueness and validity of the 'Average' patient profile.

Classification of patients with unknown cluster membership

The first sub-sample (N=331) was used to a) develop the patient clusters, b) test the classification accuracy of patients with known cluster membership and c) investigate the validity of the k-means cluster analysis. Then, making use of the cluster means and co-variance matrices of the first sub-sample, the second half (N=332) of the total sample was used to classify patients with unknown cluster membership. The same classification program and strategy for assigning patients to clusters were again used.

Table 6.5 presents the number, percentage and average posterior probability of the four clusters for sub-samples one and two. Seventeen patients (5.1%) could not

MPI-DLV CLUSTERS

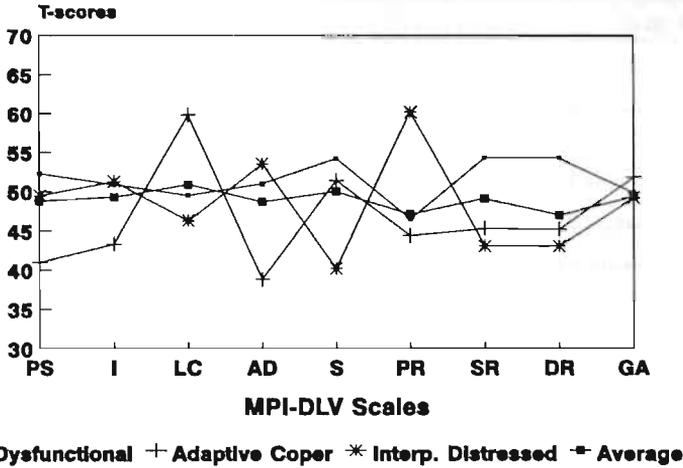


Figure 6.1 Mean T-scores for the patient profiles of the MPI-DLV.

be assigned to one of the four clusters (and were therefore designated Anomalous). It was tested whether there were differences between the two sub-samples with regard to the proportion of classified patients and posterior probability. Regarding the posterior probability, there was neither a significant sub-sample effect ($F(1,627) = 2.08, p = 0.15$) nor a significant sub-sample x cluster effect ($F(3,627) = 0.04, p = 0.78$). Also, the proportion of patients classified to the clusters did not differ significantly ($z_1 = 1.26, z_2 = -0.26, z_3 = 0.50, z_4 = 1.07$; all $P_s > 0.104$). The MPI-DLV scale scores of the clusters were compared in another test of the comparability of the sub-samples. The multivariate test indicated that there was no significant difference between the scores on the MPI-DLV scales ($F(\text{Hot}) = 0.99, p = 0.45$).

Some tests were then carried out to eliminate potential confounds. More specifically, investigations were carried out to determine whether there were differences between the clusters with regard to sex, age, IASP site and duration of pain. There were no differences between the four clusters with regard to sex or IASP-primary pain locus ($\chi^2 = 1.15, \text{d.f.} = 3, p = 0.76$; $\chi^2 = 11.3, \text{d.f.} = 9, p = 0.26$).

There was a significant age effect ($P < 0.05$), showing that Dysfunctional (42.4 years) and Interpersonally Distressed patients (43.8 years) are slightly older than Adaptive Copers and Average patients (both 41.4 years). After correcting for age, the MPI-DLV scale differences between the clusters remained intact ($F(\text{Hot}) = 72.7, p < 0.0001$). The four clusters also differed significantly with regard to duration of pain ($p = 0.05$), indicating a shorter duration in Adaptive Copers patients (6.9 years compared to 10.9, 14.9 and 11.6 years in Dysfunctional, Interpersonally Distressed and Average patients respectively). As was the case with the variable

age, the cluster differences remained significant ($F(\text{Hot}) = 14.8, p < 0.0001$) after correction for duration of pain.

Table 6.5 Number, percentage and mean posterior probability of MPI-DLV clusters (sub-sample 1 en 2)

	N	%	Posterior probability	
			Mean	s.d.
Dysfunctional				
- sub-sample 1	117	35.3	0.86	0.10
- sub-sample 2	102	30.7	0.86	0.11
Interpersonally Distressed				
- sub-sample 1	76	23.0	0.89	0.10
- sub-sample 2	79	23.8	0.89	0.11
Adaptive Coper				
- sub-sample 1	40	12.1	0.81	0.10
- sub-sample 2	36	10.8	0.78	0.08
Average				
- sub-sample 1	88	26.6	0.55	0.07
- sub-sample 2	98	29.5	0.54	0.06

Table 6.6 Mean maximum Watt-level of the MPI-DLV clusters as measured on a walking-to-tolerance bicycle ergometer test

	N	Watt-level	
		mean	s.d.
Dysfunctional	46	117.8	27.4
Interpersonally Distressed	23	121.1	31.6
Adaptive Coper	21	134.5	37.6
Average	19	130.8	41.0

Finally, an attempt was made to validate the MPI-DLV cluster differences using an external measure: maximum Watt level obtained on a working-to-tolerance bicycle ergometer test. Because of a) the differences regarding the variable duration of pain and b) the fact that this variable might interfere with physical performance, it was decided to take duration of pain as a co-variate in the analysis. The results are summarized in Table 6.6. Comparison of the Adaptive Coper profile and the other three together revealed a marginal effect ($t=1.50, p=0.07$). The difference between the Adaptive Coper and Dysfunctional group was significant ($t=1.99, p=0.03$).

6.5 Discussion

The aim of this study was to replicate the MPI cluster findings reported by Turk and Rudy (1988). Although there are some notable differences, the results of the cluster analyses of the Dutch version of the MPI are to a large extent comparable to those of the American version. Some points need further clarification.

There are some differences with regard to sample characteristics. Firstly, the present sample size used to carry out the cluster analysis is very much larger (331 versus 122). Secondly, the present sample seems to be more heterogeneous in terms of diagnosis and primary pain location. More specifically, the Dutch sample consisted of a large number of fibromyalgia patients. Also, in contrast to the Dutch patient sample, all the patients in Turk's study met specific admission criteria for an outpatient pain clinic. Thirdly, the sex-distribution is different in the two studies (79% men in Turk's sample versus 22.3% in the present study). A combination of these differences in sample characteristics, as well as cultural differences, may account for the slight discrepancies between the two cluster solutions.

The most striking and clinically important difference between the Dutch and American cluster solutions is the higher level of psychological distress of the Interpersonally Distressed type in the Dutch sample. This specific MPI-DLV cluster characteristic could, moreover, be confirmed by the validation analyses. Attention to the affective / psychological profile components, in addition to focussing on interpersonal problems, may be central to a treatment program for the Interpersonally Distressed patient.

A fourth cluster, designated 'Average', was distinguished. In Turk's clustering solution these patients were called "Hybrid". In the American sample patients of this type are relatively scarce; 4-5% of the cases (Rudy, 1989). In the present study the amount of patients thus classified was considered large enough (27%) to be treated as a new patient type. The results of the validation analyses added weight to this decision. The term Average should be interpreted in combination with the profile configurations of the other three types. Although the pain does interfere with their lives to a high degree and the general activity level is relatively low, these patients do not suffer psychologically as much as patients classified as Dysfunctional or Interpersonally Distressed. Interestingly, in a study investigating profile characteristics of the four patient types (Lousberg et al., 1993f), it was found that patients of the Average type have a marginally significant higher level of intelligence. It can be hypothesized, consequently, that these patients fill in the MPI-DLV in more a rational, intellectualizing manner. The answers to the questions may be somewhat more nuanced and less extreme than those of the other patient types - hence the resulting 'mean' profile. Yet more research is needed to map out the characteristics and distinctive features of this patient type.

The issue of potential confounds in empirical taxonomies is a crucial point. The development of empirically derived patient clusters based of the MPI scores would be a useless enterprise if the differences in clusters could be entirely explained by key variables such as duration of pain and degree of medical-physical pathology. In

this study it was found that Adaptive Copers were marked by a shorter duration of pain. Consequently, this could be interpreted as the reason these patients report less pain, are more active, etc. However, the MPI-DLV cluster differences were large enough to remain significant after correcting for the duration of pain effect. Interestingly, Turk and Rudy found that Adaptive Copers do have a lower degree of medical-physical pathology. Again, this difference can be viewed as an artefact of the clustering; since Adaptive Copers have a lower degree of pathology, it is logical that they report less pain, etc. But here too, after correcting for medical findings, the "cluster differences could not be accounted for solely by the degree of physical pathology" (Turk and Rudy, 1988). Unfortunately, data regarding medical-physical pathology was not available for this study. Future research should answer the question as to whether 'medical-physical pathology' should be viewed as a serious potential confound in the MPI-DLV taxonomy.

With regard to the concurrent validation of the four MPI-DLV patient profiles, significant correlations with scores from other questionnaires were observed. Also, experimental evidence was obtained indicating that Adaptive Copers do perform (marginally) better in a working-to-tolerance bicycle ergometer test. Compared to the performance level of patients classified as Dysfunctional, this effect was significant.

It can therefore be concluded that (in the Netherlands) pain patients can be assigned to one of four sub-groups. It has been shown that individual patients can be accurately assigned to these sub-groups. Evidence was found to support the validity and uniqueness of the patient sub-groups. Future experimental research, focusing on the distinctiveness and valid interpretation of the MPI(-DLV) patient profiles (cf. Pearce and Morley, 1989), is warranted.

VALIDATING THE MPI-DLV USING EXPERIENCE SAMPLING DATA¹

7.1 Summary

The present study investigates the relationship between scale scores on the Dutch version of the Multidimensional Pain Inventory (MPI-DLV) and data derived from comparable items from an experience sampling procedure. The experience sampling items were constructed in such a way as to resemble as far as possible the key concept of the corresponding MPI-DLV scale. Fifty-seven chronic pain patients participated in the study which lasted six consecutive days. Special attention was given to the relationship between the MPI-DLV pain intensity score and the mean experience sampling pain intensity score. Significant correlations were found between the MPI-DLV scales Pain Severity, Interference, Solicitous, Punishing and Distracting Responses, Household Chores and their experience sampling analogues. A marginally significant correlation was found with regard to the MPI-DLV Life Control scale. The General Activity and Affective Distress scales had no relationship with the analogous experience sampling items. The significant correlations were regarded as a further validation of the MPI-DLV. Reasons for the non-significant correlations are given. A regression analysis revealed that 58% of the variance of the experience sampling pain intensity score could be explained by the MPI-DLV present pain intensity item score. In addition, low scores on the MPI-DLV pain item were accompanied by somewhat higher experience sampling pain item scores and high MPI-DLV pain scores by somewhat lower experience sampling pain scores. The implications of these findings are discussed.

7.2 Introduction

The use of self-report questionnaires in multidisciplinary pain assessment has become increasingly popular, as is evidenced by the rapidly growing number of self-report questionnaires measuring one or more pain related aspects, and the fact that scores derived from questionnaires often form a major part of the effect variables in treatment evaluation studies.

¹Submitted for publication by Lousberg, R., Schmidt, A.J.M., Groenman, N.H., Vendrig, L. & Dijkman, C.I.M. (1993e).

One of the questionnaires often used in pain research is the Multidimensional Pain Inventory (MPI) (Kerns et al., 1985). This list is designed to assess the psychosocial and behavioral aspects of pain. The psychometric properties (reliability and validity) are good. The MPI has recently been translated into German (MPI-D) (Flor et al., 1990b) and into Dutch (MPI-DLV) (Lousberg et al., 1993c). Apart from some small differences, the factor structure of the American version could be replicated in both studies. Also, the reliability and validity analyses of the MPI-D and MPI-DLV have yielded satisfactory results. In addition, evidence is reported that these three MPI versions are sensitive to treatment effects.

Although there are some items which ask for a judgment to be made pertaining to that particular moment (e.g., "Rate the level of pain *at the present moment.*"), the MPI can be considered to be a list which measures relatively stable behavioral and emotional aspects of pain (from the patient's perspective). For example, the MPI-scale 'Interference' measures the degree to which the pain generally interferes with the patient's daily life. In clinical practice, one of the objectives in the assessment phase of a chronic pain problem is, indeed, to get an idea of how a patient functions, thinks and behaves in daily life. However, since the MPI is usually completed in a clinical setting, and since it is more or less stable / 'trait' information that is obtained, inferences regarding the patient's behavior, pain experience, etc., in daily life cannot really be drawn. It is therefore important, from a clinical point of view, to acquire greater certainty as to the generalizability, ecological validity, of the MPI scale scores.

The results of the Flor et al. (1990b) study mentioned above, have shed more light on this issue. Making use of diary data, it was found that the MPI-D Pain Intensity scale had a significantly positive correlation (0.61) with the mean diary pain intensity score (based on one week's data). A significantly positive correlation (0.48) was also found between the MPI-D Interference scale and an analogous question in the diary. Correlations between the MPI-D activity sub-scales and the corresponding diary activity scales did not, however, reach significance, neither did the MPI-D General Activity scale correlate significantly to the diary general activity scale ($r = -0.27$; n.s.). Apart from the MPI-D General Activity (sub)-scale(s), these results can be seen as a further, ecological, validation of the MPI. The Flor et al. (1990b) study has prompted at least two other research issues. Firstly, it would be of value to know the connection between daily records and several other MPI scales. Secondly, the relationship between the MPI-D Pain Severity scale and the mean diary score was expressed as a correlation, indicating the strength of the relationship. It would be useful to obtain a better insight into the **nature** of this relationship.

It was decided to employ an Experience Sampling Method (ESM) to investigate these two issues. ESM is a technique for obtaining a representative sample of moments in a person's daily life. The reliability and validity of this technique have already been demonstrated by Csikszentmihalyi and Larson (1987), and, as was argued by DeVries (1987), this method avoids the problem of general and retro-

spective recall (see also Clark, 1988). Another advantage of ESM is that anticipation at moments of measurement is prevented.

In the present study a watch was used which indicated to the patient (in a random time schedule) when to answer some questions in a booklet. Key items were created to represent / mirror as far as possible the content of the MPI-DLV scales. It was predicted that the mean score of the experience sampling items would have a positive correlation with corresponding MPI-DLV scale scores. Further, a regression analysis was carried out to investigate the nature of the relationship between the present pain estimate item on the MPI-DLV, and the mean experience sampling pain intensity rating.

7.3 Method

Subjects

All patients were referred to the pain-center of the Academic Hospital Maastricht. After a patient was seen by the neurologist he/she was asked to participate in the study, if the following criteria were met:

- pain duration longer than 6 months;
- pain is not caused by cancer;
- age between 18 and 65;
- no alcohol or other drug dependence;
- having a significant other;
- absence of severe psychopathology;
- no major difficulties with reading or writing.

Patients were told the purpose of the experience sampling experiment, and their informed consent was obtained. They were told that the experience sampling experiment would start the next day and would last 6 full days. They were then given instructions on how to handle the watch and what to do with the booklets, etc. They were told that they could turn off the watch's signal function (if, for example, they were in church), but that it was important that the booklets were filled in completely. They were also assured that, should they have any questions or problems, they could phone the interviewer. The patients then completed the MPI-DLV. Finally, they were given 6 booklets (each for one day), the watch and a set of instructions on both the experience sampling questions and the watch. An appointment was made for a debriefing session at which the watch and booklets would be returned. The briefing period took about 30 minutes.

During the debriefing session the booklets were checked for missing data. If there were any problems, these were discussed. Patients were also asked whether, during the six days of the experience sampling procedure, anything had happened which might have had a major influence on the experiment.

Fifty-seven chronic pain patients (26 males and 31 females) satisfied the inclusion criteria and gave their informed consent. The mean age was 42.33 (range 21-64, s.d. = 10.0). The mean duration of pain was 7.1 years (range 1-33, s.d. = 7.5). Table 7.1 presents the IASP primary pain-locus distribution.

Table 7.1 Primary IASP pain locus

	N	%
Head, face and mouth	6	10.5
Cervical region	2	3.5
Upper shoulder and upper limbs	5	8.8
Thoracal region	4	7.0
Abdominal region	2	3.5
Lower back, lumbar spine, sacrum and coccyx	18	31.6
Lower limbs	5	8.8
Pelvic region	1	1.8
Anal, perianal and genital region	1	1.8
More than three major sites	13	22.9

Sampling method

In most cases, the experience sampling experiment started on a Tuesday morning and lasted until Sunday night. A Seiko RC-1000 watch was used to indicate to the subjects when they should fill in/answer the items/questions. The signals (beeps) sounded eight times a day between 08.30 and 22.30, in a random time schedule. In addition, patients had to answer some questions the moment they woke up and just before going to sleep. Responses made more than 10 minutes after a beep were considered invalid.

Experience sampling booklets

The first five scales on the MPI-DLV are Pain Severity, Interference (by the pain with the patient's daily life), Life Control, Affective Distress and Support. The next three scales measure Punishing, Solicitous and Distracting Responses (by the partner to the patient's pain behavior). The last scale, General Activity, measures the frequency of some common daily activities. The General Activity scale is divided into three sub-scales: Household Chores, Outdoor Work and Social Activities / Activities Away From Home.

Table 7.2 shows the ESM items used to mirror the MPI-DLV scales (with the exception of the Support scale). All the items were scored on 7 point Likert scales (0-6). The Likert scales of the first three ESM items ranged from "no pain / interference / control at all" (0) to "a great deal of pain / interference / control" (6). The Likert scales of ESM items representing the partner response scales ranged from "not at all" (0) to "very often" (6). The physical activity item ranged

from "resting, lying down, doing nothing" (0) to "hard physical work" (6). The change in mood item ranged from "normal mood" (0) to "a sharp change" (6). If a patient indicated a change (i.e., a score from 1 to 6), he had to fill in whether this change was perceived as pleasant or unpleasant.

Table 7.2 MPI-DLV scales with their corresponding ESM item

MPI-DLV scale	ESM item
Pain Severity	Pain Severity
Interference	Pain-interference
Life Control	Control on situation
Affective Distress	Change in mood
Punishing Responses	Irritation
Solicitous Responses	Takes over / take work off
Distracting Responses	Attention distraction
General Activity	Physical activity

The item 'change in mood' was chosen to represent the MPI-DLV Affective Distress scale, rather than the more simple and straightforward item 'mood'. This was due to a specific interest in mood changes (reported elsewhere). In order to carry out the planned correlation analysis, the ESM mood change score was transformed as follows: A 0 score (normal mood) was recoded 3. Mood change scores in the 'pleasant' direction were recoded 2.5, 2, 1.5, 1, 0.5 or 0. Mood change scores in the 'unpleasant' direction were recoded 3.5 to 6 (at 0.5 intervals). In other words, the transformed mood change scale ranged from 0 (a very pleasant change) via 3 (normal mood) to 6 (a sharp change in the unpleasant direction). It was predicted that this transformed scale would show a positive correlation with the MPI-DLV Affective Distress scale.

Since the MPI-DLV General Activity scale contains three kinds of activities (Household Chores, Outdoor Work and Social Activities / Activities Away From Home) it could scarcely be represented by one key item. Based on a previous finding (the positively significant correlation between bicycle ergometer performance and the MPI-DLV General Activity scale), it was hypothesized that a measure of actual physical activity level would correlate positively with the MPI-DLV General Activity scale. In order to get more specific information about daily activities, three open questions were added which had to be answered at each signal. These were: What are you doing? Where are you? and Who are you with? These three daily life activity sub-categories were coded. Previous research, using the same coding system had shown that the inter-rater reliability was high (Kappa's > 0.90; Lousberg, 1993). On the basis of the coded data a 'household activity frequency score' was computed for each patient. This score was correlated with the MPI-DLV General Activity sub-scale Household Activities.

