IMPROVING NON-OPTIMAL RESULTS IN CHRONIC PAIN TREATMENT

A TRIPARTITE APPROACH

Carola A.J. Mes

Address of correspondence:

Carola Mes
Roessingh Research & Development
PO Box 310
7500 AH Enschede
The Netherlands
+31 53 4875777
c.mes@rrd.nl

Cover image provided courtesy of Medtronic, Inc.
Printed by Gildeprint Drukkerijen B.V., www.gildeprint.nl, Enschede

ISBN 90-365-2435-0

IMPROVING NON-OPTIMAL RESULTS IN CHRONIC PAIN TREATMENT

A TRIPARTITE APPROACH

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof.dr. W.H.M. Zijm,
volgens besluit van het College voor Promoties
in het openbaar te verdedigen
op vrijdag 12 januari 2007 om 15.00 uur

door

Carola Adriana Johanna Mes
geboren op 17 januari 1973
te Eindhoven

Dit proefschrift is goedgekeurd door de promotoren:

Prof. dr. G. Zilvold
Prof. dr. D.C. Turk

En assistent promotor:
Dr. R. Lousberg

De promotiecommissie is als volgt samengesteld:

Voorzitter en secretaris:	
Prof.dr. P.H. Hartel	Universiteit Twente
Promotoren:	
Prof. dr. G. Zilvold	Universiteit Twente
Prof. dr. D.C. Turk	University of Washington, Seattle, USA
Assistent-promotor:	
Dr. R. Lousberg	Universiteit Maastricht, Roessingh Research & Development, Enschede
Leden:	
Prof.dr. M.I. Hasenbring	Ruhr-University Bochum, Germany
Prof.dr.ir. H.J. Hermens	Universiteit Twente
Prof.dr. W.H. van Harten	Universiteit Twente
Dr. P.A.E.G. Delespaul	Universiteit Maastricht
Dr. M.M.R. Vollenbroek-Hutten	Roessingh Research & Development, Enschede
Referent:	
Dr. K.M.G. Schreurs	Revalidatiecentrum Het Roessingh, Enschede

Paranimfen:
Anke Kottink
Yvette Bulthuis

This study was supported by grants from the Nardy Roeloffzen Foundation.

The publication of this thesis was generously supported by:

- Roessingh Research & Development, Enschede;
- Biomedical Signals and Systems group, University of Twente, Enschede;
- The Dr. G.J. Van Hoytema Foundation, Enschede. The aim of this foundation is mainly to promote the education of health care professionals, including the financial support of PhD students and academic chairs.

CONTENTS

OUTLINE OF THESIS

Chronic pain is a major medical and social problem in the industrialized world. In his review of the epidemiology of chronic pain, Bonica summarized that "... more than a third of the American population has chronic painful conditions, and of these, 50 - 60% are partially or totally disabled for periods of days (e.g., recurrent headaches), weeks (e.g., reflex sympathetic dystrophies, myofascial syndromes), months (e.g., low back pain), and some permanently (e.g., arthritis) ..."[1]. In The Netherlands, 44.4% of the population (>= 25 years old) have musculoskeletal pain complaints[2]. Evidently, chronic pain contributes to disability on a large scale. Additionally, chronic pain is expensive[3]. Increased health care consumption, loss of income, inability to work, compensation and litigation all incur costs that make the problem of chronic pain a pricey one. For these reasons, it is important to discover how this major problem can best be treated.

Based on the available scientific knowledge about the causes of chronic pain and mechanisms that maintain it, numerous methods of pain treatment have been developed. Despite this, many patients still do not completely recover from their complaints, regardless of the kind of treatment applied. Although the *mean* result of a given treatment may be positive from both a statistical and clinical point of view, a relatively large range around this average estimate is often observed. In practice, this means that within a 'successfully' treated group there are also patients who report no improvement at all or even a further deterioriation in health. From a "good/optimal" treatment, the expectation is an improvement in *all* patients, at least to a certain extent. Cases where a treatment fails to show improvement in all patients may be regarded as non-optimal treatment results.

Treatment results reported as non-optimal may be due to several causes[4]. One reason may be a lack of fit between a standard treatment protocol to be offered to all patients and the needs and characteristics of individual patients or patient subgroups. It is also possible that the variability in treatment results is partially caused by program failures or theory failures. A third possibility is that there is an insufficient amount of basic scientific knowledge about the causes of chronic pain and the mechanisms by which it is maintained. Obviously, a treatment program based on insufficient knowledge is unlikely to succeed in obtaining an optimal result for all patients.

The studies described in this thesis are related to the three mentioned explanations for non-optimal treatment results. A general overview of chronic pain and its treatment is presented in Chapter 1. In Chapter 2, the effectiveness of a multidisciplinary pain management program, conducted at the Roessingh Center for Rehabilitation in The Netherlands, is evaluated. The results of this study will be used to demonstrate the phenomenon of non-optimal treatment results. Chapter 3 describes the ways in which variability in results can be explained by the existence of patient subgroups. Whether the variability in treatment responses is related to possible program or theory failures is investigated in Chapter 4. In Chapters 5, 6 and 7, the results of an experiment focusing on expectations and experiences of pain for chronic pain patients in their daily living situations are discussed. This discussion endeavours to explore the mechanisms by which chronic pain is maintained. Lastly, Chapter 8 presents a general discussion concerning a) the extent to which the different studies succeed in explaining treatment variability; and b) how the findings may contribute to improving future treatment results.