Analyses

Analyses were carried out as follows: For each ESM variable a mean score was computed by adding up all non-missing scores and dividing this total by the number of valid observations². Correlation analyses were carried out to assess the strength of the relationship between the MPI-DLV scales and the corresponding experience sampling items. The ESM irritating response variable showed a very positively skewed distribution. A logarithmic transformation was carried out on this variable. (The correlation between the transformed and non-transformed variable was 0.97). A regression analysis was carried out, taking the MPI-DLV item "Rate the level of pain *at the present moment*" as the independent variable and the mean score on the related ESM item as the dependent variable.

7.4 Results

Compliance

All but one of the patients who gave their informed consent completed the study without serious difficulties. The one who did not, gave up after a day and a half because he felt the experiment was too stressful. This patient was excluded from the analyses.

A further measure of compliance was derived from the amount of missing data. There were two kinds of missing data: the 'real' missing data (no answer at all to a question) and the invalid answers (given more than 10 minutes after a signal). Missing data is reported here as a percentage (number of missing observations on a variable divided by the maximum number of possible observations on this variable). Table 7.3 shows these percentages for each experience sampling variable. As can be seen from this table, the total amount of missing data on the ESM variables is relatively large (except for on the partner response scales). The non-response to the item 'physical activity' was very high. A closer inspection of the missing data in the ESM-booklets indicated that, in almost all cases, there was at least one long period (about half a day or longer) during which the watch's signal function was turned off. Another explanation given by some patients for missing data was that in noisy situations they simply could not hear the signal. The average number of non-missing data (valid observations) varied from 32, out of 48, (physical activity) to

²The maximum number of valid observations for the analogues of the MPI-DLV (sub)scales Interference, Life Control, Affective Distress, General Activity and Household Chores was 48 (8 times a day multiplied by 6 days). The Pain Severity item had a maximum of 60: ten ratings per day (8 beeps, one score when the patient woke up and one just before going to sleep). For the analogues of the Punishing, Solicitous and Distracting Responses scales, the maximum number of observations was 6 (one rating per day, made at the end of that day). One rating per day was decided upon for these three variables because in many cases there is a structural absence of the patient's partner (at work).

50, out of 60, (pain intensity), and this was considered sufficient for obtaining reasonable estimates of the ESM variable means. With regard to the partner response items, for almost all patients the mean value could be based on the maximum of six observations.

Table 7.3 Percentage of missing data on the ESM-variables

	PS	I	LC	AD	GA	HC	SR	PR	DR
Not filled in	11	13	13	24	27	12	6	5	6
Invalid	5	6	6	2	6	6	0	0	0
Total	16	19	19	26	33	18	6	5	6

PS: Pain Severity, I: Interference, LC: Life Control, AD: Affective Distress, GA: General Activity, HC: Household Chores, SR: Solicitous Responses, PR: Punishing Responses, DR: Distracting Responses.

Table 7.4 presents the correlations between the MPI-DLV scales and the related ESM items, together with their significance. As can be seen, there are (relatively) strong positive relationships between the MPI-DLV scales Pain Severity, Interference, Solicitous Responses, Punishing Responses, Distracting Responses and Household Chores and their corresponding ESM variables. The correlation between the MPI-DLV Life Control scale and its corresponding ESM variable reached only marginal significance. The ESM items change in mood and physical activity did not correlate with the corresponding MPI-DLV scales.

Table 7.4. The relationship between MPI-DLV scales and corresponding ESM items

	Pearson correlation	P-value
Pain Severity	0.75	< 0.001
Interference	0.60	< 0.001
Life Control	0.25	0.06
Affective Distress	0.20	0.14
General Activity	0.16	0.22
Household Chores	0.40	0.002
Solicitous Responses	0.52	< 0.001
Punishing Responses	0.23	0.04
Distracting Responses	0.31	0.02

Frequency distributions showed an approximately normal distribution for both variables, allowing a regression analysis to be carried out. The resulting regression equation was: $ESM = 1.19 + 0.70 * MPIDLV$. The F -statistic of the regression model was highly significant ($F(1,55)=52.6, p<0.0001$) as were the t -value of the regression coefficient ($t=7.30, p<0.0001$) and constant ($t=3.40, p=0.001$).

Forty-nine percent of the variance could be explained by the model. Residual analyses, as well as a visual inspection of the scatterplot showed that there was one 'clear' outlier (MPI-DLV value 0; ESM value 4.12; standardized residual 2.9). Disregarding this case resulted in a remarkably improved data fit ($F(1,54)=73.9$; $p<0.0001$; Multiple-R = 0.76; R-squared = 0.58). The new regression equation was: $ESM = 0.78 + 0.81 * MPI-DLV$. The constant in the equation was significantly different from 0 ($t=2.32$, $p=0.02$). Figure 7.1 contains a scatter plot of the standardized scores on the item "rate the level of pain at the present moment" on the MPI-DLV Pain Severity scale and its ESM equivalent.

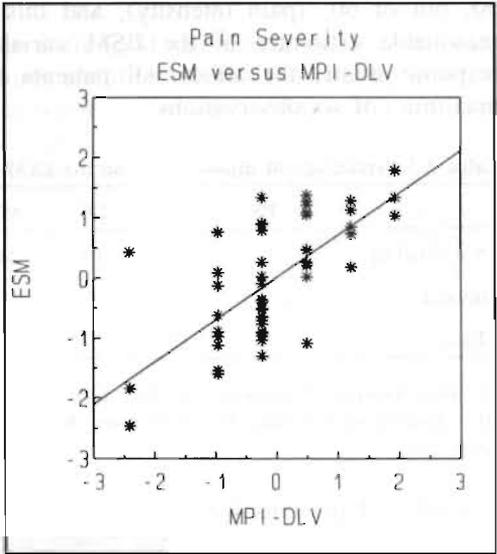


Figure 7.1 Pain intensity standardized scores: ESM versus MPI-DLV

Finally, the difference between the mean pain severity score of the MPI-DLV and the related ESM item was tested. The mean MPI-DLV pain severity score was 3.52, the mean ESM score was 3.31. This difference was not significant ($t=1.02$, $p=0.314$).

7.5 Discussion

In principle, a validation process is never-ending, in that all information on new validity facets merely adds to a more complete understanding of the properties of the instrument under consideration. As stated above, validation research on the MPI has already been successful. One of the purposes of the present investigation was to obtain additional information on the ecological validity of the MPI-DLV; in particular, on the relationship between relatively stable / trait information from the MPI-DLV scales and analogous items (measuring state information) in an experience sampling procedure.

Six MPI-DLV scales showed significant correlations with the corresponding ESM items. These were Pain Severity, Interference, Solicitous Responses, Punishing Responses, Distracting Responses and Household Chores. The correlations between the MPI-DLV Pain Severity and Interference scales and their ESM analogues are comparable to those reported by Flor et al. (1990b). The ESM situation control item correlated marginally significantly to the MPI-DLV Life Control scale. The results relating to these MPI-DLV scales would seem to show

that valid information can be obtained about a patient's behavior and thoughts in daily life, using a comparatively short instrument (the MPI-DLV).

No relationship was found between the MPI-DLV scales Affective Distress and General Activity and their related ESM items. As has already been mentioned, the problem relating to the General Activity scale had to do with the lack of an adequate analogous item for the ESM. Nevertheless, based on the results of a working-to-tolerance bicycle ergometer test (Lousberg et al., 1993c) where a positive correlation had been found between the patients' performance levels and the General Activity scales, expectations were that the ESM 'physical activity' item would also show a positive correlation to this scale. As a matter of fact, a moderate positive correlation (0.16) was found but it failed to reach significance. It is interesting, however, that the MPI-DLV General Activity sub-scale Household Chores did show a fairly high correlation to a (reliably) coded household activities score. As was previously mentioned, Flor and co-workers (1990b) also failed to find a significant correlation between the diary and MPI-D General Activity scale. These findings concur with Linton's (1985) conclusion that there is a discrepancy between the patient's perception of activities and the objective activity level. All in all, it is not clear at the moment whether the MPI-DLV General Activity scale adequately represents patient's physical activity level.

With respect to the ESM 'mood change' item, the lack of significance can most likely be attributed to the content of the item (the emphasis on change instead of on mood alone). It is left to future research to investigate the relationship between the MPI-DLV Affective Distress scale and an experience sampling analogue.

Another aim of this study was to investigate the nature of the relationship between the MPI-DLV and ESM with regard to present pain intensity. Examination of the data revealed one outlier. The problem of what to do with outliers is still unsolved (Cohen and Cohen, 1983). There were no reasons - such as a failure to carry out an instruction, a low intelligence level, etc. - which could account for the large discrepancy, and therefore provide a reasonable argument for disregarding the case. Nevertheless, because of the relatively small number of subjects and the strong influence of this patient's score on the regression model, the authors still consider the second regression model (with the outlier omitted) to be a better estimation of the relationship. The parameters of the second regression equation convey useful information. A low score on the MPI-DLV pain intensity item is probably an underestimation of patient's mean daily pain intensity level. On the other hand, a high score on the MPI-DLV pain item should probably be interpreted as a slight overestimation of the mean daily pain level. As to the small discrepancies in the low and high regions, it can be argued that this is a phenomenon of 'regression to the mean'. It is very likely that a patient with a maximum score (6) on the MPI-DLV pain item will report at least some lower scores during the experience sampling period. Similarly, a minimum MPI-DLV score of 0 will very likely be transformed into some higher scores measured in daily life (if not, the term chronic pain patient becomes questionable). In general, though, it can be

stated that the MPI-DLV score on pain intensity reflects the patient's mean pain experience level in daily life.

A major point of concern in the present study was the amount of missing data. Although the planned analyses could be carried, in future research studies, measures will have to be taken to minimize the possibility of missing data, especially when fluctuations in pain are under investigation (Affleck et al., 1991; Jamison and Brown, 1991).

In conclusion, the present results lend further support to the validity of six of the MPI-DLV scales. Given the scores on these MPI-DLV scales, valid conclusions can be drawn about some aspects of a patient's behavior / thoughts in daily life. In addition, the finding that data from a short questionnaire strongly resembles that yielded by extensive and patient-loaded information-acquisition methods, is very important. As far as the MPI-DLV scales Affective Distress and General Activity are concerned, further research is needed before conclusions can be drawn about their generalizability to daily life situations.

PROFILE CHARACTERISTICS OF THE MPI-DLV CLUSTERS¹

8.1 Summary

This article describes an investigation into whether there are differences between the four pain patient clusters from the Multidimensional Pain Inventory - Dutch Language Version (MPI-DLV), in terms of prescribed analgesic medication, level of intelligence and other personality characteristics (MMPI). It was predicted that patients classified as 'Dysfunctional' would be distinguished by a higher level of analgesic medication intake. This hypothesis was confirmed. Patients classified as 'Average' were marked by a marginally significant higher level of intelligence. The personality profile of 'Interpersonally Distressed' patients indicated an elevated neurotic triad and a tendency to a passive-aggressive personality structure. Patients classified as Dysfunctional and Average were marked by the so-called conversion-V. There were no elevated scales in the Adaptive Coper profile. The clinical implications of the findings are discussed.

8.2 Introduction

The Multidimensional Pain Inventory (MPI) has been proven to be a reliable and valid questionnaire (Kerns et al., 1985) measuring several pain related aspects. An interesting feature of this questionnaire is that individual pain patients can be reliably assigned to three meaningful subgroups (Turk and Rudy, 1988). In addition to being used for assessment purposes, the MPI can also be used as a research instrument (for instance to measure therapy-outcome effect). This instrument has been translated into German and Dutch, resulting - respectively - in the MPI-D and MPI-DLV, and attempts at reproducing its factor structure have been successful (Flor et al., 1990b; Lousberg et al., 1993c). The psychometric qualities of these two translated versions have also been proven to be good and comparable to those of the American version. A cluster analysis of a large pain patient sample in the Dutch situation revealed, apart from some small differences, the same three patient clusters (Dysfunctional, Interpersonally Distressed and Adaptive Coper) as in the Turk and Rudy study (1988). An additional fourth cluster, designated Average, was

¹Submitted for publication by Lousberg, R., Groenman, N.H. & Schmidt, A.J.M. (1993f).

identified. The uniqueness and validity of these four patient clusters was demonstrated.

Interpretation of the characteristics of the MPI-DLV sub-groups is based on their mean score profile configuration from nine MPI-DLV scales. These scales are Pain Severity, Interference, Life Control, Affective Distress, Support, Punishing Responses, Solicitous Responses, Distracting Responses and General Activity. Although these nine scales represent a fairly broad range of important pain-related variables, a more complete and detailed picture of the MPI-DLV cluster characteristics could be gained from additional information on such profile features as the use of prescribed analgesia, intelligence, and other personality characteristics². The choice of these three aspects was made on practical grounds, in that these variables formed a part of a previous study (Lousberg et al., 1993c).

The following hypothesis regarding prescribed analgesia was tested. Patients classified as Dysfunctional take prescribed analgesia more frequently than the other three MPI-DLV sub-types. This hypothesis was based on a similar observation made by Turk and Rudy (1988). No specific hypotheses were formulated with regard to the variables intelligence and other personality characteristics.

8.3 Method

A computer software program which employs a classification technique (based on multivariate generalized squared distances and Bayesian posterior probabilities) described by Tatsuoka (1988), was used to assign individual patients to one of the four MPI-DLV sub-types.

Where a patient could not be reliably assigned to one of the four clusters, the classification 'Anomalous' was given. Patients classified as Anomalous were excluded from the analyses.

Data on the use of prescribed analgesia was taken from the experience sampling procedure described by Lousberg et al. (1993e). This involved 57 chronic pain patients filling in experience sampling booklets over six consecutive days. The beep of a watch (which sounded eight times a day, at random intervals), indicated to patients that they should answer some questions in a booklet. Use of medication (whether analgesia or not) since the last beep, was one of the variables to be measured. Medication use was coded according to the category system given in the "Farmacotherapeutisch Kompas" (Ziekenfondsraad, 1992), a Dutch catalogue of registered medication. The present study focused on the category 'Analgesia and antirheumatics'.

²In scientific literature there is no consensus on the definition of personality. Noting that one of the definitions emphasizes the fact that personality can be viewed as an organized system which differentiates persons (see Hall and Lindzey, 1978), intelligence is considered here as a personality characteristic.

The measurement of intelligence and other personality characteristics forms part of the psychological assessment procedure at the Pain Center of the Academic Hospital Maastricht. Scores on the shortened version of the Groninger Intelligence Test (Luteijn and Ploeg van der, 1983) were available from 83 patients. The GIT intelligence level (IQ) has a mean distribution of 100 and a standard error of 5 (shortened GIT).

The Minnesota Multiphasic Personality Inventory (MMPI) was used to measure other personality characteristics. MMPI and MPI-DLV data from 86 patients was available. The MMPI profiles were interpreted along the lines described by Graham (1987), Greene (1980) and Bigos et al. (1991).

Since no a-priori hypotheses were specified in advance for intelligence and personality, all tests (ANOVAs) were performed two-tailed.

8.4 Results

Use of prescribed analgesic medication

On the basis of the MPI-DLV scores, 53 of the 57 cases could be assigned to one of the four patient clusters (the remaining four patients were classified as Anomalous). The total number of medication intakes (during the 6 days of the ESM experiment), coded to the analgesia category, was computed for each patient. Table 8.1 shows the number of patients belonging to the each cluster, the mean number of analgesic intakes per person in the clusters, and the mean score on the MPI-DLV Pain Severity scale.

Table 8.1 Mean number of analgesic intakes in the MPI-DLV clusters

	N	Number of analgesic intakes		MPI-DLV Pain Severity
		mean	s.d.	
Dysfunctional	18	8.33	8.5	4.2
Interpersonally Distressed	4	4.25	5.7	3.4
Adaptive Copier	13	0.53	1.2	2.4
Average	18	6.11	7.3	3.6

Unfortunately, only four patients were classified as Interpersonally Distressed. The results of the subsequent analyses should therefore be interpreted with care. Based on the appropriate cluster contrast, an analysis of variance indicated that Dysfunctional patients do have a significantly higher analgesic intake ($t=2.20$, $p=0.03$) than patients in the other three clusters³.

³An investigation was carried out to see whether one or a few patients with a very large analgesic intake, caused the effect. However, inspection of the analgesic intake frequency distribution revealed that this was not the case.

These results confirmed the hypothesis given above and thus replicated Turk's finding. However, taking the mean MPI-DLV Pain Severity levels into account, the value of these results has to be seen in context, as was the case with the potential cluster confounds (see Chapter 6). In other words, is it not merely a logical sequence that someone who reports higher levels of pain takes more pain medication? Or, alternatively, can a higher analgesic intake be regarded as a patient characteristic 'in its own right'. To answer this question an analysis of co-variance was carried out using the MPI-DLV Pain Severity score as a co-variate. The effect of the co-variate was marginally significant ($F=3.65$, $p=0.06$) and explained enough of the variance for the significant cluster-effect to disappear ($F=1.17$, $p=0.17$). However, when the two most different groups with respect to analgesic usage (Dysfunctionals and Adaptive Copers) - groups containing sufficient numbers of cases - were compared, the cluster effect was significant ($t=1.81$, $p=0.038$).

Intelligence

The mean intelligence levels of the four MPI-DLV clusters are given in Table 8.2. The IQ's differed significantly between the clusters ($F=3.34$, $p=0.02$).

Table 8.2 Mean intelligence levels of the MPI-DLV clusters

	N	Shortened GIT score	
		mean	s.d.
Dysfunctional	34	98.4	11.1
Interpersonally Distressed	17	105.9	17.9
Adaptive Coper	17	104.9	7.6
Average	15	109.0	10.2

Separated post-hoc comparisons indicated that Dysfunctional patients have a lower GIT score than patients in the other three MPI-DLV clusters ($t=3.04$, $p=0.003$). A comparison between the patients classified as Average and the other three clusters taken together, revealed a marginal effect ($t=1.70$, $p=0.09$). After comparison, the IQ of the Average group was found to be considerably higher than the mean intelligence level of a large, representative norm-group taken from the Dutch population at large ($N=1570$).

Other personality characteristics

Of the 86 patients on whom MMPI data was available, thirty-six were classified as Dysfunctional. Of the remaining patients, 16 were assigned to the Interpersonally Distressed cluster, 17 to the Adaptive Coper cluster and 15 to the Average cluster. Two cases were classified as anomalous and were left out of the analyses. The mean MMPI profile was computed for each of the MPI-DLV clusters. The K-

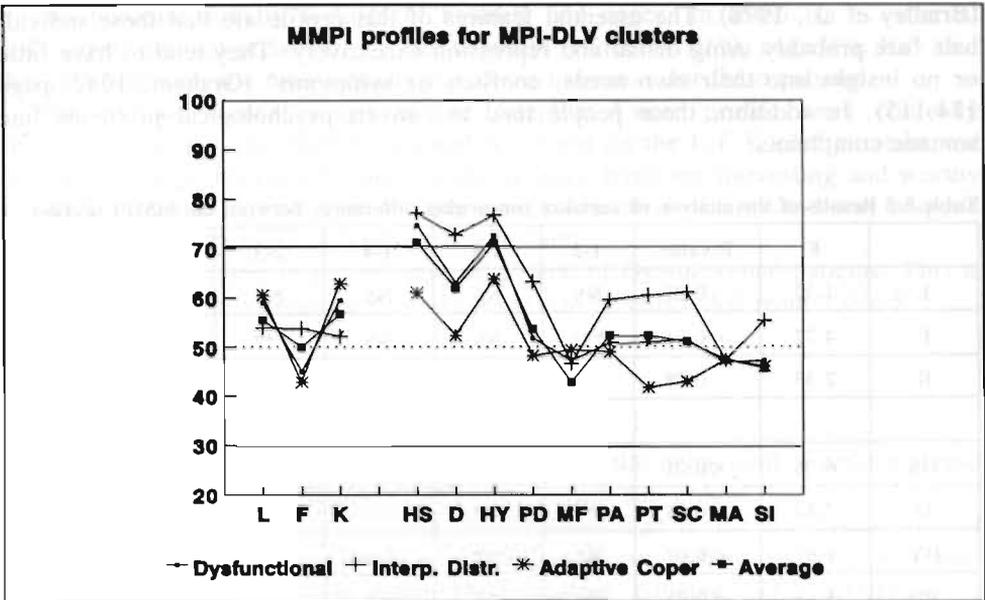


Figure 8.1 MMPI profiles for the MPI-DLV clusters.

corrected T-scores are presented in Figure 8.1⁴ in the form of a graph. A multivariate analysis of variance was carried out to test the differences between the four MMPI profiles. The results are summarized in Table 8.3. Taking all MMPI-scales together, the four profiles differ quite markedly from one other, as is indicated by the highly significant multivariate *F*-value. The results of the univariate *F*-tests showed that significant differences were found in eight of the thirteen MMPI scales. In order to obtain more detailed information on the differences between the MPI-DLV clusters on each of the MMPI-scales, analyses were carried out on the basis of specific contrasts. The significance of these contrasts can be seen in the right-hand section of Table 8.3. The MPI-DLV clusters Dysfunctional, Interpersonally Distressed, Adaptive Coper and Average are numbered 1, 2, 3 and 4 respectively. For example, column '1-2' lists the significance levels obtained by testing the T-score on a specific MMPI-scale of the Dysfunctional profile against the T-score of the Interpersonally Distressed profile on that same scale.

It was concluded, on the basis of the L-F-K configurations, that there were no reasons to assume invalid MMPI profiles.

The MMPI profile of the Dysfunctional type is marked by high (> 70) scores on the scales Hypochondriasis and Hysteria, in combination with a slightly elevated depression scale - the conversion-V profile already well known in the literature

⁴The usual way of report MMPI results is by scale scores which are converted to T-scores (mean = 50, S.D. = 10). Scale scores higher than 70 or lower than 30 are both of statistical significance and clinical relevance.

(Bradley et al., 1978). The essential features of this profile are that these individuals "are probably using denial and repression excessively. They tend to have little or no insight into their own needs, conflicts or symptoms" (Graham, 1987, page 114-115). In addition, these people tend to convert psychological problems into somatic complaints.

Table 8.3 Results of the analysis of variance concerning differences between the MMPI profiles

	F	P-value	1-2	1-3	1-4	2-3	2-4	3-4
L	1.32	0.27	NS	NS	NS	NS	NS	NS
F	4.72	<0.01	***	NS	NS	***	NS	*
K	2.33	0.08	*	NS	NS	**	NS	NS
HS	6.76	<0.01	NS	***	NS	***	NS	*
D	7.84	<0.01	**	***	NS	***	**	*
HY	4.81	<0.01	NS	***	NS	***	NS	*
PD	5.61	<0.01	***	NS	NS	***	**	NS
MF	1.55	0.21	NS	NS	NS	NS	NS	*
PA	2.71	0.05	**	NS	NS	**	NS	NS
PT	7.74	<0.01	***	**	NS	***	*	**
SC	6.66	<0.01	**	**	NS	***	*	*
MA	0.02	0.99	NS	NS	NS	NS	NS	NS
SI	2.26	0.09	*	NS	NS	*	*	NS

MMPI scale abbreviations. L: Lie; F: False; K: Correction; HS: Hypochondriasis; D: Depression; HY: Hysteria; PD: Psychopathic Deviation; MF: Masculine Feminine; PA: Paranoia; PT: Psychasthenia; SC: Schizophrenia; MA: Hypomania; SI: Social Introversion.

Multivariate $F(\text{Hot})=1.85, p=0.003$

*: $p \leq 0.05$

** : $p \leq 0.01$

***: $p \leq 0.005$

NS: Not Significant

All three scales of the neurotic triad do have clinically elevated scores in the Interpersonally Distressed profile. Reactivity to stress and the avoidance of responsibility by developing physical symptoms are the usual characteristics of this profile. A depressive diagnosis is often made. There is also a slight tendency towards a "Scarlett O'Hara V" (a low MF scale relative to the PD and PA scales). This configuration is most often seen in women. Analyses did indeed reveal that the "V-shape" was sharper for the nine women than for the seven men in the Interpersonally Distressed group. A 'Scarlett O'Hara V' is indicative of an inability

to communicate directly feelings of hostility and anger. In addition, "... marital difficulties, familial problems, are common" (Greene, 1980, page 99). It is said that psychotherapeutic interventions are contra-indicative for these patients.

None of the clinical scales of the Adaptive Coper profile were clinically elevated. However, the strongly marked V-pattern on the L-F-K configuration and the rather less pronounced V-form on the neurotic triad are interesting and worthy of mention. These patients tend to deny (psychological) problems and may respond to psychological stressors with physical complaints.

The Average profile is almost identical to that of Dysfunctional patients. This is underlined by the non-significant contrasts, as can be concluded from Table 8.3.

8.5 Discussion

The present data has provided greater insight into the behavioral, intellectual and other personality characteristics of MPI-DLV patient sub-groups.

With reference to analgesic usage, the analyses lead to the conclusion that, if compared on the basis of pain level, Adaptive Coper patients do take less prescribed analgesic medication than Dysfunctional patients. The difference in medication intake between these two MPI-DLV patient sub-groups cannot, however, be viewed **solely** as a function of pain level. As far as the clinical implications of this finding are concerned, it could be argued that a treatment program for Dysfunctional patients should pay special attention to the role and function of analgesic medication usage. However, since the number of patients was very low, these results have to be interpreted with care. Replication of this study based on a larger sample is necessary to obtain more reliable results.

Investigations regarding intelligence level produced some unexpected differences in IQ. One tentative conclusion can be drawn. The mean scale scores are characteristic of the Average type in the MPI-DLV profile. As was argued earlier (Chapter 6, page 85) this patient type seems to respond to the MPI-DLV questions in a nuanced way. It is therefore unlikely that answers from this patient group will produce large numbers of extreme scores. This profile characteristic accords with the (marginally) higher level of intelligence of the Average type. This finding not only leads to a better understanding of how the Average type differentiates from the other three MPI-DLV profiles, but could also have therapeutic implications. In setting up specific treatment programs for the patient sub-groups, it should perhaps be borne in mind that insight-giving therapeutic components requiring a higher than average intellectual level would probably prove more effective in a program designed for the Average type. Consideration might also be given to modifying the treatment program in a general sense: e.g., the way in which the program is presented, the manner of communicating with the patients, etc. The issue of tailoring treatment programs to particular MPI-DLV patient sub-groups will be discussed in more detail in Chapter 9.

Although the psychometric properties of the MMPI have been criticized (e.g., Turk, 1990), this instrument is still used worldwide in personality assessment. As was already stated in Chapter 1 (page 13), the popularity of the MMPI in pain research has resulted in an almost incalculable number of articles. In the present study, it was the link between the MPI-DLV pain patient clusters and personality characteristics as measured on the MMPI that was under investigation. It should be noted that the MMPI-profiles are mean profiles. Comparing the mean MMPI configurations of the MPI-DLV clusters with those reported by Costello et al. (1987) has revealed some interesting similarities. In outline, Costello's N(ormal) type seems to resemble the Adaptive Coper MMPI profile, and Costello's A (conversion-V) profile is almost identical to that of the Dysfunctional and the (MPI-DLV) Average cluster type. The situation regarding the Interpersonally Distressed MMPI profile is less clear, but there are some characteristics which also appear in Costello's P(sychopathological) profile: the high neurotic triad in combination with the V-form on the PD-MF-PA scales, and the slightly elevated PT and SC scales. The scales PD, PA, PT and SC may, however be considered too low to establish any real resemblance to the P(sychopathological) profile. Rather, it may be argued that the Interpersonally Distressed MMPI profile more closely resembles the Type-I profile, where only the scales from the neurotic triad have elevated scores.

It was, however, possible to draw the following conclusions about the MMPI-profiles. It would seem, from these results, that the MPI-DLV clusters really do represent unique patient-subgroups in terms of aspects of personality. Knowledge of their personality structure will probably lead to a better understanding of patients' behavior. The MMPI profiles were found to be broadly in line with the expectations. The Adaptive Coper MMPI profile showed no clinically elevated scales, suggesting a patient type with a relatively stable and balanced personality. One of the essential MMPI-profile features of Interpersonally Distressed patients, problems with (marital) conflicts, is totally in line with the MPI-DLV scale configuration. The MMPI profiles of the Dysfunctional and Average groups represent the characteristics associated with the often cited conversion-V.

It would seem, then, that the differences between the MPI-DLV clusters, in terms of prescribed analgesic medication, intelligence level and personality characteristics, have shown themselves to be of both statistical significance and clinical importance. Future studies will be needed to overcome some of the problems mentioned above and to investigate the value of the clinical recommendations made.

GENERAL DISCUSSION

This Chapter starts with a critical consideration of the results of the MPI-DLV studies (Chapters 5 to 8) and the operant conditioning experiments (Chapters 2-4)¹. This is followed by an attempt to integrate these two lines of research. The Chapter concludes with recommendations for future research.

9.1 MPI-DLV development and patient sub-types

Chapter 5 described the development of the Multidimensional Pain Inventory - Dutch Language Version. Apart from some small differences, the attempt at replicating the MPI factor structure was successful and the psychometric qualities of the MPI-DLV were proved to be adequate. In addition, an experimental study revealed the construct validity of the MPI-DLV General Activity scale, and an experience sampling study showed the ecological validity of some scales. Evidence was also found of the MPI-DLV's capacity to measure treatment effects.

Four patient sub-groups were identified on the basis of the MPI-DLV scale scores and the application of multivariate statistical techniques. The profiles of three of these sub-groups highly resembled the three patient groups reported by Turk and Rudy (1988) and were given the same names as their American counterparts. The fourth MPI-DLV cluster was called Average, because of the mean score profile. The results of several analyses provided further evidence of the distinctiveness and uniqueness of the four MPI-DLV profiles. 'New', individual patients can be reliably assigned to one of the four patient sub-groups. Additional profile characteristics of the four patient clusters were presented in Chapter 8. A summary of the essential features of each MPI-DLV patient type is given in the following paragraphs.

The Dysfunctional patient type shows a high level of pain severity, interference and affective distress and a low level of life control. Unlike the Interpersonally Distressed type, Dysfunctional patients report having a very supportive environment. His/her partner mainly reacts to the patient's pain behaviour in terms of solicitous and distracting responses. There is also evidence that the Dysfunctional type takes more analgesic medication than the other MPI-DLV types (even after

¹For the convenience of the reader, these paragraphs also contain a brief summary of the most important results.

correction for the high pain severity level). The Dysfunctional cluster type seems to have a somewhat lower intelligence level than the other three MPI-DLV types.

The Interpersonally Distressed patient is also marked by a high degree of suffering (pain severity, affective distress, etc.). However, he/she also suffers on the interpersonal level, as is indicated by the low level of environmental support. The patient's pain behavior is likely to be 'punished' by the partner.

The Adaptive Copier sub-type shows relatively low levels of pain and psychological distress. The level of environmental support is somewhat lower than that of the Dysfunctional type, but considerably higher than that of the Interpersonally Distressed type. This patient type is the most active, takes the lowest level of analgesic medication and has the shortest duration of pain.

The characteristics of the Average type, as the name suggests, have to be 'situated somewhere between the other three MPI-DLV patient types'. The only distinguishing feature seems to be the relatively high intelligence level.

9.2 MPI-DLV development and patient sub-groups: evaluative comments

The first two points are related to the psychometric development of the MPI-DLV and the five remarks which follow relate to the patient sub-groups.

1. The importance of the experimental assessment of validity has been stressed. The bicycle ergometer test and the experience sampling experiment are two examples of how this kind of assessment can be carried out. However, more research is needed to obtain a more detailed validation of all MPI-DLV scales. For example, part three of the MPI-DLV measures three relatively highly inter-correlated activity factors: 'Household Chores', 'Outdoor Work' and a combined 'Social Activities / Activities Away From Home' factor. These three factors together form the General Activity level. The positive correlation between the performance level on the bicycle ergometer test and the score on the General Activity level provides information on the construct validity of the three activity factors combined. However, nothing can be said on the basis of these results, about the discriminant validity of the three activity factors. In the experience sampling study, the MPI-DLV Household Chores activity scale showed a significant correlation to the frequency of household activities, as measured in the experience sampling booklets. The validity of the factors Outdoor Work and Social Activities / Activities Away From Home, however, have still to be investigated. Similar critical observations can be made regarding the discriminant validity of other correlated factors (Pain Severity and Interference, Support and Solicitous Responses).
2. The psychometric analyses were based on a patient sample containing a majority of (female) fibromyalgia patients. That this fact questions the generalizability of the results to other pain diagnoses or pain loci cannot be ignored. Although one analysis indicated that there were no differences

- between fibromyalgia and back pain patients with regard to the MPI-DLV factor structure, analysis of a patient sample of a more evenly balanced composition is needed before definite conclusions can be drawn.
3. The interpretation of profile characteristics must always be considered in context (i.e., in comparison with the other three MPI-DLV patient clusters). This is especially true where the Adaptive Coper type is concerned. These patients obviously experience enough pain to make them seek professional help. Another point of caution that should be borne in mind when interpreting the MPI-DLV cluster assignment, relates to the probabilistic nature of the classification. Commenting on this, Rudy said: "Perhaps the most important basis for using quantitative approaches to the classification of patients is that they provide an objective and operationally specified way of describing how individual patients get assigned to particular groups or categories. From a purely research standpoint, these methods have substantial advantages over the use of subjective classification methods that are not likely to be understood and used in the same way by different clinicians" (Rudy, 1989, page 35). The present author would also emphasize that the interpretation of the MPI-DLV results should take place in combination with other assessment data, such as that from other questionnaires, data obtained on intake, the medical status data already known, etc.
 4. The MPI-DLV measures several pain-related psychosocial and behavioral variables. The goal of multi-axial assessment (MAP) is, however, not achieved until this information is supplemented by pain-relevant medical information (e.g., the MEDICS system). At the time the sub-group analyses were carried out, no quantifiable measure of physical pathology (such as the MEDICS score) was available. Not only does this mean that the MAP taxonomy was not complete, but medical information would have been very useful with regard to the issue of potential confounds. As was found by Turk and Rudy (1988), the Adaptive Coper can be expected to have a lower physical pathology level than the other three types. The results of an analysis of co-variance, testing the differences between the MPI-DLV clusters after correcting for physical pathology (like analyses in Chapter 6), would no doubt shed further light on this critical 'potential confound' issue.
 5. In principle, assignment to one of the four patient clusters requires that there be no missing MPI-DLV scale scores². Particular care should be taken with MPI-DLV part two. In this part the patient has to answer questions relating to the reactions of a significant other to his/her pain behavior. Most patients live with a partner. When they do not, the significant other should be a person

²In the MPI computer program user's manual (version 2.0), Rudy states that a missing scale score could be replaced by the mean score of the T-distribution (50). This substitution allows classification of the case to be attempted. However, it is reported that if more than two MPI scales scores are missing, the classification process would produce an unreliable result.

who frequently visits the patient³. It might well be that differences with respect to the frequency and nature of contact with the significant other, have serious consequences for the cluster assignment. This issue has also yet to be investigated.

6. The analyses regarding analgesic medication intake (Chapter 8) are based on a relatively small sample. The number of patients classified as Interpersonally Distressed (N=4), in particular, is extremely low. A larger sample size is to be highly recommended, since the findings are interesting and might have implications for therapy. If the findings were replicated using a larger sample, the fact that analgesic medication intake is not merely a function of pain severity level may lead to the formulation of other hypotheses. For example, medication intake, as a pain behavior, might be controlled principally by operant learning factors. If this hypothesis (pertaining to an underlying mechanism for the Dysfunctional patient's behavior) could be confirmed in an experimental study, this could have clear consequences for therapy. This issue will be discussed in more detail below.
7. A striking finding to emerge from the profile characteristics is the difference in intelligence level. The mean intelligence level of the Average type is more than ten IQ-points higher than that of the Dysfunctional type. This finding has to be seen in context⁴. Compared to a large sample of the Dutch population the intelligence level of the Dysfunctional type falls within the mean range. The intelligence level of the Average type falls within the "above mean" range. Care should also be taken with respect to an interpretation in terms of a cause-effect relationship (intelligence level 'causes' a specific MPI-DLV sub-type). Future research is called for, since further examination of the observed differences in intelligence level would no doubt be of value. It might be that the difference is 'caused' by different verbal performance levels or, alternatively, by different performal functioning. As was argued in Chapter 8, significantly different intelligence levels can have major consequences for the composition of a treatment program and the way in which it is presented.

Given the above, it may be concluded that the MPI-DLV is a psychometrically sound instrument which can be applied in clinical assessment as well as for research purposes.

³As a third possibility, a patient may report having no significant other. In this case no cluster assignment can be made.

⁴Care should be always be taken when drawing conclusions on a socially loaded concept such as intelligence. In scientific literature there is not even consensus regarding the definition of intelligence (see the discussion between Eysenck and Kamin, 1981).

9.3 Operant conditioning of pain: evaluative comments and post-hoc analyses

Chapters 2 to 5 dealt with an investigation into whether pain behavior and the experience of pain could be conditioned in an operant way.

The treadmill test experiment, described in Chapter 2, is not a direct test of the operant 'developmental' mechanism of pain. Rather, this experiment tested a prediction based on the conviction that pain can be learned operantly. That is, the operant learning theory, when applied to pain, suggests that, once pain has been learned operantly, the presence of a solicitous spouse may function as a discriminative stimulus for pain behavior. Since pain behavior can take many forms - the pain behavior construct can be divided into several dimensions (Vlaeyen et al., 1987) - endurance (operationalized as walking time and heart rate) was measured in addition to verbal pain complaint. The results have shown that, in the presence of a solicitous spouse, patients report more pain and show poorer physical performance. It should be noted that these results, in accordance with the hypothesized effect, were only found in the spouse's pain-response ratings. The predicted effect was not found on the patient's pain-response ratings. In discussing this discrepancy, it was argued that the spouse's opinion of his own responses is perhaps more objective than that of the patient. However, the meager internal consistency of the spouse-response scale precluded a definite conclusion.

Chapter 3 described an experiment which was set up to replicate the findings reported by Linton and Götestam (1985). In this experiment the operant developmental mechanism of chronic pain was tested **more** directly. Healthy subjects were subjected to a series of acute pain stimuli. In the natural situation, in almost all cases, acute pain behavior will be followed by environmental consequences, whatever they may be. This natural situation was reproduced in the Linton and Götestam experiment and in those described in Chapters 3 and 4. It was thought that words of praise and approval from the experimenter would act as potent rewarding reinforcers for occasioning increases in pain ratings. Using these, and 'punishing remarks' in response to decreases in pain, an attempt was made to condition pain ratings to higher levels. Another aspect of the investigation was to determine whether pain-related psychophysiological responses could also be conditioned. Although the results from this first attempt at replication (a heightened pain report, a heightened magnitude matching of the last stimulus and increased skin conductance responses in the up-conditioning group) accorded fully with the hypothesized effects, the second replication experiment (Chapter 4) shed a completely different light on the results of first conditioning study. The methodological modification in the second study, with regard to the conditioning process, was related to the concentration-centered punishing responses. After the content of the punishing remarks had been 'neutralized', however, the 'conditioning' effects of the first study could not be replicated. As was already stated on page 54, the increasing VAS and SCR curves most likely reflect an attention shift towards the pain. It was also stated that, based on their observations in the laboratory, the two

experimenters felt that conditioning had probably taken place in a sub-group of subjects. The most salient results of the post-hoc analyses have already been given in the discussion section of Chapter 4. Post-hoc analyses may provide a fertile ground for new hypotheses. Therefore a more detailed report of this strongly statistically oriented exercise will be given here.