References

1. Bonica, J.J. (1990). *The Management of Pain*, (2nd Ed). Philadelphia: Lea and Febiger.
2. Picavet, H.S.J., Gils, H.W.V. van, & Schouten, J.S.A.G. (2000). *Klachten van het bewegingsapparaat in de Nederlandse bevolking: prevalenties, consequenties en risicogroepen*. RIVM rapport 266807 002. Bilthoven: RIVM i.s.m. CBS.
3. Loeser, J.D. (1999). Economic implications of pain management. *Acta Anaesthesiol Scand, 43,* 957-959.
4. Lousberg, R. (1994). *Chronic pain: Multiaxial diagnostics and behavioral mechanisms*. Maastricht: University Press.

CHAPTER 1
General introduction

1.1 What is (chronic) pain?

Until the 1960s, pain was viewed as a simple sensory response to tissue damage. With the gate control theory of Melzack and Wall, this perspective has been changed[1]. Nowadays, it is generally agreed that pain can also exist in the absence of tissue damage. Pain is defined by the International Association for the Study of Pain (IASP) as "an unpleasant sensory and emotional experience, associated with actual or potential tissue damage, or described in terms of such damage"[2]. This definition implies that pain may exist even when no physical substrate can be demonstrated. There is no such thing as *the* pain. Moreover, the IASP definition states that pain is a *subjective* experience: each individual experiences and interprets pain in his or her own way.

In the acute phase of pain, the experience of pain is seen as functional and can be considered as a normal response to tissue damage. In most cases, this pain will subside in accordance with the healing of damaged tissue. If this does not happen, one may consult health care practitioners in order to find relief. In such cases, a diagnosis is made and a therapy is prescribed. For most patients, the pain then disappears during the therapy. In some cases, pain persists. This may be due to persistent noxious stimulation by chronic pathological processes in somatic structures or viscera, by prolonged dysfunction of parts of the peripheral or central nervous system, or by both (e.g. cancer pain). However, in some cases, pain persists even after tissue damage has been resolved and no physical substrate can be demonstrated. In other cases, a large discrepancy between patients' pain complaints and underlying somatic pathology exists. Both cases are illustrative of a type of pain called chronic pain. Chronic pain is defined by the IASP as 'pain that persists beyond the normal time of healing'[3]. Unfortunately, there is no general consensus on the period of time the pain must persist before it can be considered chronic. Until recently, a time period of 6 months was the norm. Today, a shift towards the acceptance of shorter period can be seen. The IASP has created categories for classification that include the periods of less than one month, one to six months, and greater than six months[4].

Today, there is international agreement about the multidimensional nature of chronic pain. There is overwhelming evidence that chronic pain is influenced by not only physiological, but also psychological, social and cultural factors[1,5,6,7,8,9,10,11,12]. Accordingly, procedures for intake and treatment of chronic

pain patients, as well as the measurement of pain patients' functioning and outcome of treatment, should be of a multidimensional nature.

1.2 The need for a multidimensional approach: mechanisms of chronic pain

The fact that chronic pain is a multidimensional phenomenon implies that theoretical models explaining the development and maintenance of chronic pain should also be multidimensional. In this section, several multidimensional, biobehavioral models for the development of chronic pain, are presented.

1.2.1 The diathesis-stress model of chronic pain

The diathesis-stress model of pain, developed by Flor, Birbaumer and Turk, describes how physically and mentally stressful events may lead to chronic pain through the hyperactivity of the back muscles[9]. Using a respondent conditioning model, they assume that pain leads to reflex muscle spasm and to sympathetic activation. Over time, these processes become classically conditioned to otherwise harmless stimuli, thereby producing pain[13]. In a study by Flor, Turk and Birbaumer, a group of patients with back pain, a group of patients with other pain syndromes, and a group of healthy individuals were exposed to two personally relevant stressors, as well as to a general and a neutral stressor, while the electromyograhic (EMG) level of their back muscles was measured[14]. The data show that the majority of the back pain patients reacted to the stimuli with strong increases in EMG but only in the case of personally relevant stressors. No effects were found with respect to the general stressor or the neutral stressor. The other two groups showed minimal reactions to all situations. The results suggest that a specific relationship exists between personally relevant stress, muscle tension and back pain.

1.2.2 The fear-avoidance model of pain

The fear-avoidance model of pain was first introduced by Lethem et al.[15] and further elaborated upon by Vlaeyen et al. (Figure 1.1)[16]. The fear-avoidance model assumes that an individual with acute pain will tend to reduce or avoid physical activity in order to stimulate tissue healing. This theory contends that some patients maintain avoidance behavior even after the body tissue is healed. A patient may no longer perform certain activities because he fears that these activities will increase pain and suffering. Later on, this protective pain avoidance behavior may persist in

anticipation or expectation of pain, instead of as a response to it[16,17]. This is relevant as the avoidance of physical activity can have significant detrimental consequences in the long term[18,19]. Physical consequences may include the loss of mobility, decreased muscular strength, and reduced physical condition. Psychological consequences may include depression or a loss of self-esteem. Such consequences are likely to ensure that the pain experienced by the patient becomes more severe and thus reinforce further avoidance behaviour, thereby creating a negative cycle of avoidance and reinforcement[15].