A visual inspection of the **individual** VAS-profiles showed that there were decreasing / habituation-like curves, more or less stable curves, and increasing / sensitizing-like curves in **all** three groups (up-conditioning, down-conditioning and control). It was therefore decided to divide the profiles of each of the three conditions into three sub-groups. Taking the VAS responses first, difference scores were computed in which the VAS difference score equalled the mean VAS score from trials 6 to 12, minus the mean VAS score from trials 1 to 5. Thus, a negative VAS difference score denotes a decrease in VAS-scores in the conditioning phase - compared to the baseline -, whereas a positive difference score represents an increase in VAS-scores. Similarly, difference scores were computed for the SCRs. Difference scores were also used with regard to the magnitude matchings, as was outlined on page 50. The sub-groups were created on the basis of the VAS-difference score distribution thus obtained. The 33% split was chosen as the means of defining the sub-groups, because of the relatively small total sample size. In other words, the 33% highest, the 33% lowest and the 33% remaining VAS-difference values, were respectively assigned to a 'sensitizing', 'habituating' and 'stable' subgroup.

Table 9.1 gives the mean values of the VAS difference scores for the sub-groups thus created. The mean values of the sub-groups are reflected in the designations given them. The three sensitizing sub-groups are marked by a positive VAS-difference score and the habituating subgroups are marked by a clear negative VAS score. The mean scores of the stabilizing sub-groups are slightly below zero, indicating a habituating tendency. It was argued that if conditioning of pain had taken place in a sub-group of subjects, the following post-hoc hypotheses would hold true:

1. The mean pain-rating score from the 'sensitizing sub-group' of the up-conditioning group would be **higher** than those from the sensitizing sub-groups of the down-conditioning and control groups;
2. The mean pain-rating score from the 'habituating' sub-group of the down-conditioning group would be **lower** than those from the habituators in the up-conditioning and control groups.

With regard to SCR and magnitude matching, it was hypothesized that:

3. SCRs would be higher and magnitude matching differences (relatively) more positive in the sensitizing sub-group of the up-conditioning group than in the other two sensitizing sub-groups;

4. SCRs would be lower and magnitude matching differences (relatively) more negative in the habituating sub-group of the down-conditioning group than in the other two habituating groups.

Table 9.1 VAS-difference scores of the sub-groups

Condition	Sub-group	mean	s.d.
Control	habitators	-14.9	5.89
	stabilizers	- 4.7	2.69
	sensitizers	4.5	4.75
Up	habitators	-11.9	5.77
	stabilizers	- 2.1	1.97
	sensitizers	10.4	5.38
Down	habitators	-11.9	7.03
	stabilizers	- 0.2	1.70
	sensitizers	6.7	3.46

Sensitizer groups

With regard to post-hoc hypothesis 1, an analysis of variance was carried out with VAS-difference as the dependent variable and group (sensitizers, stabilizers and habitators) and condition (up, down and control) as the independent variables. The following interaction contrast was analyzed: up contrasted against down and control, in combination with sensitizers contrasted against habitators and stabilizers. This interaction contrast produced a significant effect ($t=1.73$, $p=0.04$). This effect, however, requires careful interpretation. As can be concluded from the data in Table 9.1, this interaction effect is not only due to a relatively stronger increase from the stabilizer to the sensitizer group of the up-conditioning group (12.5 points compared to 9.2 points in the control group and 6.9 points in the down-conditioning group). It is also due to a slightly smaller than expected increase in the down-conditioning sensitizer sub-group. If the sensitizer sub-group of the down-conditioning group had shown an equal increase (from stabilizer to sensitizer sub-group) to that of the control group, the increasing effect in the up-conditioning group would probably have been too small to reach significance. In conclusion, there *seems* to be an indication that up-conditioning *in a sub-group* of subjects did take place.

A similar analysis was carried out to test post-hoc hypothesis 3 (regarding the SCRs). Table 9.2 shows the mean SCRs in the three sub-groups of each group. There was no significant group x condition interaction effect ($t=0.34$, $p=0.37$). As can be seen in table 9.2, the mean SCR difference score of the sensitizing sub-group of the up-conditioning group is even more negative than the mean SCR difference score from subjects in the stabilizer sub-group of the up-conditioning group. On the other hand, and as would be expected, the SCR difference scores in the sensitizing sub-groups of the control and down-conditioning groups are less negative (less habituating) than the habitators and stabilizers in these conditions. This puzzling situation immediately becomes clear when the pattern of the tension

scores is inspected (see Table 9.3). In the control and down-conditioning groups subjects become increasingly tense along the group-sequence: habituators, stabilizers and sensitizers. It should be noted, however, that this was not the case in the up-conditioning group. Remarkably, the sensitizers in the up-conditioning group became considerably more relaxed than the stabilizers in the up-conditioning group. When the pattern of the SCR difference scores and the tension-change scores of the three sub-groups are compared, the striking resemblance leads to the conclusion that the SCRs might largely reflect tension levels, and not merely pain. This was confirmed by a non-significant correlation ($r=0.15$, $p=0.20$) between SCR and VAS pain severity difference scores after the effect caused by tension-change scores had been partialled out.

Table 9.2 Difference scores of skin conductance responses

Condition	Sub-group	mean	s.d.
Control	habituaors	-0.32	0.66
	stabilizers	-0.19	0.10
	sensitizers	-0.01	0.40
Up	habituaors	-0.22	0.18
	stabilizers	-0.09	0.31
	sensitizers	-0.11	0.33
Down	habituaors	-0.29	0.25
	stabilizers	-0.12	0.10
	sensitizers	-0.07	0.14

Table 9.3 Change of tension-level during experiment

Condition	Sub-group	mean	s.d.
Control	habituaors	33.3	19.1
	stabilizers	36.9	19.2
	sensitizers	50.0	21.4
Up	habituaors	32.2	16.8
	stabilizers	44.6	18.5
	sensitizers	34.6	19.1
Down	habituaors	33.1	11.1
	stabilizers	38.4	19.2
	sensitizers	42.8	20.5

The results of the magnitude matching are presented in Table 9.4. Visual inspection of the data clearly reveals that the difference scores along the sub-group sequence habituators, stabilizers and sensitizers, show an increasing trend. The hypothesized interaction effect (post-hoc hypothesis 3), however, was not significant ($t=0.16$, $p=0.44$). It was investigated by means of computing a partial correlation, whether tension-changes also affected the correlation between magni-

tude matchings and VAS-pain intensity scores. This was not the case, as is illustrated by the highly significant correlation: $r=0.478$, $p<0.0001$.

Table 9.4 Magnitude matching (relative difference scores) in milli-Ampere

Condition	Sub-group	mean	s.d.
Control	habituaors	-0.18	0.13
	stabilizers	-0.19	0.19
	sensitizers	-0.01	0.08
Up	habituaors	-0.07	0.14
	stabilizers	-0.07	0.45
	sensitizers	0.07	0.13
Down	habituaors	-0.22	0.12
	stabilizers	-0.10	0.11
	sensitizers	-0.01	0.15

Habituator groups

The testing of post-hoc hypothesis 2 resulted in a non-significant VAS-pain intensity difference interaction effect ($t=0.18$ $p=0.43$). The interaction effects relating to the SCR difference scores and magnitude matching differences did not reach significance either ($P_s >0.19$).

Table 9.5 Effects of rewards and punishments on VAS-score of subsequent trial

Condition	Sub-group	mean(mm.)	s.d.
Effect reward			
Up	habituaors	-6.52	13.61
	stabilizers	-2.30	6.31
	sensitizers	0.34	9.91
Down	habituaors	1.50	9.93
	stabilizers	5.28	9.92
	sensitizers	0.52	7.70
Effect punishment			
Up	habituaors	2.08	11.87
	stabilizers	3.15	7.43
	sensitizers	4.00	15.73
Down	habituaors	-8.83	17.51
	stabilizers	-6.48	15.00
	sensitizers	0.81	9.59

Finally, tests were carried out to investigate the effects of rewards and punishments on the VAS scores in the up- and down-conditioning sub-groups. The mean effects on the VAS-scales on the subsequent trial (in the conditioning phase) are given in Table 9.5. There are two groups of particular interest: the sensitizers in the up-conditioning group and the habituaors in the down-conditioning group.

The increase in the VAS scores of the sensitizers in the up-conditioning group can be largely explained by the effect of a punishment (on the average 4.00 mm) and, to a lesser extent, by the effect of a reward (0.34 mm increase). The increasing VAS scores of the down-conditioning sensitizer sub-group, on the other hand, were occasioned by the equal effects of rewards and punishments (0.52 and 0.81 mm respectively).

It is interesting that the habituators in the down-conditioning group show sharp decreases in VAS-pain intensity scores (8.83 mm) after punishments, whereas a reward in this sub-group generally leads to a small increase (1.50 mm). An inverse effect can be observed in the up-conditioning habituator sub-group. In sum, it may be said, then, that the effects detected in subjects who can be seen to have responded in the target direction, are principally caused by their reactions to punishments.

In the discussion section in Chapter 4 it was stated that the results of these post-hoc analyses have to be considered with caution.

Stringing all the results together, there is more evidence, at least within the present design, to refute than to support the notion that pain report can be operantly conditioned.

The scientific value of the post-hoc analyses has been that they have offered some guidelines for future research. The following recommendations are made:

1. Conditioning will probably be facilitated if a non-student group is selected (see page 52).
2. The presence of the experimenter (thus imitating in as far as possible the natural situation) may cause problems as to standardization. It is not impossible that different interactions with subjects might have profoundly negative consequences for standardization. A point also related to the experimenter, has to do with variables such as sex, age and physical appearance. Dependent upon the sex and age of the subject, a pompous, stern-looking 50-year-old male experimenter will probably elicit quite different responses to those elicited by a pretty 25-year-old women of gentle appearance.
3. The supposition that, in the present design, people who can easily be operantly conditioned are more shy, docile, submissive, etc., can be tested by using the appropriate questionnaires.
4. In the present design the people in the control condition were subjected to a neutral reinforcement system. This 'control' condition, however, was probably not the most 'pure'. A control group which was subjected to a random system of rewards and punishments would probably be preferable.

9.4 MPI-DLV patient sub-groups and pain mechanisms: an integration

The major impetus behind the MPI-DLV and the operant conditioning studies was the existence of 'non-optimal treatment results'. It was argued that the demarcation

of valid patient sub-groups, would facilitate the development of tailor made treatment programs, which should in turn lead to better treatment results. The aim of the other research line was to gain deeper insight into pain mechanisms - in this case, operant conditioning. The results of fundamental research into etiological and maintaining pain mechanisms can be used to modify treatment techniques.

The results described in Chapter 6 would seem to suggest that the aim of developing valid patient sub-groups has been achieved. Each of the MPI-DLV sub-groups are marked by specific features. It was stated earlier (e.g., page 28/29 or 73) that this descriptive psychosocial and behavioral information could be used in the setting-up of tailored cognitive-behavioral treatment programs. For instance, the essential characteristics of the Interpersonally Distressed profile would seem to indicate the need for a cognitive-behavioral treatment intervention in which special attention is paid to interpersonal communication / conflicts. This 'a vu' method of setting up a treatment module can, however, be criticized from a theoretical point of view. This can be illustrated by the following, perhaps rather extreme example. Consider the phenomenon fever, a state in which the body's thermoregulatory mechanisms behave as if they were adjusted to maintain body temperature at a higher than normal level. On the basis of this specific characteristic, it might be argued that a person suffering from fever should be placed in a refrigerated room for some time, in the hope that the temperature might go down. 'Mechanistic research', however, has revealed that fever is most commonly caused by a bacterial or viral infection. Without wishing to go into too much detail on the mechanism of fever, it is known that the use of antipyretic agents (e.g., aspirin) usually stops prostaglandin and, thereby, reduces body temperature. Clearly, then, experimental research on etiological and/or maintaining mechanisms provides a more solid basis for treatment strategies than does the 'a vu' method⁵. Interpreted in terms of chronic pain problems, and specifically the MPI-DLV sub-types, information on the question "What are the pathogenic mechanisms that promote the pain sub-types?" is very important. Investigation needs to be carried out to determine whether the MPI-DLV typology is related to different etiological and/or maintaining mechanisms⁶. It is like asking "Why does this patient type behave the way he does?"⁷ or "What factors predict that an acute pain patient will become a

⁵From a *clinical* and *practical* perspective, however, it might be argued that it is better not to wait until fundamental research has been completed.

⁶Even so, it might also be possible that the MPI-DLV patterns reflect idiosyncratic reactions to pain. If the latter were proved to be true, the treatment implications are less clear.

⁷This way of thinking is also applied in a completely different research field, that of bulimia nervosa. In discussing a biopsychological boundary model of bulimia nervosa (Herman and Polivy, 1984), Jansen criticizes the non-explanatory power of this model: "The boundary model is of a descriptive nature and does not explain why dieters regularly pass the diet boundary" (Jansen, 1990, page 26).

Dysfunctional chronic pain patient?". Research directed at unraveling these mechanisms is strongly recommended. In discussing future directions for chronic pain research Keefe (1992) states: "... research needs to identify the mechanisms by which behavioral and cognitive-behavioral interventions may alter pain behavior. Behavioral interventions may alter pain by changing the way a spouse responds to the patient's pain behavior. Cognitive interventions may work by altering biological mechanisms (e.g., endogenous opioid activation) that underlie pain perception. To identify mechanisms, behavioral researchers need to incorporate a broader range of behavioral, perceptual, and biological measures into their studies"⁸.

Based on the distinguishing features of the MPI-DLV clusters, hypotheses can be drawn up regarding the underlying mechanisms. These hypotheses can be tested in strictly experimental designs. Ideas regarding experimental research related to the MPI-DLV clusters are briefly discussed below.

9.5 Hypothetical underlying mechanisms for MPI-DLV patient clusters

The Dysfunctional type

As was outlined above, the essential features of the Dysfunctional type are the high level of pain report, psychological distress and environmental support, in combination with a perceived low level of control and a low general activity level. There seems, in addition, to be a high level of analgesic intake. The connection between these characteristics might be explained by operant conditioning factors. The high level of the MPI-DLV Pain Severity scale and low level of General Activity are probably related to frequent verbal pain complaints and other forms of pain behavior in the natural environment. High levels of environmental support and solicitous responses from the patient's partner may serve as potent reinforcers of the patient's pain behavior. It thus can be hypothesized that operant conditioning factors may control pain behaviors in Dysfunctionals more than in Interpersonally Distressed and Adaptive Coper patients. The results of some studies (Flor et al., 1987a; Gil et al., 1987) concur with the hypothesized operant conditioning mechanism in Dysfunctional patients. It should be noted, however, that in these two studies no analyses were carried out on MPI patient sub-groups. Rather, these studies provided evidence for (1) a positive relationship between solicitous responses from the patient's partner on the one hand and lower activity levels and higher pain severity reports on the other, and (2) the positive relationship between the level of pain behavior and the satisfaction with social support. Two experimental studies (Block et al., (1980); Lousberg et al., 1992) also support these findings. If the question to be answered is whether operant conditioning factors control pain behavior to a higher degree in Dysfunctional patients than in the other two MPI

⁸The phrase "altering biological mechanisms that underlie pain perception" refers to the work of Bandura et al. (1987).

patient types (in other words the link between Chapters 2 to 4 and 5 to 8), then the experimental designs used in the Block et al. (1980) and Lousberg et al. (1992) studies are to be preferred⁹. In addition, if, in future studies, the 'Linton conditioning technique' is capable to produce (valid) operant conditioning curves yet, the hypothesis can be tested that the conditioning curve (from VAS or other parameters) shows a different course in the Dysfunctional type.

The Interpersonally Distressed type

The MPI-DLV scale configuration of the Interpersonally Distressed type is roughly comparable to that of the Dysfunctional type, except on four scales - the Support scale and the three partner pain response scales. The low level of environmental and partner support and the partner's punishing responses to pain behavior most likely have a very negative effect on the level of psychological well-being of the Interpersonally Distressed type. The inability to reach a satisfactory resolution of problems (e.g., interpersonal conflicts, or more generally, an inability to cope with stress) could be an explanation. If this factor is already present, when the person is experiencing acute pain (a migraine attack, a disk herniation, etc.) tension levels may increase, which might in turn have a negative effect on the pain experience (see Chapter 1 on the antecedent control of pain). In time, a prolonged experience of pain and long-lasting (interpersonal) problems which cannot be controlled by the patient, may become sources of severe stress.

These hypothetical statements can be tested in an experimental design such as described by Flor et al. (1985). Three groups of subjects participated in this experiment: back pain patients, non-back pain patients and healthy controls. Subjects were given four tasks: the discussion of personal stress and pain, reciting the alphabet and mental arithmetic. Paraspinal and frontalis EMG, heart rate and skin resistance were recorded throughout. One of the most important findings in this study was that chronic back pain patients display elevations and delayed return to baseline, only in their paravertebral musculature and only when discussing personally relevant stress. (It is worth noting in passing that this study gave further support to the diathesis-stress model (Turk and Flor, 1984)). The implications for MPI-DLV sub-group research are clear. The Interpersonally Distressed type can be contrasted to the other three MPI-DLV types on the basis of stress responsiveness.

Further, research aimed at investigating the influence of stress in the work environment is to be recommended. Characteristics of the work environment are found to be a) predictive of the affective and evaluative dimensions of pain (Feuerstein et al., 1985) and b) a strong predictor of low back pain disability (Cats-Baril and Frymoyer, 1991). Another type of research which could be carried out on the Interpersonally Distressed type is related to the negative influence of various types of pain on interpersonal conflicts. Faucett and Levine (1991) found that the

⁹Unfortunately, at the time the experiment described in Chapter 2 was set up, only part two of the MPI was available. A computation of the MPI-DLV cluster membership was therefore not possible.

impact of conflicts in close relationships is stronger in patients suffering from a myofascial disorder than in arthritis patients. They argue that arthritis pain and related disability, in contrast to myofascial disorders, is widely validated and accepted, which probably leads to fewer conflicts and higher levels of support across the social network.

The Adaptive Copier type

Adaptive Copers display relatively low levels of pain, affective distress and interference by the pain with the daily life. They are the most active and there is evidence to suggest that they take the least amount of analgesic medication (after controlling for pain severity). It seems to be that these patients do possess a coping repertoire which enables them to deal with the pain adequately, as compared to the other three MPI-DLV patient clusters. The fact that these patients are able to deal with their pain does not mean that no experimental research is required in this area.