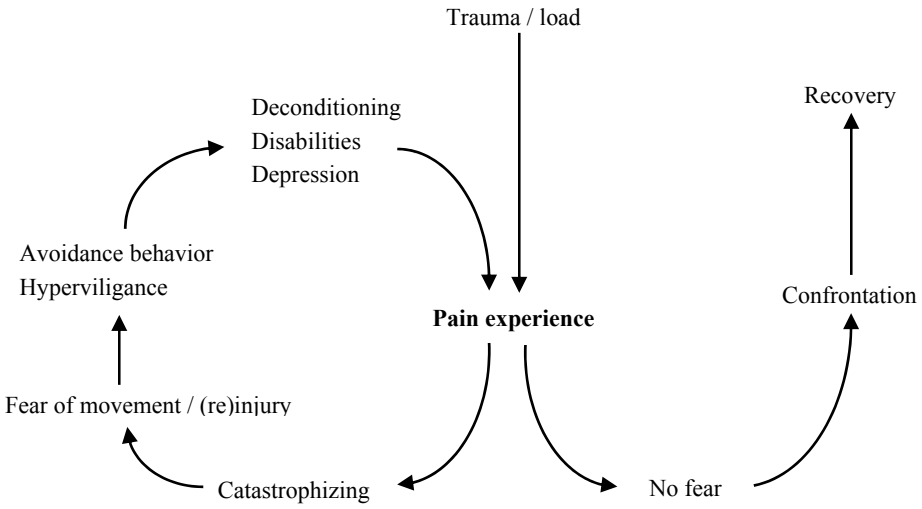


Figure 1.1: The fear-avoidance model according to Vlaeyen et al.[16]

1.2.3 The avoidance-endurance model of pain

The avoidance-endurance model of chronic pain (Figure 1.2) is presented as an expansion of the fear-avoidance model of pain[20,21]. In addition to to the established pathway from catastrophizing through avoidance behavior and muscle insufficiency towards chronic pain, they also distinguished a pathway from minimalizing thoughts through avoidance behavior and muscle hyperactivity towards chronic pain. A study from Hasenbring et al. indicated that the assessment of pain-related coping modes yielded an important differentiation between subgroups of low back pain patients six months after surgery[22]. Patients adhering to an endurance coping strategy displayed more signs of overuse in their daily

behavior in spite of pain. The one fear avoidance coper in their study tended to do less physical activity, which may be regarded as underuse.

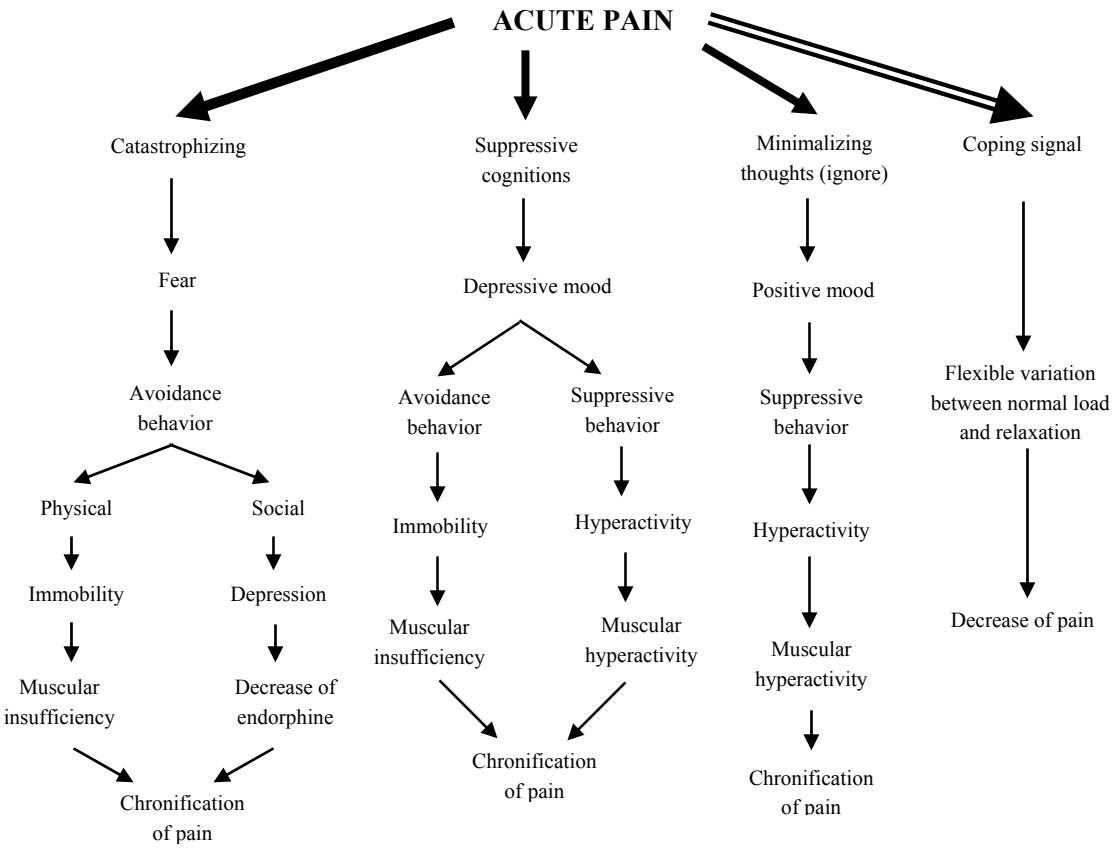


Figure 1.2: The avoidance-endurance model according to Hasenbring et al.[20,21]