In the first place, experiments could be carried out to confirm the meaning of Adaptive Copier. The primary purpose of such experiments would therefore be one of validation. A great many studies have been carried out to investigate the effects of controllability and predictability on the pain experience. Control is defined as a belief that a person has at his disposal a response that can influence the aversiveness of an event (Thompson, 1981). Predictability of aversive events refers to a relationship between signs that may signal the occurrence of the event and the occurrence of the aversive event itself (Seligman, 1971). Several studies have reported evidence that perceived self-control over the (experimentally induced) pain stimulus results in increased pain tolerance and reduced pain reports (Bowers, 1968; Kanfer and Seidner, 1973; Rosebaum, 1980; Staub et al., 1971). Also, coping strategies such as neutral (mental arithmetic) distraction (Hodes et al., 1990), external (music) distraction (Anderson et al., 1991) and self-observation (Gilligan et al., 1984) have shown to be effective tools for increasing tolerance levels and reducing distress. (An extensive review is provided by Turk et al., 1983). The disturbing effects of an experimental manipulation of the predictability of pain intensity in healthy subjects, have been demonstrated by Arntz and co-workers. An unannounced sudden increase in the pain stimulus (in a series of pain stimuli) had long-term effects on fear of pain, uncertainty of predictions and expected aversiveness to the painful stimulus, and dishabituating effects on the experienced pain intensity and even on psychophysiological responses (Arntz and Hout, van den, 1988; Arntz and Lousberg, 1990). Further, these unannounced sudden increases in pain intensity level contributed to avoidance behavior (Arntz et al., 1991). The aforementioned studies have given weight to the idea that the level of pain experience can be modulated by means of psychological strategies. Less is known, however, about the mechanism by which they enhance pain tolerance. An intriguing laboratory study by Bandura et al. (1987) has shed more light on this issue. They set up an experiment to investigate the relationship between opioid involvement and perceived self-efficacy (Bandura, 1977a). Subjects were assigned to one of three conditions: a cognitive coping condition, a placebo condition and a

control condition. In each of the conditions one half of the subjects received an injection of naloxone (an opiate antagonist) and the other half, a placebo injection of saline. Pain tolerance was assessed by means of the cold pressure task. The results showed that a) the stronger the perceived self-efficacy to withstand pain, the longer subjects could endure the pain, b) cognitive copers who were given naloxone were less able to tolerate pain than were their saline counterparts and c) higher levels of perceived self-efficacy in reducing pain were accompanied by greater opioid activation. From this study it can be concluded that psychological factors such as controllability of pain by means of perceived self-efficacy affects neurophysiological systems (more specifically, endogenous opioids) which are assumed to play a direct and important role in the modulation of pain (cf. Frederickson, 1988). In spite of these obviously clear results, the debate on the positive effect of perceived control on endogenous opioid release (especially in chronic pain states) still continues (Arntz, 1991, page 154). In conclusion, if Adaptive Copers really do have a better coping repertoire, they must, in the experimental designs described above, behave differently to the other three types.

As to etiological/maintaining research, based on the MPI-DLV scale configuration of the Adaptive Coper profile, there is no 'obvious' hypothesis regarding an underlying mechanism. Research aimed at mapping out more fully the profile characteristics (see Chapter 8) may well suggest some directions. (The higher intelligence level of the Average type was, in fact, found in this way). Prospective studies (e.g., Philips and Grant, 1991; Bigos et al., 1990; Cats-Baril and Frymoyer, 1991) would probably also yield valuable information. A large sample of acute pain patients could provide information on several variables considered to be of importance, such as the ability to cope with stress, marital and work satisfaction, etc. There are at least two important research questions here. Can these variables predict chronicity and can they predict MPI-DLV cluster membership in chronic pain states?

In summary it may be said that one of the strengths of MPI-DLV sub-groups is that the (reliably obtained) grouping variable 'cluster membership' can be used in a large variety of experiments.

The Average type

In discussion on the underlying mechanisms for the Adaptive Coper, it was mentioned that there is no abundantly clear hypothesis. This is also true with respect to the Average type. This type seems to have no salient characteristics, apart from the higher intelligence level. Research testing the significance and implications of this feature requires special attention.

9.6 General conclusions

The undertaking to develop pain patient sub-groups originated in non-optimal treatment results. It is thought that if a reliable and valid system of sub-grouping can be

created, this will lead to better treatment results. The importance of research aimed at unravelling etiological and/or maintaining mechanisms before tailoring treatment procedures, has been stressed. This opinion is reinforced by the following example. Recently, Schulte and co-workers (1992) set up a well controlled experiment to test the more-effect of tailor-made versus a 'standard' therapy in phobic patients. There were three groups. The first group was characterized by an individualized tailored therapy, planned by the therapist. The second group was a control group undergoing a standardized - exposure in vivo - therapy (no tailoring) and the third was a yoked control group undergoing a variable standardized therapy. Surprisingly, it was not the individualized tailored therapy, but the standardized therapy which proved to be the most successful. It was concluded that "too much flexibility and too much adaptation can be disadvantageous, at least for the treatment of phobic patients" (Schulte et al., 1992). At a first glance this study seems to be a bad precursor for a future MPI-DLV sub-group tailoring study. A major point of criticism of the Schulte et al. study, however, is that it lacked 'mechanistic' research. That is, the individualized tailored treatment programs were based on a problem analysis of each single case. Although the tailored treatment started with hypotheses regarding the cause and/or maintaining mechanisms of the individual problem, the correctness of these hypotheses were **not** tested before the start of treatment.

In accordance with recent international ideas on pain assessment and treatment, a Multiaxial Assessment of Pain (MAP) approach and multi-model treatment of chronic pain is recommended. This line of thought is backed by an increasing number of studies reporting promising results in this field. Nevertheless, a large number of questions remain unanswered and this calls for international homogeneous pain assessment, marked by good communication and extensive collaboration between individual pain research laboratories, which would undoubtedly go a long way towards improving the quality of and rapid progress in pain research.

REFERENCES

- Affleck, G., Tennen, H., Urrows, S. & Higgins, P. (1991). Individual differences in the day to day experience of chronic pain: a prospective daily study of rheumatoid arthritis patients. *Health Psychology*, 10, 419-426.
- Anderson, R.A., Baron, R.S. & Logan, H. (1991). Distraction, control and dental stress. *Journal of Applied Social Psychology*, 21, 156-171.
- Armentrout, D.P., Moore, J.E., Parker, J.C., Hewett, J.E. & Felz, C. (1982). Pain patient MMPI subgroups: the psychological dimensions of pain. *Journal of Behavioral Medicine*, 1, 201-211.
- Arntz, A. & Hout van den, M. (1988). Generalizability of the match/mismatch model of fear. *Behaviour Research and Therapy*, 26, 207-223.
- Arntz, A. & Lousberg, R. (1990). The effects of underestimated pain and their relationship to habituation. *Behaviour Research and Therapy*, 28, 15-28.
- Arntz, A. (1991). *Pain: attention, emotion, prediction and control*. Academic thesis, Meppel (the Netherlands): Krips Repro.
- Arntz, A., van Eck, M. & Jong de, P. (1991). Avoidance of pain of unpredictable intensity. *Behaviour Research and Therapy*, 29, 197-201.
- Arntz, A., Hout van den, M., Meijboom, A. & Berg van den, G. (1991). The effects of incorrect pain expectations on acquired fear and pain responses. *Behaviour Research and Therapy*, 29, 547-560.
- Aronoff, G.M., Evans, W.O. & Enders, P.L. (1983). A review of follow-up studies of multidisciplinary pain units. *Pain*, 16, 1-11.
- Arrindel, W.A. & Ettema, J.H.M., SCL-90. (1986). *Handleiding bij een psychopathologie-indicator*. Lisse (the Netherlands): Swets & Zeitlinger.
- Bandura, A. (1977a). *Social learning theory*. London: Prentice-Hall.
- Bandura, A. (1977b). Self efficacy: toward a unifying theory of behavior change. *Psychological Review*, 84, 191-215.
- Bandura, A., O'Leary, A., Taylor, C.B., Gauthier, J. & Gossard, D. (1987). Perceived self-efficacy and pain control: Opioid and nonopioid mechanisms. *Journal of Personality and Social Psychology*, 53, 563-571.
- Bates, M.S., Edwards, W.T. & Anderson, K.O. (1993). Ethnocultural influences on variation in chronic pain perception. *Pain*, 52, 101-112.
- Beecher, H.J. (1956). Relationship of significance of wound to the pain experienced. *Journal of the American Medical Association*, 161, 1609-1613.
- Berge ten, J.M.F. (1986). Rotatie naar perfecte congruentie en de Multipiele Groep Methode. *Nederlands Tijdschrift voor de Psychologie*, 41, 218-225.
- Bernstein, D.A. & Borcovic, Th.D. (1973). *Progressive relaxation training*. Champaign Ill.: Research Press.
- Bernstein, I.H. & Grabin, C.P. (1983). Hierarchical clustering of pain patients' MMPI profiles: a replication note. *Journal of Personality Assessment*, 47, 171-172.

- Bigos, S.J., Battié, M.C., Spengler, D.M., Fisher, L.D., Fordyce, W.E., Hansson, T.H., Nachemson, A.L. & Wortley, M.D. (1991). A prospective study of work perceptions and psychosocial factors affecting the report of back injury. *Spine*, 16, 1-6.
- Black, R.G. (1975). The chronic pain syndrome. *Surgery Clinics of North America*, 55, 999-1011.
- Blanchard, E.B., Andrasik, F., Neff, D.F., Arena, J.G., Ahles, T.A., Juerish, S.E., Pallmeyer, T.P., Saunders, N.L. & Teders, S.J. (1982). Biofeedback and relaxation training with three kinds of headache: treatment effects and their prediction. *Journal of Consulting and Clinical Psychology*, 50, 562-575.
- Block, A., Kremer, E. & Gayler, M. (1980). Behaviour treatment of chronic pain: the spouse as a discriminative cue for pain behavior. *Pain*, 9, 243-252.
- Bowers, K. (1968). Pain, anxiety, and perceived control. *Journal of Clinical and Consulting Psychology*, 32, 569-602.
- Bradley, L.A., Prokop, C.K., Margolis, R. & Gentry, W.D. (1978). Multivariate analyses of the MMPI profiles of low back pain patients. *Journal of Behavioral Medicine*, 1, 253-272.
- Brennan, A.F., Barret, C.L. & Garretson, H.D. (1986). The prediction of chronic pain outcome by psychological variables. *International Journal of Psychiatric Medicine*, 16, 373-387.
- Butterworth, J.C. & Deardorff, W.W. (1987). Psychometric profiles of craniomandibular pain patients: identifying specific subgroups. *The Journal of Craniomandibular Practice*, 5, 225-232.
- Cats-Baril, W.L. & Frymoyer, J.W. (1991). Identifying patients at risk of becoming disabled because of low-back pain. *Spine*, 16, 605-607.
- Chapman, C.R., Chen, A.C.N. & Harkins, S.W. (1979). Brain evoked potentials as correlates of laboratory pain: a review and perspective. *Advances in Pain Research and Therapy*, 3, 791-803.
- Chapman, C.R. & Turner, J.A. (1990). Psychologic and psychosocial aspects of acute pain. In: Bonica, Lea & Fetger (Eds.), *The management of pain*. London.
- Clark, D.A. (1988). The validity of measures of cognition: a review of the literature. *Cognitive Therapy and Research*, 12 1-20.
- Cohen, J. & Cohen P. (1983). *Applied multiple regression/correlation analyses for the behavioral sciences*. London: Hillsdal, New Jersey.
- Conally, G.H. & Sanders, S.H. (1991). Predicting low back pain patients' response to lumbar sympathetic nerve blocks and interdisciplinary rehabilitation: the role of pretreatment overt pain behavior and cognitive coping strategies. *Pain*, 44, 139-146.
- Costello, R.M., Hulse, T.L., Schoenfeld, L.S. & Ramamurthy, S. (1987). P-A-I-N: a four cluster MMPI typology for chronic pain. *Pain*, 30, 199-209.
- Cott, A., Parkinson, W., Fabich, M., Bedard, M. & Marlin, R. (1992). Long-term efficacy of combined relaxation: biofeedback treatments for chronic headache. *Pain*, 51, 49-56.
- Craig, K.D. (1978). Social modeling influences on pain. In R.A. Sternbach (Ed.) *The Psychology of Pain*. New York, Raven Press.
- Craig, K.D. (1988). Consequences of caring: pain in the human context. *Canadian Psychologist*, 28, 311-321.

- Craufurd, D.I.O., Creed, F. & Jayson, M.I.V. (1990). Life events and psychological disturbance in patients with low-back pain. *Spine*, 15, 490-494.
- Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- Czikszenmihalyi, M. & Larson, R. (1987). Validity and reliability of the experience-sampling method. *Journal of Nervous and Mental Disease*, 175, 526-536.
- Demjen, S. & Bakal, D. (1986). Subjective distress accompanying headache attacks: evidence for a cognitive shift. *Pain*, 25, 194-197.
- Derogatis, L.R. (1983). SCL-90R Manual II. Clinical Psychometric Research. Maryland.
- Descartes, R. (1664). L'homme. Translated by M. Foster, *Lectures on the history of physiology during the 16th, 17th, and 18th centuries*, Cambridge University Press Cambridge, 1901.
- DeVries, M. (1987). Investigating the mental disorders in their natural settings. *Journal of Nervous and Mental Disease*, 175, 509-513.
- Duncan, G.H., Gregg, J.M. & Ghia, J.N. (1978). The pain profile: A computerized system for assessment of chronic pain. *Pain*, 5, 275-284.
- Eysenck, H.J. & Kamin, L. (1981). *Intelligence: the battle for the mind*. London: The Macmillan Press.
- Faucett, J.A. & Levine, J.D. (1991). The contributions of interpersonal conflict to chronic pain in the presence or absence of organic pathology. *Pain*, 44, 35-43.
- Feuerstein, M., Sult, S. & Houle, M. (1985). Environmental stressors and chronic low back pain: life events, family and work environment. *Pain*, 22, 295-307.
- Feuerstein, M., Labbe, E. & Kuczmierczyk, A.R. (1986). *Health psychology: a psychological perspective*. New York: Plenum Press.
- Feuerstein, M., Papciak, A.S. & Hoon, P.E. (1987). Biobehavioral mechanisms of chronic low back pain. *Clinical Psychology Review*, 7, 243-273.
- Flor, H. & Turk (1984). Etiological theories and treatments for chronic back pain I: Somatic models and interventions. *Pain*, 19, 105-121.
- Flor, H., Turk, D.C. & Birbaumer, N. (1985). Assessment of stress-related psychophysiological reactions in chronic back pain patients. *Journal of Consulting and Clinical Psychology*, 53, 354-364.
- Flor, H., Kerns, R. & Turk, D.C. (1987a). The role of spouse reinforcement, perceived pain and activity levels of chronic pain patients. *Journal of Psychosomatic Research*, 31, 251-259.
- Flor, H., Turk, D.C. & Rudy, T.E. (1987b). Pain and families. II Assessment and treatment. *Pain*, 30, 29-45.
- Flor, H. & Turk, D.C. (1988). Chronic back pain and rheumatoid arthritis: predicting pain and disability from cognitive variables. *Journal of Behavioral Medicine*, 11, 251-265.
- Flor, H., Birbaumer, N. & Turk, D.C. (1990a). The psychobiology of chronic pain. *Advances in Behaviour Research and Therapy*, 12, 47-84.
- Flor, H., Rudy, T.E., Birbaumer, N. & Schugens, M.M. (1990b). Zur Anwendbarkeit des West Haven-Yale Multidimensional Pain Inventory im deutschen Sprachraum. *Der Schmerz*, 4, 82-87.
- Flor, H., Fydrich, T. & Turk, D.C. (1992). Efficacy of multidisciplinary pain treatment centers: a meta-analytic review. *Pain*, 49, 221-230.

- Flor, H., Behle, D.J. & Birbaumer, N. (1993). Assessment of pain-related cognitions in chronic pain patients. *Behaviour Research and Therapy*, 31, 63-73.
- Foekens, R. (1992). Rugklachten; waar zit nu de pijn? *Tijdschrift voor Verzekeringsgeneeskunde*, 30, 98-103.
- Fordyce, W.E., Fowler, R.S., Lehmann, J.F., DeLateur, B.J., Sand, P.L. & Trieschmann, R.B. (1973). Operant conditioning in the treatment of chronic pain. *Archives of Physical Medicine and Rehabilitation*, 54, 399-408.
- Fordyce, W.E. (1974). Pain viewed as learned behavior. In J.J. Bonica (Ed.) *Advances in Neurology*, vol 4, 415-422. New York: Raven Press.
- Fordyce, W.E. (1976). *Behaviour methods for chronic pain and illness*. St. Louis, MO: Mosby Company.
- Fordyce, W.E., Roberts, A.H. & Sternbach, R.A. (1985). The behavioral management of chronic pain: a response to critics. *Pain*, 22, 113-125.
- Fordyce, W.E., Brockway, J.A., Bergman, J.A. & Spengler, D. (1986). Acute back pain: a control-group comparison of behavioral vs traditional management methods. *Journal of Behavioral Medicine*. 9, 127-140.
- Fordyce, W.E., Roberts, A.H. & Sternbach, R.A. (1988). The behavioral management of chronic pain: a critique of a critique. *Pain*, 33, 385-387.
- Frederickson, R.C.A. & Chipkin, R.E. (1988). Endogenous opioids and pain: status of human studies and new treatment concepts. In: H.C. Fields & J.M. Besson (Eds.), *Progress in Brain Research*, Vol 77, 407-417.
- Gentry, W.D. & Bernal, G.A.A. (1977). Chronic pain. In: R.B. Williams and W.D. Gentry (Eds.), *Behaviour approaches to medical treatment*. Ballinger, Cambridge, MA, pp. 173-182.
- Gil, K., Keefe, F., Crisson, J. & Dalfsen van, P. (1987). Social support and pain behavior. *Pain*, 29, 209-217.
- Gilligan, R.M., Ascher, L.M., Wolper, J. & Bochachevsky, C. (1984). Comparison of three cognitive strategies in altering pain behaviors on a cold pressure task. *Perceptual and Motor Skills*, 59, 235-240.
- Glass, G., McGaw, B. & Smith, M.L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.
- Goldfried, M.R. (1977). The use of relaxation and cognitive relabeling as coping skills. In: R.B. Stuart (Ed.), *Behaviour Self-Management: Strategies, Techniques and Outcomes*, New York: Brunner / Mazel, 82-116.
- Graham, J. (1987). *The MMPI, A Practical Guide*. New York: Oxford University Press.
- Greene, R.L. (1980). *The MMPI, An Interpretive Manual*. New York: Grune & Stratton.
- Groenman, N.H., Vlaeyen, J.W.S., Bakker-Boerrigter, A., Eek van, H., Schuerman, J.A. & Snijders, A.M.J. (1988). Chronische pijn als aangeleerd gedrag. *Bijblijven - cumulatief geneeskundig nascholingsstelsel*, Bohn, Scheltema & Holkema (Eds.), 4, 18-30.
- Groenman, N.H., Dingemans, W.A., Köke, A.J.A. & Lousberg, R. (1990). Behandeling van chronische benigne pijn vanuit het perspectief van het drie-factoren model. *Bewegen en Hulpverlening*, 3, 257-266.
- Groenman, N.H., Rober, I.M., Reitsma, B., Tuymelaar-Koldenhof, C.T.H. & Lousberg, R. (1993). Patienten met chronische benigne pijn: nader onderzoek met de SCL-90. *NVB Pijnbulletin*, 3, 9-12.