1.2.4 The match-mismatch model of pain

A model closely related to the fear-avoidance model and avoidance-endurance model is the match-mismatch (MM) model of pain, which is based on the MM model of fear[23,24]. The MM model of fear hypothesizes that an underestimation or underprediction of the experienced fear expected in a fear provoking situation has a negative effect. Arnzt and Van den Hout have explored the generalizability of this

model in an attempt to create a more general model of human processing of aversive experiences[25]. The MM mechanism contends that the most important immediate consequence of a mismatch between expected and experienced intensity of a fearful or painful event is a shift of the subsequent pain expectation in the direction of the last pain experience (Figure 1.3). For example, when an event is more aversive than expected (underprediction), the expectation for the next event is increased. Vice versa, when the event is less aversive than expected (overprediction), the future expectations probably will be decreased. When the expectation is accurate, the expectations will most probably remain constant. These hypothetical mechanisms have been demonstrated in a series of experiments by Arntz and coworkers[26,27,28]. These studies show that indeed the most important immediate consequence of a mismatch between expected and experienced pain is an adaptation of the expectation for the next experience in the direction of the last experience. Among others, a study by Arntz et al. have pointed out the negative influence mismatches may have on avoidance behavior and increases in the experience of pain[29].

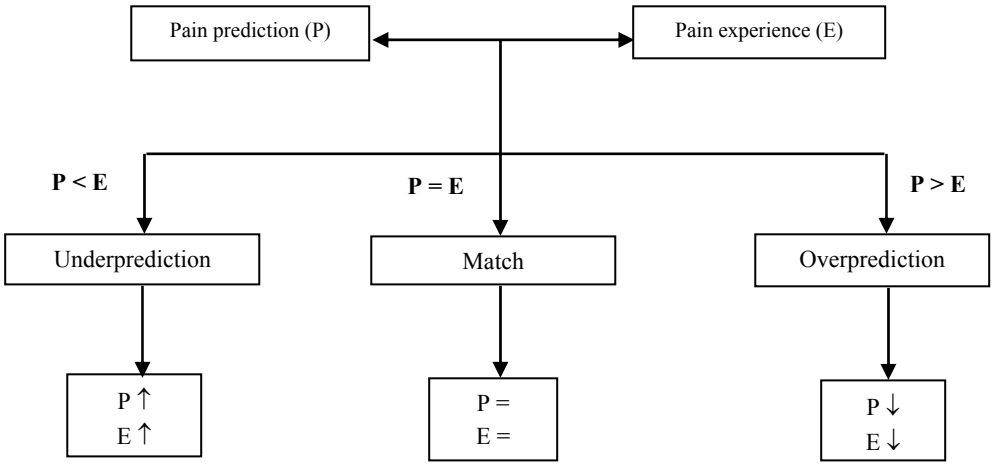


Figure 1.3: Schematic representation of the match-mismatch mechanism of pain

1.2.5 Recent advances

The use of certain measurement techniques that attempt to objectivate chronic pain is, nowadays, becoming increasingly popular in multidimensional research on the underlying mechanisms of chronic pain. One of these techniques is the surface electromyography (sEMG). This technique has been applied by Flor and co-workers[13,14,30,31]. In a study focusing on the influence of enhanced aversive conditioning, Schneider, Palomba and Flor demonstrated enhanced muscular response in chronic pain patients. They suggested that a dissociation of muscular and central processes during aversive conditioning in the patients may contribute to the problem of chronicity[31]. EMG measurements have also been applied in studies of the 'Neuromuscular assessment of the Elderly Worker (NEW)' project. The NEW project seeks to develop a better understanding of the relationship between muscle activation patterns and muscle properties in musculoskeletal disorders (particularly neck and shoulder region) experienced by elderly workers and psychosocial and work characteristics[32,33,34]. So far, these studies have shown clear differences and similarities between cases and controls with regard to both physiological (EMG activation) and psychological variables.

Recent studies that relate physiological measures to psychological mechanisms of pain have also provided encouraging results by measuring central nervous processing. For example, Lousberg et al. investigated the relationship between pain-related psychosocial aspects and Event Related Potentials (ERPs) in the central nervous system[35]. Their study's objective was to explore whether or not the self-reporting of pain can be increased and decreased by operant conditioning. Lousberg and colleagues attempted to demonstrate that the neurophysiologic basis of verbal pain report, defined by pain ERPs, is affected by the conditioning procedure. On the basis of the results of their study, they concluded that the subjective report of pain as well as a specific pain-related ERP component could be operantly conditioned. A similar study measuring processes in the central nervous system was performed by Flor, Diers and Birbaumer[30]. In their study, 16 chronic back pain (CBP) patients, 16 tension headache (TH) patients and 16 healthy controls (HC) were exposed to four series of ten electric stimuli while an electro-encephalogram (EEG) was recorded from three sites. The electric stimuli were specified at three levels, namely perception threshold, pain threshold and 10% below the pain threshold. The results showed that the CBP patients had significantly lower pain threshold and pain tolerance values than the HC and THA patients. THA patients displayed higher pain tolerance values and habituation was

less present in the CBP group. EEG results further showed that central reactivity was not significantly different between the groups nor were there significant group differences in the peripheral measures. However, since the stimulation intensity was significantly lower in the CBP patients, these data were indicative of both enhanced central and peripheral reactivity.

Another objective measurement technique recently applied is functional Magnetic Resonance Imaging (fMRI). Seminowicz, Mikulis and Davis used fMRI to show how a cognitively demanding task modulates pain-related brain activations and, conversely, how pain modulates attention-related activity[36]. Their findings suggested that cortical regions associated with pain could be modulated by cognitive strategies. In a study by Koyama et al., a combination of psychophysical and fMRI techniques were used to characterize brain activation related to the intensity of expected pain and experienced pain[37]. This study revealed that a mental representation of an impending sensory event can significantly shape the neural processes that underlie the formulation of the actual sensory experience. In short, this study illustrated how positive expectations diminish the severity of chronic disease states.

Evidently, the above mentioned studies demonstrate how physiological measures can be related to psychological mechanisms of pain. They thus support the contention that chronic pain has a multidimensional nature. Of course, additional studies are necessary in order to determine all physical-psychosocial mechanisms of chronic pain, but the results of the above mentioned studies may very well provide general guidelines for influencing physiological pain processes during treatment in order to improve psychosocial mechanisms of pain and vice versa.

1.3 Treatment of chronic pain

From a traditional medical approach, there is a wide variety of possibilities for the treatment of chronic pain. These possibilities vary from somatic oriented interventions (rest, medication, tissue stimulation, orthotics, pharmacological therapies, surgery, nerve block therapy, epidural steroid injection, trigger point injection, and physical therapy) to psychosocial oriented interventions (behavior therapy, family therapy, operant treatment)[7,8]. Although most of the forementioned interventions are somewhat effective, they are mostly unidimensional in nature. Applying a therapy that exclusively deals with either somatic or psychological

components of pain ignores the multidimensional aspects of chronic pain. A meta-analysis by Flor, Fydrich and Turk illustrates the superior effect of multidisciplinary pain treatment[38]. They showed that multidisciplinary pain treatment is superior to no treatment, waiting list control and more importantly, single-discipline, unidimensional forms of treatment. One of the most well-known multidimensional treatment approaches is cognitive-behavioural therapy.

1.3.1 Cognitive-behavioral therapy

Of all biobehavioral approaches to chronic pain, the cognitive-behavioral approach seems to be the most effective. Like the gate-control theory of Melzack and Wall[1], the cognitive-behavioral approach emphasizes the important contribution of psychological variables to pain experience and pain behavior[6,39]. The basis of this approach is that emotional reactions to and behaviors in certain situations are determined by the cognitions one has in such situations, which is, in this case, chronic pain.

Assumptions of the cognitive-behavioral approach are[6,39,40].

- An individual is an active processor of information and not a passive reactor;
- Thoughts (judgements, expectations, and convictions) are able to provoke and influence mood. They are able to influence physiological processes, have social consequences and can serve as an impetus for behavior. Inversely, mood, physiology, environmental factors and behavior can influence the nature and content of thinking processes;
- Behavior is mutually determined by individual as well as environmental factors;
- An individual is able to learn a more adjusted way of thinking, feeling, and behaving; and
- An individual must be an active and cooperating representative in changing his poorly adjusted thoughts, feelings, and behavior.

The aim of the cognitive-behavioral approach is to change both the patient's view of his pain and the poorly adjusted ways in which he copes with pain. Cognitive-behavioural therapy (CBT) endeavours to increase the patient's feeling of control over the pain experience and his life in general[41]. All cognitive-behavioral approaches to pain treatment emphasize being active, making time limitations and creating structure. The therapists are not merely providers of information, but also teachers, coaches and trainers. They work in cooperation with the patient (and

sometimes their relatives) to reach certain mutually determined goals. The cognitive-behavioral therapist is, like the behaviour therapist, concerned with using environmental manipulations. The difference between the two is that the cognitive-behavioral therapist sees manipulations as informational feedback trials that provide an opportunity for the patient to question and reappraise their situation, thereby promoting self-control over maladaptive thoughts, feelings, behaviors and physiological responses[39].

CBT is frequently used to help patients manage chronic pain. There are numerous studies that report the efficacy of CBT[42,43,44,45,46,47,48]. Unfortunately, very few studies show the efficacy of CBT in a randomized controlled trial (RCT). Nicholas, Wilson and Goyen compared CBT to non-psychological treatment such as standard physiotherapy and discussion sessions[45]. Their results indicated that, immediately after treatment and at 6-months follow-up, the CBT condition improved significantly more than the non-psychological treatment on measures of functional impairment, the employment of active coping strategies and self-efficacy beliefs. Another study, conducted by Basler, Jäkle and Kröner-Herwig, showed that experimental subjects receiving CBT and medical treatment reported less pain, better control over pain, more pleasurable activities and feelings, less avoidance and less catastrophizing than subjects receiving only medical treatment (control group)[46]. This study further showed that disability, in terms of social roles**,** physical functions and mental performance, in the experimental group was reduced. This was also the case 6 months after the initial experiment. Another example is a study by Becker et al., in which outpatient multidisciplinary CBT was compared with general medical practice[47]. Pain intensity was significantly reduced in patients who received CBT. These patients also showed a general improvement in health and quality of life. To further explore the effectiveness of CBT, Morley, Eccleston and Williams conducted a meta-analysis[49]. They showed that CBT produces significant changes in measures of pain experience, mood/affect, cognitive coping and appraisal, pain behavior and activity level, and social role functioning. They also concluded that active psychological treatments based on the principle of CBT are indeed effective. More recently, a study by Spinhoven et al. reported that patients following a CBT showed improvement with respect to depression, pain behavior, activity tolerance, catastrophizing and internal pain control both on a short term and on a long term basis[48].