- Guck, T.P., Meilman, P.W., Skultety, F.M. & Poloni, L.D. (1988). Pain patient Minnesota Multiphasic Personality Inventory (MMPI) subgroups: evaluation of long term treatment outcome. *Journal of Behavioral Medicine*, 11, 159-169.
- Hall, C.S. & Lindzey, G. (1978). Theories of personality. New York: John Wiley & sons.
- Hart, R. (1984). Chronic pain: replicated multivariate clustering of personality profiles. *Journal of Clinical Psychology*, 40, 129-132.
- Härkäpää, K., Järvikoski, A., Mellin, G., Hurri, H. & Luoma, J. (1991). Health locus of control beliefs and psychological distress as predictors for treatment outcome in low-back pain patients: results of a 3-month follow-up of a controlled intervention study. *Pain*, 46, 35-41.
- Herman, C.P. & Polivy, J. (1984). A boundary model for the regulation of eating. In A.J. Stunkard, & E. Steller (Eds.) *Eating and its disorders*. New York: Raven Press.
- Hilgard, E.R. (1975). The alleviation of pain by hypnosis. *Pain*, 1, 213-231.
- Hodes, R.L., Howland, E.W., Lightfoot, N. & Cleeland, C.S. (1990). The effects of distraction on responses to cold pressure pain. *Pain*, 41, 109-114.
- Hugdahl, K. (1981). The three systems-model of fear and emotion - a critical examination. *Behaviour Research and Therapy*, 19, 75-85.
- Hunter, M., Phillips, C., & Rachman, S. (1979). Memory for pain. *Pain*, 6, 35-46.
- Ignelzi, R., Sternbach, R. & Timmermans, G. (1977). The pain ward follow-up analyses. *Pain*, 3, 277-280.
- International Association for the Study of Pain (IASP) (1979). Pain terms: a list with definitions and notes on usage. *Pain*, 6, 249-252.
- Jacobson, E. (1938). Progressive relaxation. Chicago: University of Chicago Press.
- Jamison, R.N., Rock, D.L. & Parris, W.C.V. (1988). Empirically derived Symptom Checklist 90 subgroups of chronic pain patients: a cluster analysis. *Journal of Behavioral Medicine*, 11, 147-158.
- Jamison, R.N. & Brown, G.K. (1991). Validation of ourly pain intensity profiles with chronic pain patients. *Pain*, 45, 123-128.
- Jansen, A. (1990). *Binge eating, notes and data*. Academic dissertation, Maastricht (the Netherlands): Datayse.
- Jöreskog, K.G. & Sörbom, D. (1988). *Lisrel-7, a guide to the program and applications*. Gorinchem (the Netherlands): INC.
- Kanfer, F.H. & Seidner, M.L. (1973). Self-control: factors enhancing tolerance of noxious stimulation. *Journal of Personality and Social Psychology*, 25, 381-389.
- Keefe, F.J. & Block, A.R. (1982). Development of an observation method for assessing pain behavior in chronic low back pain. *Behavior Therapy*, 13, 363-375.
- Keefe, F.J. (1989). Behaviour measurement of pain. In: Chapman, C.R. & Loesser, J.D. (Eds.) *Issues in Pain Measurement*. New York: Raven Press.
- Keefe, F.J., Bradley, L.A. & Crisson, J.E. (1990). Behaviour assessment of low back pain: identification of pain behavior subgroups. *Pain*, 40, 153-160.
- Keefe, F.J. (1992). Behaviour and cognitive-behavioral approaches to chronic pain: recent advances and future directions. *Journal of Consulting and Clinical Psychology*, 60, 528-536.
- Kerns, R.D., Turk, D.C. & Rudy, T.E. (1985). The West Haven Yale Multidimensional Pain Inventory (WHYMPI). *Pain*, 20, 345-356.

- Kerns, R.D., Turk, D.C., Holzman, A.D. & Rudy, T.E. (1986). Comparison of cognitive-behavioral and behavioral approaches to the outpatient treatment of chronic pain. *The Clinical Journal of Pain*, 1, 195-203.
- Kiers, H.A.L. & Berge ten, J.M.F. (1989). Alternating least squares algorithms for simultaneous component analysis with equal weight matrices in two or more populations. *Psychometrika*, 54, 467-473.
- Kiesler, D. (1971). Experimental designs in psychotherapy research. In: Bergin, A.E. & Garfield, S.L. (Eds.) *Handbook of psychotherapy and behavior change: an empirical analysis*. New York: John Wiley, 36-74.
- King, S.A. & Snow, B.R. (1989). Factors for predicting premature termination from a multidisciplinary inpatient chronic pain program. *Pain*, 281-287.
- Kleinke, C.L. & Spangler, A.S. (1988). Predicting treatment outcome of chronic back pain patients in a multidisciplinary pain clinic: methodological issues and treatment implications. *Pain*, 33, 41-48.
- Kloot van der, W.A. & Vertommen, H. (1989). *De MPQ-DLV, een standaard Nederlandse versie van de McGill Pain Questionnaire*. Lisse (the Netherlands): Swets & Zeitlinger.
- Kole-Snijders, A.J.M. (1988). *Personal Communication*.
- Kores, R.C., Murphy, W.D., Rosenthal, T.L., Elias, D.B. & North, W.C. (1990). Predicting outcome of chronic pain treatment via a modified self-efficacy scale. *Behaviour Research and Therapy*, 28, 165-169.
- Lakatos, I. & Musgrave, A. (1974). *Criticism and the growth of knowledge*. Cambridge University Press.
- Lange, A., (1983). *Interactionele Probleem Oplossings Vragenlijst IPOV*. Deventer (the Netherlands): Van Loghum Slaterus.
- Lee, P.W.H., Chow, S.P., Lieh-Mak, K.C. & Wong, S. (1989). Psychosocial factors influencing outcome in patients with low-back pain. *Spine*, 14, 838-843.
- Lefebvre, M.F. (1981). Cognitive distortion and cognitive errors in depressed psychiatric and low back pain patients. *Journal of Consulting and Clinical Psychology*, 49, 512-525.
- Levey (1980). Measurements units in psychophysiology. In Martin & Venables P.H. (Eds), *Techniques in Psychophysiology* (pp. 597-628). New York: Wiley.
- Liebeskind, J.C. & Paul, L.A. (1977). Psychological and physiological mechanisms of pain. *Annual Review of Psychology*. 28, 41-60.
- Linton, S.J. (1982). A critical review of behavioural treatments for chronic benign pain other than headache. *British Journal of Clinical Psychology*, 21, 321-337.
- Linton, S.J. (1985). The relationship between activity and chronic back pain. *Pain*, 21, 289-294.
- Linton, S.J. (1986). Behaviour remediation of chronic pain: a status report. *Pain*, 24, 125-141.
- Linton, S.J. & Götestam, K.G. (1985). Controlling pain reports through operant conditioning: a laboratory demonstration. *Perceptual and Motor Skills*, 60, 427-437.
- Lipton, J.A. & Marbach, J.J. (1984). Ethnicity and the pain experience. *Social Science & Medicine*, 19, 1279-1298.
- Loesser, J.D. (1980). Low back pain. In J.J. Bonica (Ed.) *Research Publications: Association for research on nervous and mental disease: Vol. 58* (page 363-377). New York: Raven Press.

- Lousberg, R. (1989). *Unpublished report*.
- Lousberg, R. (1992). *Personal communication* with Dr D.C. Turk.
- Lousberg, R. (1993). *Personal communication* with Drs P. Delespaul.
- Lousberg, R., Schmidt, A.J.M. & Groenman, N.H. (1992). The relationship between spouse solicitousness and pain behavior: searching for more experimental evidence. *Pain*, 51, 75-79.
- Lousberg, R., Schmidt, A.J.M., Groenman, N.H. & Gielen, A.A.C.M. (1993a). Operant conditioning of pain experience. *Submitted for publication*.
- Lousberg, R., Schmidt, A.J.M. & Groenman, N.H. (1993b). Operant conditioning of pain experience: a failed replication. *Submitted for publication*.
- Lousberg, R., Groenman, N.H., Schmidt, A.J.M., Arntz, A. & Winter, F.A.M. (1993c). Psychometric properties of the Multidimensional Pain Inventory - Dutch Language Version (MPI-DLV). *Submitted for publication*.
- Lousberg, R., Groenman, N.H., Schmidt, A.J.M., Schouten, E.G.W. & Meijer, S. (1993d). Pain patient subgroups derived from the Multidimensional Pain Inventory Dutch Language Version (MPI-DLV). *Submitted for publication*.
- Lousberg, R., Schmidt, A.J.M., Groenman, N.H. & Vendrig, L. (1993e). Validating the Multidimensional Pain Inventory - Dutch Language Version using experience sampling data. *Submitted for publication*.
- Lousberg, R., Groenman, N.H. & Schmidt, A.J.M. (1993f). Profile characteristics of the MPI-DLV clusters. *Submitted for publication*.
- Love, A.W. & Peck, C.L. (1987). The MMPI and psychological factors in chronic low back pain: a review. *Pain*, 28, 1-12.
- Luteijn, F. & Ploeg van der, F.A.E. (1983). *Groninger Intelligentie Test*. Lisse (the Netherlands): Swets & Zeitlinger.
- March, H.W. & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First- and higher order factor models and their invariance across groups. *Psychological Bulletin*, 97, 562-582.
- McCreary, C. (1985). Empirically derived MMPI profile clusters and characteristics of low back pain patients. *Journal of Consulting and Clinical Psychology*, 53, 558-560.
- McGill, J.C., Lawlis, G.F., Selby, D., Mooney, V. & McCoy, C.E. (1983). The relationship of Minnesota Multiphasic Personality Inventory (MMPI) profile clusters to pain behaviors. *Journal of Behavioral Medicine*, 6, 77-92.
- McNemar, Q. (1969). *Psychological Statistics*. New York: John Wiley & Sons.
- Meijboom, A.C. (1992). *Klassieke Conditionering van pijn en analgesia: Een onderzoek naar de vraag of pijn klassiek te conditioneren is*. Doctoraalscriptie. Rijksuniversiteit Limburg, Maastricht (the Netherlands).
- Melzack, R. & Wall, P. (1965). Pain mechanisms: A new theory. *Science*, 50, 971-979.
- Melzack, R. (1975). The McGill Pain Questionnaire: major properties and scoring methods. *Pain*, 1, 277-299.
- Melzack, R. & Wall, P. (1982). *The challenge of pain*. New York: Penguin books.
- Merskey, H. (1986). Classification of chronic pain. Descriptions of chronic pain syndromes and definitions. *Pain*, suppl. 3, S1-S225.
- Ministerie WVC. (1991). *Notitie chronisch-ziekenbeleid*. Den Haag (the Netherlands).
- Moore, J.E., Armentrout, D.P., Parker, J.C. & Kivlahan, D.R. (1986). Empirically derived pain-patients MMPI subgroups: prediction of treatment outcome. *Journal of Behavioral Medicine*, 9, 51-63.

- Mowrer, O.H. (1939). A stimulus-response analysis of anxiety and its role as a reinforcing agent. *Psychological Review*, 45, 553-564.
- Murphy, M.A., Tosi, D.J. & Pariser R.F. (1989). Psychological coping and the management of pain with cognitive restructuring and biofeedback: a case study and variation of cognitive experimental therapy. *Psychological Reports*, 64, 1343-1350.
- Naliboff, B.C., Cohen, M.J., Swanson, G.A., Bonebakker, A.D. & McArthur, D.L. (1985). Comprehensive assessment of chronic low back pain patients and controls: physical abilities, level of activity, psychological adjustment and pain perception. *Pain*, 23, 121-134.
- Nathan, P.W. (1976). The gate-control theory of pain: A critical review. *Brain*, 99, 123-158.
- Nunnally, J.C. (1978). *Psychometric Theory*. New York: McGraw-Hill.
- Olesen, J. (1988). Classification and diagnostic criteria for headache disorders, cranial neuralgias and facial pain (Headache Classification Committee of the International Headache Society). *Cephalalgia*, 8 Suppl. 7, 1-96.
- Olesen, J. & Jensen, R. (1991). Getting away from simple muscle contraction as a mechanism of tension-type headache. *Pain*, 46, 123-124.
- Pauley, J.W. & Haskell, D.J. (1975). Treatment of migraine without drugs. *Journal of Psychosomatic Research*, 19, 367-374.
- Pavlov, I.P. (1927). *Conditioned reflexes*. London: Oxford University Press.
- Pearce, J. & Morley, S. (1989). An experimental investigation of the construct validity of the McGill Pain Questionnaire. *Pain*, 115-121.
- Pennebaker, J.W., (1982). *The psychology of physical symptoms*. New York: Springer-Verlag.
- Peters, M.L. (1992) *Chronic pain and pain perception*. Academic dissertation. Maastricht (the Netherlands): University Press.
- Philips, H.C. (1987). Avoidance behavior and its role in sustaining chronic pain. *Behaviour Research and Therapy*, 25, 273-279.
- Philips, H.C. (1989). Thoughts provoked by pain. *Behaviour Research and Therapy*, 27, 469-473.
- Philips, H.C. & Grant, L. (1991). The evolution of back pain problems: a longitudinal study. *Behaviour Research and Therapy*, 29, 435-441.
- Philips, H.C. & Jahanshahi, M. (1985). The effects of persistent pain: the chronic headache sufferer. *Pain*, 21, 163-176.
- Polatin, P., Gatchel, R.J., Barnes, D., Mayer, H., Arens, C. & Mayer, T.G. (1989). A psychosociomedical prediction model of response to treatment by chronically disabled workers with low-back pain. *Spine*, 14, 956-961.
- Popper, K.R. (1972). *Objective knowledge*. Oxford University Press.
- Prokop, C., Bradley, L., Margolis, R. & Gentry, W. (1980). Multivariate analysis of the MMPI profiles of patients with multiple pain complaints. *Journal of Personality Assessment*, 44, 246-252.
- Reading, A.E. (1984). Testing pain mechanisms in persons in pain. In: P.D. Wall & R. Melzack (Eds.) *Textbook of pain*. Edinburgh: Churchill Livingstone, 195-204.
- Roberts, A.H. & Reinhardt, L. (1980). The behavioral management of chronic pain: long-term follow-up with comparison groups. *Pain*, 8, 151-162.
- Robinson, M.E., Swimmer, G.I. & Rallof, D. (1989). The P-A-I-N MMPI classification system: a critical review. *Pain*, 37, 211-214.

- Rosenbaum, M. (1980). Individual differences in self-control behaviors and tolerance of painful stimulation. *Journal of Abnormal Psychology*, 89, 581-590
- Rudy, T.E. (1989). *Multiaxial assessment of pain: Multidimensional Pain Inventory. Computer Program User's Manual, version 2.0*. University of Pittsburgh.
- Rudy, T.E., Turk, D.C., Brena, S.F., Stieg, R.L. & Brody, M.C. (1990). Quantification of biomedical findings of chronic pain patients: development of an index of pathology. *Pain*, 42, 167-182.
- Sanders, S.H. (1985). The role of learning in chronic pain states. *Clinics in Anaesthesiology*, 3, 57-73.
- Santen-Hoeufft van, H.M.S. (1992). Fitnesstraining and biofeedback in fibromyalgia: a randomized controlled trial, *In preparation*.
- Schipper, E.W.M. (1991). *Persoonlijkheid en chronische pijn*. Doctoraalscriptie, Rijksuniversiteit Limburg, Maastricht (the Netherlands).
- Schmidt, A.J.M. (1985). Performance level of chronic low back pain patients in different treadmill test conditions. *Journal of Psychosomatic Research*, 29, 639-645.
- Schmidt, A.J.M. (1987). The behavioral management of pain: a criticism of a response. *Pain*, 30, 285-291.
- Schmidt, A.J.M. (1988). Reply to letter from Fordyce, Roberts and Sternbach. *Pain*, 33, 388-389.
- Schreurs, P.J.G., Van de Willige, G., Tellegen, B. & Brosschot, J.F. (1988). *De Utrechtse Coping Lijst UCL*. Lisse (the Netherlands): Swets & Zeitlinger.
- Schulte, D., Künzel, R., Pepping, G. & Schulte-Bahrenberg, T. (1992). Tailor-made versus standardized therapy of phobic patients. *Advances in Behaviour Research and Therapy*, 14, 67-92.
- Schwartz, D.P. & DeGroot, D.E. (1983). An approach to the psychosocial assessment of the chronic pain patient. *Current Concepts in Pain*, 1, 3-11.
- Seligman, M.E.P., Maier, S.F. & Solomon, R.L. (1971). Unpredictable and uncontrollable aversive events. In: Brush, F.R. (Ed.) *Aversive conditioning and learning*. New York: Academic Press.
- Skevington, S.M. (1983). Social cognitions, personality and chronic pain. *Journal of Psychosomatic Research*, 27, 421-428.
- Skinner, B.F. (1953). *Science and human behavior*. New York: MacMillan.
- Spence, S.H. (1989). Cognitive-behavior therapy in the management of chronic, occupational pain of the upper limbs. *Behaviour Research and Therapy*, 27, 435-446.
- Staub, E., Tursky, B. & Schwartz, G.E. (1971). Self-control and predictability: their effects on reactions to aversive stimulation. *Journal of Personality and Social Psychology*, 18, 157-162.
- Sternbach, R. (1974). *Pain patients: Traits and treatment*. London, Academic Press.
- Sternbach, R.A. (1978). *The psychology of pain*. New York: Raven Press.
- Stevens, S.S. (1956). The direct estimation of sensory magnitude-loudness. *The American Journal of Psychology*, 69, 1-25.
- Stove, D.C. (1973). *Probability and Hume's inductive Scepticism*. Oxford University Press.
- Suijker, L.M.A. (1989). *Vragenlijst: "Identificatie gedrags-modaliteiten voor chronische pijn-patienten*. Doctoraalscriptie Rijksuniversiteit Leiden (the Netherlands).
- Summers, J.D., Rapoff, M.A., Varghese, G., Porter, K. & Palmer, R.E. (1991). Psychosocial factors in chronic spinal cord injury pain. *Pain*, 47, 183-189.