Evidently, there is substantial evidence supporting both the cognitive-behavioral approach to chronic pain in general and the effectiveness of CBT specifically. Unfortunately, there is no general consensus on the specific content of a CBT program. This is exemplified by the fact that the label cognitive–behavioral approach has been applied to a wide range of techniques described in both clinical and research literature. Consequently, it is unclear exactly which specific components of CBT are responsible for the the effectiveness of CBT. In the following paragraph, a treatment program based on cognitive-behavioral principles, as executed by the Roessingh Center for Rehabilitation in the Netherlands, will be described.

1.3.2 The CBT program at the Roessingh Center for Rehabilitation

The CBT program at the Roessingh Center for Rehabilitation (CBT-R program) has been in place since 1984. The various treatment components are based on the cognitive-behavioral approach to chronic pain[6]. The main aim of the CBT-R program is to help patients learn how to deal with pain in an adequate way. Efforts are made to maintain the patients' highest possible quality of life despite the pain[50].

The CBT-R program is offered both as an inpatient and as an outpatient therapy. Intake procedures determine what mode of therapy (inpatient or outpatient) is most appropriate for the patient. To participate in the CBT-R programs, the following criteria are applied:

- The patient has to be referred to the program by a physician or specialist;
- The patient is not allowed to seek a diagnosis for his pain problem. In fact, he is required to decrease his medical dependency and has to agree with the program goal, which is to increase functioning despite the pain;
- The patient must have an obvious pain problem;
- The patient's pain must be related to the musculoskeletal system;
- The patient must be sufficiently instructable;
- The patient is not involved in a litigation concerning worker's compensation (as this may hinder the improvement of pain behavior); and
- An insurance company must compensate treatment costs.

The assumption of the CBT-R program is that pain is a sign of a disturbed balance between load and load capacity[50]. This means that there is a lack of balance between a) what a person wants to or has to do; b) his expectations and obligations; and c) his capabilities. The ways in which pain manifests and the duration of pain

complaints are thus seen as reactions to a disturbed balance. It is also assumed that learning mechanisms and early experiences with pain play a role in this process. As a result, the various treatments components in the CBT-R program aim to restore the balance between load and load capacity. This means that the patient has to initially focus on load relief and then, from a realistic baseline onward, gradually increase his load capacity.

In 1992, Winter evaluated the efficacy of the CBT-R program using a waiting list control period[50]. An improvement in both physical and psychological functioning was demonstrated. In accordance with the positive results of CBT mentioned in paragraph 1.3.1, Winter concluded that the CBT-R program is indeed effective. However, a post-hoc analyses of Winter's dataset shows that although the overall treatment results are positive (when looking at the group average), a great deal of variability in the therapy results does exist (unpublished data, Mes 1999). This basically means that although several patients did indeed significantly benefit from the program as evidenced by such things as a decrease of pain and a remarkable increase in physical activity, other patients reported no improvement. These findings call for a new investigation of the data. It is imperative to seek out the causes of variability and, in doing so, determine how treatment results can be further optimized.

1.4 Non-optimal treatment results and their hypothetical causes

The author contends that, in the field of chronic pain treatment (including CBT), there are very few treatment regimes that are optimal in nature. 'Optimal' treatment is here defined as treatment which leads to an improvement in *all* patients, at least to a certain extent. Failing this, a treatment can be considered non-optimal. Unfortunately, in almost every form of treatment, there are patients who show no improvement or even deteriorate following treatment. This means that most treatment regimes are non-optimal. Non-optimal treatment results are often characterized by a significant amount of variability around the mean treatment result. Three possible explanations for this variability are:

1. The existing treatment programs do not fit the individual problems, characteristics and needs of patients[51,52];
2. There are program failures or theory failures in the existing treatment programs[53,54];

3. The available knowledge about the origin and maintenance of chronic pain is insufficient[11,52,55].

In the following sections, these potential causes for variability will be discussed. For a more extensive discussion of variability in treatment results, readers are referred to the subsequent chapters.

1.4.1 Customizing treatment to patient characteristics

The mean results of a large sample of patients are usually the basis upon which treatment efficacy is determined. However, within a group with a positive mean treatment result, in addition to a number of very successfully treated patients, there are also patients who do not respond to treatment as expected or patients who even get worse following treatment. One of the hypothesized explanations for this treatment variability is heterogeneity in the group[51]. Heterogeneity in a patient group would imply that individual patients respond differently to standard treatment programs, thereby creating a large degree of variability in the overall treatment result.

One way to deal with this heterogeneity is to create subgroups of patients who have specific characteristics in common[51,56]. Once homogeneous subgroups are identified, treatment can be tailored to the specific underlying pain mechanisms, needs and characteristics of these subgroups. In addition, the selection criteria for treatment in a specific pain program can be adapted to subgroup characteristics, so that referring and admitting physicians can distinguish beforehand which patients are most likely to benefit from the program and which are not. In short, creating subgroups and using subgroup characteristics as criteria for treatment may lead to more optimal treatment results.

Subgroups can be created based on predictors for treatment success. The predictive value of variables such as age, pain history, number of operations and psychological distress has been examined to determine who benefits most from a given treatment[57,58]. Subgroups can also be identified using a statistical procedure called k-means cluster analysis[59]. With this type of analysis, subgroups of patients are created in such a way that within-group variability is minimized (patients within a cluster share as many characteristics as possible) and between-groups variability is maximized (patients of different clusters share as few characteristics as possible). These kinds of cluster analyses have been performed using pain related psychological characteristics, psychopathology and cognitive variables[60,61,62,63,64,65,66,67,68,69,70,71,72]. One of the most successful attempts to create

subgroups of patients has been done on the basis of the (West Haven-Yale) Multidimensional Pain Inventory (MPI)[73]. The MPI is used worldwide and has been translated into numerous languages. The clusters defined by the MPI have also been replicated in various different nations[56,74,75,76,77,78,79,80,81]. A number of studies have provided support for the hypothesis that customizing treatment and outcome measures based on a patient's characteristics, psychosocial needs and somatic needs may improve both treatment efficacy and the evaluation of treatment outcome.

For a more detailed description of the subject of pain treatment customization readers are referred to Chapter 3. In this chapter, the differential response of pain patients' subgroups, as defined with the MPI, on the outcome of the CBT-R program is discussed.

1.4.2 Failures underlying a treatment program

Every treatment program is guided by a theory, namely the program theory, which describes and prescribes what must be done in order to reach a desired outcome[53]. When a program is not executed or implemented according to the program theory, it's protocol and it's prescriptions, program failure is likely to result. Even when therapists do work strictly according to protocol, variability can still occur in therapy response. This is likely to occur when program developers and program executors use incorrect or insufficient theories in the design and implementation of a program or when incorrect theory is used to explain program outcomes. In these cases, theory failure has occurred.

In Figure 1.4, the process between input and output of a treatment program, as described by Chen, is represented[53]. A treatment program is developed with the intention of attaining certain goals. The treatment is thus the input which activates certain intervening mechanisms. These mechanisms are expected to lead to the desired outcome. The process, comprised of input, intervening mechanisms and outcome, is influenced by several factors in the implementation environment. Usually, research on treatment programs focuses predominantly on determining the outcome of the program. By doing so, the underlying mechanisms responsible for the occurrence of effects are often not considered. As a result, when outcome studies fail to demonstrate significant improvements, it is often unclear whether failure to show improvement is caused by program failures (such as inadequate treatment implementation or the inclusion/exclusion of certain treatment elements) or by theory failures (relating to the intervening mechanisms underlying the

treatment program). Conversely, when an outcome is positive, it is impossible to determine what the essential components of the treatment were or exactly how the treatment was conducted. Evaluation studies like this can thus be considered black box evaluations. An alternative is to do program evaluations that focus not only on output of a therapy but also on the implementation and underlying mechanisms of a program. By doing this, both the successful aspects and limiting factors of a program can be explained. This then enables program developers to adapt treatment programs in such a way that more optimal treatment effects can be obtained[53,82].

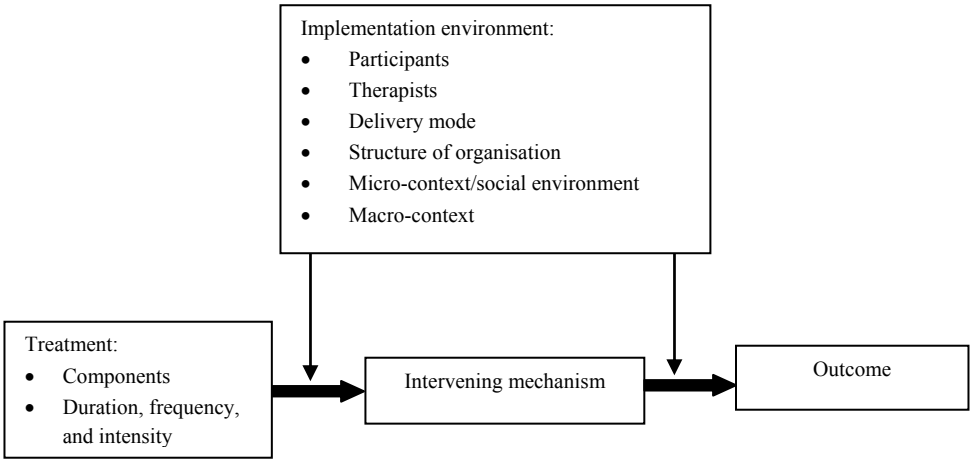


Figure 1.4: The process between input and output of a treatment program according to Chen[53]

A more extensive description of the advantages linked to a program evaluation can be found in Chapter 4. Chapter 4 also discusses the results of a process evaluation conducted on the CBT-R program.