- Swanson, D., Swenson, W., Maruta, T. & McPhee, M. (1976a). Program for managing chronic pain: I program descriptions and characteristics of patients. *Mayo Clinic Proceedings*, 51, 401-408.
- Swanson, D., Floreen, A. & Swenson, W. (1976b). Program for managing chronic pain: II short term results. *Mayo Clinic Proceedings*, 51, 409-411.
- Syrjala, K.L. & Chapman, C.R. (1984). Measurement of clinical pain: A review and integration of findings. *Advances in Pain Research and Therapy*, 7, 71-101.
- Tan, S. (1982). Cognitive and cognitive-behavioral methods for pain control: a selective review. *Pain*, 12, 201-228.
- Tatsuoka, M.M. (1988). *Multivariate analysis*. New York: Macmillan Publishing Company.
- Thompson, S.C. (1981). Will it hurt less if I can control it? A complex answer to a simple question. *Psychological Bulletin*, 90, 89-101.
- Tucker, L.R. (1951). *A method for synthesis of factor analysis studies* (Personnel Research Section Report No. 984), Washington, D.C., Department of the Army.
- Turk, D.C., Meichenbaum, D. & Genest, M. (1983). *Pain and behavioral medicine, a cognitive-behavioral perspective*. New York: Guilford Press.
- Turk, D.C. & Flor, H. (1984). Etiological theories and treatments for chronic back pain. II. Psychological models and interventions. *Pain*, 19, 209-234.
- Turk, D.C. & Rudy, T.E. (1986). Assessment of chronic pain states: MAP-ping out the terrain. *Journal of Pain and Symptom Management*, 1, 235-237.
- Turk, D.C. & Flor, H. (1987a). Pain > pain behaviors: the utility and limitations of the pain behavior construct. *Pain*, 31, 277-295.
- Turk, D.C., Flor, H. & Rudy, T.E. (1987b). Pain and families. I. Etiology, maintenance, and psychosocial impact. *Pain*, 1987, 3-27.
- Turk, D.C. & Rudy, T.E. (1988). Towards an empirically derived taxonomy of chronic pain patients: integration of psychological assessment data. *Journal of Consulting and Clinical Psychology*, 56, 233-238.
- Turk, D.C. (1990). Customizing treatment for chronic pain patients: Who, what, and why. *The Clinical Journal of Pain*, 6, 255-270.
- Turk, D.C. & Rudy, T.E. (1990). The robustness of an empirically derived taxonomy of chronic pain patients. *Pain*, 43, 27-35.
- Turner, J.A. & Chapman, C.R. (1981). Psychological interventions for chronic pain: A critical review. I. Relaxation and biofeedback. *Pain*, 12, 1-21.
- Turner J.A. & Chapman, C.R. (1982). Psychological interventions for chronic pain. A critical review. II. Operant conditioning, hypnosis and cognitive-behavioral therapy. *Pain*, 12, 23-46.
- Turner, J.A. & Clancy, S. (1988). Comparison of operant-behavioral and cognitive-behavioral group treatment for chronic low back pain. *Journal of Consulting and Clinical Psychology*, 58, 573-579.
- Vlaeyen, J.W.S., Eek van, H., Groenman, N.H. & Schuerman, J.A. (1987). Dimensions and components of observed chronic pain behavior. *Pain*, 31, 65-75.
- Vlaeyen, J.W.S., Geurts, S.M., Eek van, H., Snijders, A.J.M., Schuerman, J.A. & Groenman, N.H. (1989a). *Pijn Cognitie Lijst Experimentele versie handleiding*, Lisse (the Netherlands): Swets & Zeitlinger.

- Vlaeyen, J.W.S., Snijders, A.M.J., Schuerman, J.A., Bremer, J.J.C.B., Eek van, H. & Groenman, N.H. (1989b). Chronic pain and the three systems model of emotions: a critical examination. *Critical Reviews in Physical and Rehabilitation Medicine*, 1, 2, 67-76.
- Vlaeyen, J.W.S., Geurts, S.M., Eek van, H., Snijders, A.M.J., Schuerman, J.A. & Groenman, N.H. (1990a). What do chronic pain patients think of their pain: Development of a pain cognition questionnaire. *British Journal of Clinical Psychology*, 29, 383-394.
- Vlaeyen, J.W.S., Pernot, H., Eek van, H., Snijders, A.M.J., Schuerman, J.A. & Groenman, N.H. (1990b). Assessment of the components of observed chronic pain behavior: The checklist for interpersonal pain behavior (CHIP). *Pain*, 43, 337-347.
- Vlaeyen, J.W.S., (1991). *Assessment and treatment from a behavioral rehabilitation perspective*. Academic thesis, Lisse (the Netherlands): Swets & Zeitlinger.
- Waddell, G. (1982). An approach to backache. *British Journal of Hospital Medicine*, sept: 187-219.
- Wade, J.B., Dougherty, R.P., Hart, R.P., Rafii, A. & Price, D.D. (1992). A canonical correlation analysis of the influence of neuroticism and extraversion on chronic pain, suffering and pain behavior. *Pain*, 51, 67-73.
- Wall, P.D. (1979). On the relation of injury to pain. *Pain*, 6, 253-264.
- Watkins, R.G., O'Brien, J.P., Braugelis, R. & Jones, D.J. (1986). Comparisons of preoperative and postoperative MMPI data in chronic back patients. *Spine*, 11, 385-390.
- Watson, D. (1982). Neurotic tendencies among chronic pain patients: an MMPI item analysis. *Pain*, 14, 620-628.
- Weiner, B. & Graham, S. (1984). Attributional approach to emotional development. In: Izard, C.E., Kagan, J. & Zajonc, R.B. (Eds.). *Emotions, cognition and behavior*. London: Cambridge University Press.
- Weisenberg, M. (1975). *Pain: clinical and experimental perspectives*. St. Louis, MO: Mosby Company.
- Weisenberg, M. (1977). Pain and pain control. *Psychological Bulletin*, 84, 1008-1044.
- Winter, F.A.M. (1992). *Revalidatie voor chronische pijnpatienten*. Academic dissertation, Enschede (the Netherlands): Intermap.
- Wright, A. & Davies, I.A.I. (1985). The recording of brain evoked potentials resulting from intra-articular focused ultrasonic stimulation: a new experimental model for investigating joint pain in humans. *Neuroscience Letters*, 97, 145-150.
- Ziekenfondsraad (1992). *Farmacotherapeutisch Kompas*. Amstelveen (the Netherlands).
- Zola, I.K. (1966). Culture and symptoms - an analysis of patients presenting complaints. *American Sociological Review*, 31, 615-630.

SUMMARY

Chapter 1. The phenomenon of 'non-optimal treatment results' in chronic 'benign' pain problems has been the grounds for conducting the studies and experiments described in the present thesis.

The word benign refers to a situation in which a patient's pain behavior cannot, or cannot sufficiently, be explained by somatic pathology. This type of pain is often called psychogenic pain. Arguments have been brought forward favoring a multidimensional approach and treatment of pain, especially chronic pain. There are several theories concerning the development and maintenance of chronic benign pain. Three visions are discussed and their scientific merits are evaluated. The first is the operant model, which states that pain behavior is maintained by environmental contingencies. Next, the vision focusing on antecedent control of pain is discussed: the pain-tension cycle. It is assumed that the experience of pain will lead to a heightened level of muscle tension which in turn might lead to a worsening of the pain. Third, the cognitive-behavioral model of pain is discussed. In this perspective it is assumed that non-adaptive cognitions (e.g., catastrophizing thoughts) and pain-coping strategies have a negative effect on the experience of pain.

Practical treatment strategies are derived from these theoretical models. In **each** of these therapies the phenomenon occurs that some patients do benefit (a lot), whereas other patients with an almost identical medical status do not benefit at all (or much less). The questions for scientific investigation that logically follow concern how this finding can be explained and what kind of decisions have to be made in order to improve treatment results of chronic benign pain. In order to answer these questions two possibilities are mentioned: 1) Test (by means of experimental research) the theoretical assumptions which form the basis of the various treatments. More knowledge on etiological and maintaining mechanisms can lead to the modification of treatments (which ultimately may lead to better treatment results). 2) Identify sub-groups within the population of chronic benign pain patients and adjust the type of treatment according to the characteristics of the sub-group concerned.

These two strategies form the core of the present thesis. Three experiments are conducted in order to investigate whether both pain behavior and the experience of pain can be conditioned in an operant way. Four studies are performed to identify patient sub-groups.

Chapter 2 investigates whether the presence of a solicitous spouse functions as a discriminative stimulus for a patient's pain behavior. Twenty-two backpain patients participated twice in a treadmill test: once alone and once in the presence of the spouse. Time of walking, pain intensity level and number of heart beats were

measured. Based on the operant paradigm regarding the maintenance of pain behavior as well as on the results of previous studies, it was expected that in the presence of a solicitous spouse patients would report more pain, would have a shorter walking time and would exert themselves less physically. Spouse solicitousness was measured in two ways: from the patient's perspective as well as from that of the spouse. The results obtained on the basis of the solicitousness measure from the spouse's perspective were in accordance with the expectations: that is, in the presence of a solicitous spouse patients report more pain and walk for a shorter duration than patients with a relatively non-solicitous spouse. These effects were not found if spouse solicitousness was measured from the patient's perspective. However, the moderate internal consistency coefficient of the partner's solicitousness measure precluded drawing definite conclusions.

In Chapter 3 a replication experiment is presented to investigate whether patient's pain experience can be conditioned in an operant way. Forty-four healthy subjects were allocated to one of two conditions. Each subject received 12, previously calibrated, painful shocks of equal intensity. Subjectively reported pain intensity was measured by means of a Visual Analogue Scale (VAS). In the experimental group increases in reported pain intensity were reinforced verbally. Decreases of reported pain were 'punished' verbally by means of thinking-out-loud comments by the experimenter present. In the control group there was no conditioning. In addition, it was investigated whether the experience of pain was influenced also. To this end, both a pain-related psychological variable, 'magnitude matching', and a pain-related psychophysiological variable, 'Skin Conductance Response' (SCR), were measured. The results confirmed previous findings: There was a clear increase of reported pain in the experimental group. The increase of pain report could be explained chiefly by the punishment of decreases of reported pain. The results concerning the magnitude matching and the SCR indicated that the experience of pain can also be conditioned operantly.

Chapter 4. Given the results of the experiment in Chapter 3 a new experiment with two aims was carried out: to replicate the results of the study in Chapter 3 and to attempt to down-condition reported pain and the experience of pain. In order to rule out an alternative explanation for the results of the previous experiment - an attentional shift towards the pain - the verbal punishments were adjusted. There were three conditions. The first condition tried to condition reported pain and the pain experience upwards (to a higher level). The second condition was marked by down-conditioning and in the third condition no conditioning took place. Both the attempt at replication and down-conditioning of pain failed. It is thought that especially the modified verbal punishments led to the inconsistent results. Therefore, the explanation of the 'conditioning-effect' found in the experiment in Chapter 3, the result of an attentional shift towards the pain, remains plausible.

Chapter 5 describes a psychometric study. The main aim was to test the psychometric qualities of the MPI-DLV: The Dutch language version of the often-used Multidimensional Pain Inventory. This questionnaire has 61 items divided into three parts. The scales of the three parts refer to 1) several pain-related psychological variables; 2) the reactions of the patient's partner on his or her pain behavior; 3) the frequency with which a diversity of daily activities are performed. MPI-DLV data were available from 733 chronic pain patients. Apart from some small differences, the factor structure of the English version could be replicated. The results regarding the reliability and validity of the scales were good. The construct validity of the scale 'General Activity' was demonstrated by means of an experimental study. Patients with high scores on the scale General Activity performed better in a working-to-tolerance bicycle ergometer test than patients with low scores. The MPI-DLV factor structure of fibromyalgia and backpain patients clearly appeared to be invariant. The clinical utility of the questionnaire was underlined by the sensitivity of the MPI-DLV to measure effects due to therapeutic interventions.

The identification of pain patient sub-groups - by means of a 'K-means cluster analysis' - forms the core issue in Chapter 6. Although the initial attempt to replicate the three patient sub-groups as found by Turk and Rudy (1988) succeeded, validation analyses revealed contradictory results. A less restrictive, more explorative, analysis resulted in the identification of four pain patient sub-groups: Since three of these groups had very similar characteristics to those in the American population the cluster nomenclature was retained: the 'Dysfunctional', 'Interpersonally Distressed' and the 'Adaptive Coper' type. The Dysfunctional type is characterized by a high level of suffering (much pain, interference of the pain with daily activities, and a high level of affective distress), a high level of environmental support as experienced by the patient and a relatively low activity level. Besides the even somewhat higher level of suffering as compared with the Dysfunctional type, the typical aspect of the Interpersonally Distressed type is the low level of environmental support experienced. Compared with the other patient types, the patient's partner is more likely to react with punishing responses to pain behavior. Patients classified as Adaptive Coper report (relatively!) little pain, experience a high level of life-control, receive sufficient support from their environment and have a relatively high activity level. Compared with these three types, the fourth type is characterized by a mean score profile. Consequently, this type is called 'Average'. The cross-validation of the four subtypes in a second sample was successful.

Chapter 7. In almost all cases the MPI-DLV is administered once in the **clinical** situation, when patients are asked to make judgments regarding their pain. In fact, the MPI-DLV measures relatively stable (trait) psychological and behavioral aspects of pain. By means of an 'experience sampling' method the ecological validity of a number of scales (in other words the generalizability to the daily situation) was investigated. Each MPI-DLV scale was translated into an item which

reflected the core of that MPI-DLV scale as well as possible. These items were included in self-report booklets. The duration of the experience sampling procedure was 6 days. At 8 random moments during each day, the patient had to record the extent to which he or she agreed with the item. For 6 MPI-DLV scales significant correlations were found with the corresponding experience sampling item: Pain Intensity, Interference, Solicitousness, Punishing and Distracting Responses and the general activity subscale Household Chores. The scale Life Control had a marginally significant correlation with the corresponding experience sampling item. No significant correlations were found for the scales Affective Distress and General Activity. The significant correlations were interpreted as a sufficient ecological validity of the MPI-DLV scales. The reason for the non-significant correlations was presumed to lie in the fact that the relevant experience sampling items did not reflect clearly enough the contents of the MPI-DLV scale in question. In addition, regression analysis was used to investigate the relation between the MPI-DLV pain intensity item (Rate the level of your pain at the present moment.) and the identical experience sampling item. The result was that 58% of the variance of the experience sampling score could be explained by the MPI-DLV pain score. Furthermore, it appeared to be that low pain intensity scores on the MPI-DLV corresponded with somewhat higher scores in the experience sampling measure and high MPI-DLV scores with somewhat lower experience sampling scores.

The purpose of the three small studies described in Chapter 8 was to map out the following patient-profile characteristics: Analgesic medication usage, level of intelligence and several personality characteristics. As expected, patients classified as Dysfunctional were marked by a marginally significant, higher level of medication usage compared to the other MPI-DLV types. Compared with the Average type this effect was significant. Regarding intelligence level Average patients were found to be marginally significantly more intelligent than the other MPI-DLV types. Concerning the MMPI personality profile, the MPI-DLV types were significantly different. The MMPI profile of Interpersonally Distressed patients was marked by high scores on the neurotic triad and a tendency to a passive-aggressive personality structure. The MMPI profile of Dysfunctional and Average patients produced the so-called 'conversion-V'.

The first part of Chapter 9 contains some critical remarks concerning the MPI-DLV and the derived patient clusters. The most important ones are mentioned.

- 1) The value of experimental studies is stressed.
- 2) Since in the majority of cases patients in the studies were fibromyalgia patients, the generalizability to other pain diagnoses and pain-loci is problematic. A more representative population of patients is necessary.
- 3) The interpretation of a MPI-DLV profile should always be done 'in perspective' (i.e., in relation to the other three MPI-DLV profiles).
- 4) It should be noted that the cluster assignment is made on probabilistic grounds. The importance of integrating somato-medical aspects is mentioned.
- 5) Care should be taken with the results of the 'other' profile-characteristics described in Chapter 8. Especially

the finding of the heightened analgesic medication usage in Dysfunctional patients should be replicated. The exact meaning and the practical implication of the elevated intelligence level in Average patients is not clear. Further research is warranted.

Based on the study described in Chapter 4 it was concluded that the attempt to condition pain behavior and pain experience in an operant way failed. Some laboratory observations made the investigators think that the desired effect (up- or down-conditioning) in a number of subjects may have been achieved. A series of **post-hoc** analyses produced the following results. There are indications that up-conditioning of reported pain in the relevant sub-group was indeed successful. This effect was caused chiefly by the subject's reactions to punishing response consequences. Down-conditioning of the pain report could **not** be established in the sub-groups in question. In addition, the post-hoc analyses made clear that the SCRs correlate much less with the experience of pain than had been previously assumed, than with the subjective tension level. The final conclusion with respect to the studies in Chapters 3 and 4 is that, using the present experimental design, it is more likely that pain in healthy people cannot be conditioned than that the opposite conclusion can be drawn.

In the second part of Chapter 9 an attempt has been made to integrate the two research lines, the identification of patient sub-groups and (fundamental) experimental research. It is argued that, before developing and carrying out tailored treatment programs for the MPI-DLV clusters, information has to be obtained about what kind of etiological and pain-maintaining mechanisms are related to the MPI-DLV clusters. Based on the interpretation of some salient cluster characteristics, it is assumed that operant factors play an important role as a pain-maintaining mechanism in patients classified as Dysfunctional. Inadequate stress coping probably plays an important role in the maintenance of pain in the Interpersonally Distressed type. Concerning the Adaptive Copier and the Average type there are no clear indications with regard to underlying etiological and/or pain-maintaining factors. Finally, in order to improve the quality of pain research homogeneity of pain assessment is recommended.

SAMENVATTING

Hoofdstuk 1. De aanleiding tot het uitvoeren van de studies en experimenten die in dit proefschrift worden beschreven, vormt het fenomeen van niet optimale behandelresultaten bij chronische 'benigne' pijnproblematiek. Onder de term benigne wordt een pijnbeeld verstaan waarbij patiënt's pijngedrag niet, of in onvoldoende mate, kan worden verklaard aan de hand van somatische pathologie. Vaak wordt dit type pijn ook wel psychogene pijn genoemd. Er worden argumenten aangevoerd die een multidimensionele benadering en behandeling van (chronische) pijn ondersteunen. Er bestaan diverse theorieën over het ontstaan en de instandhouding van chronische benigne pijn. Drie visies worden besproken en op hun waarde beoordeeld. Als eerste wordt het operante model besproken waarin wordt verondersteld dat pijngedrag door omgevingscontingenties in stand wordt gehouden. Vervolgens wordt de visie besproken waarin antecedente controle van pijn centraal staat: de pijn-spanning-pijn cyclus. De ervaring van pijn wordt verondersteld te leiden tot een verhoogd spierspanningsniveau welk op zijn beurt weer zou leiden tot een verergering van de pijn. Als derde wordt het cognitief-gedragsmatig model van pijn besproken. Binnen deze visie gaat men er van uit dat niet-adaptieve cognities (bv. catastroferende gedachten) en pijn-coping strategieën een negatief effect hebben op de ervaren pijn.

Van deze theoretische modellen zijn praktische behandelstrategieën afgeleid. Bij elk van deze therapieën doet zich het fenomeen voor dat sommige patiënten (veel) baat hebben, terwijl andere patiënten met een zo goed als identiek medisch beeld, nauwelijks of geen baat hebben. De hieruit logisch voortvloeiende wetenschappelijke vragen zijn hoe deze bevinding te verklaren is en welke stappen moeten worden ondernomen om behandelresultaten van chronische benigne pijn te verbeteren. Ter beantwoording van deze vragen worden twee mogelijkheden genoemd:

- 1) Toetsing (door middel van experimenteel onderzoek) van de theoretische assumpties die aan de verschillende behandelingen ten grondslag liggen. Meer kennis over etiologische en instandhoudende mechanismen kan leiden tot het modifieren van behandelingen (met als eind-resultaat betere behandelresultaten).
- 2) Identificatie van subgroepen binnen de populatie van chronische benigne pijnpatiënten en aanpassing van de behandelingsvorm aan de kenmerken van de betreffende subgroep. In dit proefschrift staan beide strategieën centraal: Drie experimenten worden uitgevoerd waarin wordt onderzocht of zowel pijngedrag als de ervaren pijn op operante wijze geconditioneerd kunnen worden. Vier studies worden verricht teneinde patiënt subgroepen te identificeren.

In hoofdstuk 2 wordt onderzocht of de aanwezigheid van een bezorgde partner als een discriminatieve stimulus fungeert voor patiënt's pijngedrag. Tweeënveertig rugpijnpatiënten participeren tweemaal in een loopband test: éénmaal alleen en

éénmaal in aanwezigheid van de partner. Loopduur, mate van ervaren pijn en hartfrequentie worden gemeten. Op basis van het operante paradigma betreffende het instandhouden van pijngedrag alsmede op grond van de resultaten van vorige studies, werd verwacht dat in de aanwezigheid van een relatief bezorgde partner patiënten minder lang lopen, meer pijn rapporteren en zich fysiek minder inspannen. De bezorgdheid van de partner werd zowel vanuit het perspectief van de partner als dat van de patiënt gemeten. De resultaten gebaseerd op de bezorgdheidsmeting vanuit het perspectief van de partner zijn in overeenstemming met de verwachtingen. Namelijk: in aanwezigheid van een bezorgde partner rapporteren patiënten meer pijn, lopen minder lang en spannen zich fysiek minder in, vergeleken met patiënten met een relatief niet-bezorgde partner. Deze effecten werden niet gevonden wanneer patiënt's interpretatie van partner's bezorgdheid als uitgangspunt werd genomen. De matige interne betrouwbaarheidscoëfficiënt van de bezorgdheidsmeting vanuit het perspectief van de partner verhinderde echter het trekken van definitieve conclusies.

In hoofdstuk 3 wordt een replicatie-experiment gepresenteerd waarin onderzocht wordt of de door de patiënt ervaren pijn operant conditioneerbaar is. Vierenveertig gezonde proefpersonen werden verdeeld over twee condities. Elke proefpersoon kreeg 12 pijnlijke prikkels van een gelijke, tevoren gecalibreerde sterkte. Subjectief ervaren pijnintensiteit werd met een Visual Analogue Scale (VAS) gemeten. In de experimentele groep werd een stijging van pijnrapportage verbaal bekrachtigd; een daling werd verbaal 'gestraft' middels hard-op-denkende verbale uitingen van de aanwezige experimentator. In de controle groep vond geen conditionering plaats. Tevens is nagegaan of de ervaren pijn mede beïnvloed wordt. Hiertoe werden een pijn gerelateerde psychologische variabele, 'magnitude matching', en een pijn - gerelateerde psychofysiologische variabele, 'Skin Conductance Response' (SCR), gemeten. De resultaten bevestigden eerder opgedane bevindingen: er was een duidelijke stijging van de pijnrapportage in de experimentele groep. De stijging in pijnrapportage kon voornamelijk worden verklaard door het straffen van dalingen in pijnrapportage. De resultaten betreffende de magnitude matching en de SCR gaven aan dat ook de pijnervaring operant geconditioneerd kan worden.

Hoofdstuk 4. Gegeven de resultaten van het experiment in hoofdstuk 3 werd een nieuw experiment uitgevoerd met als tweeledig doel: replicatie van de resultaten van de studie in hoofdstuk 3 en een poging om pijnrapportage en pijnervaring omlaag te conditioneren. Teneinde een alternatieve verklaring voor de resultaten van het vorige experiment - een verschuiving van aandacht naar pijn - uit te sluiten, werden de verbale 'straffen' aangepast. Er waren drie condities: In de eerste conditie werd gepoogd de pijnrapportage en pijnervaring omhoog (tot een hoger niveau) te conditioneren. De tweede conditie kenmerkte zich door omlaag conditionering en in de derde conditie vond geen conditionering plaats. Zowel de poging vorige resultaten te repliceren als het omlaag conditioneren van pijn mislukte. De inconsistente resultaten werden geacht voornamelijk het resultaat te zijn van de

gemodificeerde verbale 'straffen'. De verklaring van het gevonden 'conditionerings-effect' van hoofdstuk 3 als zijnde een gevolg van attentie-verschuiving naar de pijn toe blijft derhalve plausibel.

In hoofdstuk 5 wordt een psychometrische studie beschreven. Het hoofddoel was toetsing van de psychometrische kwaliteiten van de MPI-DLV: de in het Nederlands vertaalde versie van de veel gebruikte Multidimensional Pain Inventory. Deze vragenlijst bestaat uit 61 items welke zijn verdeeld over drie delen. De schalen van de drie delen hebben respectievelijk betrekking op: 1) Diverse pijn-gerelateerde psychologische variabelen. 2) De reacties van patiënt's partner op diens pijngedrag. 3) De frequentie waarin diverse dagelijkse activiteiten worden uitgevoerd. Van 733 chronische pijnpatiënten waren MPI-DLV data beschikbaar. Op enkele kleine verschillen na kon de factorstructuur van de Engelstalige versie worden gerepliceerd. De resultaten betreffende de betrouwbaarheid en de validiteit van de schalen waren goed. De constructvaliditeit van de schaal 'General Activity' werd door middel van experimenteel onderzoek aangetoond. Hoogscorders op de schaal General Activity presteerden, vergeleken met laagscorders, beter in een 'working-to-tolerance' fiets ergometer proef. De MPI-DLV-factorstructuur van fibromyalgie- en rugpijn-patiënten is gebleken in hoge mate invariant te zijn. De aangetoonde sensitiviteit van de MPI-DLV voor veranderingen door therapeutische interventies onderstreept de klinische bruikbaarheid van de vragenlijst.

De identificatie van subgroepen pijnpatiënten - door middel van een 'K-means clusteranalyse' - staat centraal in hoofdstuk 6. Hoewel een aanvankelijke poging tot replicatie van de drie patiëntsubgroepen zoals gevonden door Turk en Rudy (1988) slaagde, produceerden validatie-analyses tegenstrijdige resultaten. Een minder restrictieve, meer exploratieve, analyse had als resultaat dat vier subgroepen pijnpatiënten konden worden geïdentificeerd: aangezien drie van deze subgroepen zo goed als dezelfde karakteristieken hadden dan die in de Amerikaanse populatie, werd de cluster nomenclatuur gehandhaafd: het 'Dysfunctional', 'Interpersonally Distressed' en 'Adaptive Coper' type. Kenmerkend voor het Dysfunctional type is de hoge 'lijdensdruk' (veel pijn, interferentie van pijn met dagelijkse activiteiten en een sterk verstoorde stemming), de hoge mate van ervaren steun door patiënt's omgeving en een relatief laag activiteiten niveau. Typerend voor het Interpersonally Distressed type is dat, naast een nog iets hogere lijdensdruk dan het Dysfunctional type, geen of weinig steun uit de omgeving wordt ervaren. Vergeleken met de andere patiënt types, reageert patiënt's partner vaker met het uiten van 'straffende responsen' op pijngedrag. De als Adaptive Coper geclassificeerde patiënt rapporteert (relatief!) weinig pijn, ervaart een hoge mate van levenscontrole, krijgt voldoende steun uit zijn omgeving en heeft een relatief hoog activiteitsniveau. Het vierde type kenmerkt zich door een, in vergelijking tot de zojuist genoemde drie types, gemiddeld score profiel. Derhalve is dit type 'Average' genoemd. De crossvalidatie van deze vier subgroepen in een tweede steekproef was succesvol.

Hoofdstuk 7. In bijna alle gevallen wordt de MPI-DLV in de **klinische** situatie afgenomen in de vorm van een eenmalige meting waarbij de patiënt gevraagd wordt uitspraken te doen over de pijn. Gemeten worden eigenlijk relatief stabiele (trait) psychologische en gedragsaspecten van pijn. Met behulp van een 'experience sampling' methode is de ecologische validiteit, met andere woorden de generaliseerbaarheid naar de dagelijkse situatie, van een aantal MPI-DLV schalen onderzocht. De oorspronkelijke items werden in verwoordingen, die zo goed mogelijk de kerninhoud van die van de MPI-DLV schalen weergeven, opgenomen in de self-report boekjes. De experience sampling procedure duurde 6 dagen waarbij de patiënt per dag op 8 willekeurige momenten de vragen beantwoordde. Voor 5 hoofdschalen en een subschaal van de MPI-DLV werden significante correlaties met het corresponderende experience sampling item gevonden, te weten de hoofdschalen Pain Intensity, Interference, Solicitous, Punishing and Distracting Responses en de activiteiten subschaal Household Chores. De schaal Life Control correleerde marginaal significant met het corresponderende experience sampling item. Geen significante correlaties werden gevonden voor de schalen Affective Distress en General Activity. De significante correlaties werden geïnterpreteerd als een voldoende ecologische validatie van de MPI-DLV schalen. De reden voor de niet-significante correlaties werd verondersteld te liggen in het feit dat de betreffende experience sampling items niet duidelijk genoeg de inhoud van de betreffende MPI-DLV schalen weerspiegelden. Daarnaast werd middels regressie-analyse de relatie onderzocht tussen het MPI-DLV pijnintensiteits item (Hoeveel pijn heeft U op dit moment?) en het identieke experience sampling item met als resultaat dat 58% van de variantie van de experience sampling pijnintensiteits score kon worden verklaard uit de MPI-DLV pijn score. Verder bleek dat lage pijnintensiteits scores op de MPI-DLV corresponderen met iets hogere scores van de experience sampling meting en hoge MPI-DLV pijn scores met wat lagere experience sampling scores.

In hoofdstuk 8 worden een drietal korte studies beschreven die tot doel hebben de volgende patiënt-profiel karakteristieken in kaart te brengen: Analgetisch medicatiegebruik, intelligentieniveau en diverse persoonlijkheidsaspecten. Naar verwachting bleek dat Dysfunctional geclassificeerde patiënten zich in vergelijking met de andere MPI-DLV types kenmerken door (marginaal significant) meer gebruik van analgetica. Vergeleken met het Average type was dit effect significant. Wat betreft intelligentieniveau werd gevonden dat Average patiënten (marginaal significant) intelligenter zijn dan de andere MPI-DLV types. Wat betreft (MMPI) persoonlijkheidsprofiel bleken de vier MPI-DLV types onderling significant verschillend. Het MMPI profiel van Interpersonally Distressed patiënten kenmerkt zich door hoge scores op de neurotische triade en een tendens tot een passief-agressieve persoonlijkheidsstructuur. Het MMPI profiel van Dysfunctional en Average patiënten vertoont de zogenoemde 'conversie-V'.

Hoofdstuk 9 start met een aantal kritische kanttekeningen aangaande de MPI-DLV en de hieruit afgeleide patiënten clusters. De voornaamste worden genoemd.

1) De waarde van experimenteel validatie onderzoek wordt benadrukt. 2) Aangezien het merendeel van de patiënten in deze studies fibromyalgie-patiënten betrof, komt de generaliseerbaarheid naar andere pijndiagnosen en pijnloci in het geding. Een meer evenwichtig samengestelde patiënt-populatie is noodzakelijk. 3) De interpretatie van een MPI-DLV profiel dient immer 'in perspectief' (in relatie tot de andere drie MPI-DLV profielen) te worden beoordeeld. 4) De op probabilistische gronden geënte wijze van clustertoekenning als ook de integratie van somato-medische aspecten behoren in acht te worden genomen. 5) Voorzichtigheid moet worden betracht met de overige profielkenmerken beschreven in hoofdstuk 8. De bevinding dat gebruik van analgetica met name bij Dysfunctional patiënten is verhoogd, behoeft replicatie. De precieze betekenis en praktische implicatie van het verhoogd intelligentieniveau bij Average patiënten is nog onduidelijk en vraagt om nader onderzoek. Op grond van de studie in hoofdstuk 4 werd geconcludeerd dat de poging pijngedrag en de pijnervaring operant te conditioneren was mislukt. Sommige laboratorium observaties hebben ertoe geleid dat de onderzoekers de indruk kregen dat het bij een aantal mensen toch gelukt was om het gewenste effect (omhoog dan wel omlaag conditionering) te bewerkstelligen. Een serie **post hoc** analyses hebben de volgende bevindingen opgeleverd. Er zijn aanwijzingen dat omhoog conditionering van pijn-rapportage in de betreffende subgroep gelukt is. Dit effect werd voornamelijk veroorzaakt door reacties van proefpersonen op straf-fende respons-consequenties. Omlaag conditionering van de pijn-rapportage kon bij de betreffende subgroepen **niet** worden aangetoond. Verder gaven de post-hoc analyses duidelijk aan dat de SCRs veel minder sterk, dan aanvankelijk verondersteld werd, met de ervaring van pijn gecorreleerd zijn, dan met het subjectief spanningsniveau. De uiteindelijke conclusie met betrekking tot de studies in hoofdstuk 3 en 4 luidt dat het meer aannemelijk is dat het niet mogelijk is om, gebruikmakend van het huidige experimentele design, bij gezonde mensen pijn operant te conditioneren, dan dat dit wel het geval zou zijn.

Vervolgens wordt in hoofdstuk 9 een poging gedaan om de twee onderzoekslijnen, zijnde identificatie van patiënt subgroepen en (basaal) experimenteel onderzoek, te integreren. Beargumenteed wordt dat, alvorens toegesneden behandelprogramma's voor de MPI-DLV clusters te ontwikkelen en uit te voeren, informatie moet worden verkregen over welke etiologische en pijn-instandhoudende mechanismen aan de MPI-DLV clusters verbonden zijn. Interpretatie van enkele saillante clusterkenmerken leidt tot de veronderstelling dat operante factoren met name bij Dysfunctional geclassificeerde patiënten een belangrijke rol spelen als pijn-instandhoudend mechanisme. Bij het Interpersonally Distressed type speelt waarschijnlijk inadequate stress-coping een belangrijke rol in de instandhouding van pijn. Wat betreft het Adaptive Coper en het Average type zijn er geen directe aanwijzingen voor onderliggende etiologische en/of pijn-instandhoudende factoren. Tenslotte: Teneinde de kwaliteit van het pijn-onderzoek te verbeteren, wordt homogeniteit van pijn-assessment aanbevolen.

DANKWOORD

Dit boekwerk is het resultaat van 4 jaren, met plezier!, ploeteren en zwoegen. De dissertatie zou simpelweg niet tot stand zijn gekomen zonder de hulp van diverse anderen.

Een goede thuis- en privébasis is van het grootste belang: Mia en Aimée, bedankt voor alle steun, medewerking en verder voor al het geduld dat jullie in de afgelopen jaren regelmatig hebben opgebracht. Birgit, onze vriendschap was een onmisbare bron van inspiratie bij het schrijven van het proefschrift.

Zeer geleerde Groenman, beste Nico. Aan jou ben ik veel dank verschuldigd. Jij bracht mij op de juiste momenten in contact met de juiste personen. Van de hoeveelheid MPI-DLV data die we in amper twee jaren hebben verzameld, sta ik nu nog versteld. Onze samenwerking was altijd plezierig; naast de kneepjes van het vak psychodiagnostiek bij chronische pijnpatiënten leerde je mij onder andere hoe met 'stress' om te gaan. Dit kwam af en toe goed van pas! Daarnaast heb je mij in minder gemakkelijke tijden immer gesteund. Nogmaals mijn dank. Overigens, de Tour de Mont Blanc loopt niet weg!

Zeer geleerde Schmidt, beste Ton. Dank wil ik je zeggen voor de vier prettige jaren waarin we gezamenlijk pijn-experimenteel onderzoek hebben gedaan. Jouw scherp en kritisch denkvermogen heeft de experimenten uitstekend doen verlopen en kwam de kwaliteit van mijn artikelen en proefschrift zeer ten goede. Naast het pijnonderzoek denk ik eveneens met veel plezier terug aan ons hypochondrie-project. Bovendien heb ik ervaren dat je niet alleen wetenschappelijke kwaliteiten hebt, ook op de fiets ben je een waar talent.

Hooggeleerde van den Hout, beste Marcel. Als promotor heb je je waar gemaakt als een bekwaam supervisor. In recordtijd retourneerde je mijn artikelen, voorzien van een zeer bruikbaar en opbouwend commentaar. Als geen ander kun je iemand, wanneer hij/zij met tegenslagen te maken heeft, opnieuw enthousiast maken voor het bedrijven van het spel der wetenschap: dank hiervoor!

Naast mijn promotor maakten Joost Bremer, Henk Schmidt, Huda Huijjer Abu-Saad, Frans Zitman en Theo Bouman deel uit van de beoordelingscommissie. Mijn dank voor de kritische beoordeling van het manuscript.

Zeer geleerde Arntz, beste Arnold, 'leermeester van het eerste uur'. Graag wil ik jou hartelijk danken voor de moeite en energie die je in mijn stage-periode en ook daarna hebt gestoken in het mij bijbrengen van data-verwerkings- en analysetechnieken. Daarnaast heb ik regelmatig dankbaar gebruik gemaakt van jouw vakinhoudelijke deskundigheid.

Zeer geleerde Winter, beste Frits. 't Roessingh is een ware data-goudmijn gebleken. Ik ben je zeer erkentelijk voor de mogelijkheid die je mij hebt geboden om de MPI-DLV data te verzamelen. Weledelgeleerde Wulferink, beste Anneke, dank wil

ik jou zeggen voor de contentieuze wijze waarop je de administratie en dataverwerking hebt uitgevoerd.

Dr T. Rudy, Dr D. Turk en Dr J. Vlaeyen zeg ik dank voor statistische adviezen en vruchtbare vakinhoudelijke discussies.

Mijn dank aan Erik Schouten: niet alleen voor alle statistische problemen, maar ook voor elk PC-euvel stond jij immer gereed.

Dorien Wolfs en Alan Ralston zeg ik dank voor het opzetten en uitvoeren van het 'D(orien)A(lan)R(ichel)T(on)' onderzoek; Frans Verstappen voor het beschikbaar stellen van het bewegingslab.

Gerrie van Wunnik en Theo van Aerts ben ik dank verschuldigd voor alle hulp in het gedragslab.

Drs Dingemans, beste Wout. ik ben je zeer erkentelijk voor het feit dat ik veel van de patiënten van de pijnwerkgroep heb mogen benaderen voor deelname aan een van de experimenten. Hierbij is de secretariële ondersteuning van Emilienne Christophe en Marjan Groeneweg onmisbaar geweest.

Van de afdeling Medische Psychologie wil ik Christianne Vertommen danken voor haar nauwgezette administratie.

Dank ook aan de stagiaires en onderzoeksassistenten die aan een of meer van de studies / experimenten hebben meegewerkt: Maud de Bruin, Anneke Gielen, Bart Koldenhof, Margreet Krijgsman, Suzan Meijer, Heike Mersch, Nicolle Sey, Elsebeth Schipper, Lucienne Suijker, Lex Vendrig en Gert Wolters.

Joyce Angenent, Isel van Noppen, Yvonne Stevens en Margriet van Woerden: Dank jullie wel voor alle secretariële hulp (Isel, ook mijn dank voor de laatste puntjes op de i). De gezellige sfeer maakt(e) het zo goed als onmogelijk het secretariaat te passeren zonder het maken van een vlotte babbel.

Om de experimenten die in dit proefschrift worden beschreven uit te kunnen voeren, was de medewerking van patiënten en proefpersonen natuurlijk onmisbaar. Mijn dank is groot.

Margret Jones wil ik danken voor de correctie van de Engelse tekst van het manuscript.

Een laatste woord van dank wil ik richten aan mijn plotseling overleden vader. Beste Paul, helaas kan ik je het eind-resultaat van de AIO-job, onderliggend boekje, niet persoonlijk overhandigen. Hoe dan ook, mijn dank voor je betrokkenheid en alle zorgen!

Richel.

CURRICULUM VITAE

Richel Lousberg werd geboren op 18 mei 1966 te Sittard. In 1984 behaalde hij het eindexamen gymnasium- β . In datzelfde jaar begon hij aan de studie Gezondheids-wetenschappen, afstudeerrichting Geestelijke Gezondheidskunde, aan de Rijksuniversiteit Limburg. Deze studie ronden hij in mei 1989 af.

Van juni 1989 tot juni 1993 was hij als assistent in opleiding verbonden aan de vakgroep Medische Psychologie van de Rijksuniversiteit Limburg.

In september 1993 trad hij in dienst bij de vakgroep Onderwijsontwikkeling en Onderwijsresearch van de Rijksuniversiteit Limburg.