1.4.3 Incomplete theoretical knowledge

In the existing scientific literature, several multidimensional, biobehavioral mechanisms have been proposed. These attempt to explain a) how the transition from acute to chronic pain occurs, and b) the ways in which chronic pain is maintained. Some of the today's most well recognized mechanisms have already been discussed in section 1.2. It has also been proposed that the insufficient or incorrect application of these mechanisms in chronic pain treatment (theory failures) may lead to a non-optimal treatment result. Another potential reason for

non-optimal treatment results is that the existing knowledge about the origin and the maintenance of chronic pain itself is insufficient or even incorrect. This may very well explain why existing treatment programs often times do not sufficiently fit the needs of patients.

The match-mismatch (MM) model of pain, as described in section 1.2.4, offers an experimental paradigm for the investigation of cognitive processing in chronic pain. The value of the model in explaining the ways in which chronic pain develops has already been illustrated in the past[26,29,83,84]. However, the majority of studies that focus on the M&M model have been conducted in laboratory situations. A relevant question is whether or not the model is also generalizable to the daily lives of chronic pain patients. There are several variables in the real life situation that may influence the behaviour of patients. For this reason, it is imperative that MM mechanisms are studied in the daily life of the patients who participated in the CBT-R program. This investigation seeks to explain the non-optimal results of the CBT-R program and the results of this investigation are expected to contribute significantly to the existing knowledge base on the underlying mechanisms of chronic pain. For a more detailed description of the MM model, as well as the studies conducted on this model, please see Chapters 5 through 7.

1.5 Aim of this thesis

Obviously, chronic pain is a major medical, social and economic problem. It has also become evident that the issue of non-optimal results is present in almost all chronic pain treatment programs, This means that, at this point in time, a substantial problem for many people lacks a sufficient solution. It is thus important to investigate the reasons why variability in treatment results occurs. By doing this, current treatment programs can hopefully be optimized.

In this dissertation, the three above mentioned explanations for the causes of variability in therapy results are extensively described. Additionally, the extent to which these underlying causes of variability are responsible for the treatment variability of the CBT-R program are explored. By doing so, an increase of the existing scientific knowledge on chronic pain is expected to be attained. With this increase in understanding and knowledge, programs such as the CBT-R program and other types of pain treatment programs can be adapted to produce more

optimal results. Ultimately, these adaptations may lead to better and more effective chronic pain treatment.

References

1. Melzack, R., & Wall, P.D. (1965). Pain mechanisms: a new theory. *Science, 150,* 971-979.
2. International Association for the Study of Pain (IASP) (1979). Pain terms: a list with definitions and notes on usage. *Pain, 6,* 249-252.
3. International Association for the Study of Pain (IASP) (1986). Classification of chronic pain. Description of pain syndromes and definitions of pain terms. *Pain Suppl*. S3.
4. Task Force on Taxonomy (1994). *Classification of chronic pain: Descriptions of chronic pain syndromes and definitions of pain terms.* Seattle: IASP Press.
5. Fordyce, W.E. (1974). Pain viewed as learned behavior. In J.J. Bonica (Ed.). *Advances in neurology*, vol. 4. New York: Raven Press, 415-422.
6. Turk, D.C., Meichenbaum, D., & Genest, M. (1983). *Pain and behavioral medicine, a cognitive-behavioral perspective*. New York: Guilford Press.
7. Flor, H., & Turk, D.C. (1984). Etiological theories and treatments for chronic back pain I: Somatic models and interventions. *Pain, 19,* 105-121.
8. Turk, D.C., & Flor, H. (1984). Etiological theories and treatments for chronic back pain. II. Psychological models and interventions. *Pain, 19,* 209-233.
9. Flor, H., Birbaumer, N., & Turk, D.C. (1990). The psychobiology of chronic pain. *Adv Behav Res Ther, 12,* 47-84.
10. Vlaeyen, J.W.S. (1991). *Assessment and treatment from a behavioral rehabilitation perspective*. Academic thesis. Lisse: Swets & Zeitlinger.
11. Feuerstein, M., & Beattie, P. (1995). Biobehavioral factors affecting pain and disability in low back pain: mechanisms and assessment. *Phys Ther, 75,* 267-280.
12. Turk, D.C. (1996). Biopsychosocial perspective on chronic pain. In R.J. Gatchel & D.C. Turk (Eds.). *Psychological approaches to pain management: a practitioner's handbook*. New York: Guilford Press.
13. Flor, H., & Birbaumer, N. (1994). Psychophysiological methods in the assessment and treatment of chronic musculoskeletal pain. In: J.G. Carlson, A.R. Seifert, & N. Birbaumer (Eds.). *Clinical applied psychophysiology.* New York: Plenum.
14. Flor H, Turk D, Birbaumer N (1985). Assessment of stress-related psychophysiological reactions in chronic back pain patients. *J Consult Clin Psychol, 27*, 195-202.
15. Lethem, J., Slade, P.D., Troup, J.D.G., & Bentley, G. (1983). Outline of a fear-avoidance model of exaggerated pain perception-I. *Behav Res Ther, 21,* 401-408.
16. Vlaeyen, J.W.S., Kole-Snijders, A.M.J., Boeren, R.G.B., & Eek, H. van (1995). Fear of movement/(re) injury in chronic low back pain and its relation to behavioral performance. *Pain, 62,* 363-372.
17. Philips, H.C. (1987). Avoidance behaviour and its role in sustaining chronic pain. *Behav Res Ther, 25,*: 273-279.
18. Vlaeyen, J.W.S., Seelen, H.A.M., Peters, M., Jong, P. de, Aretz, E., Beisiegel, E., & Weber, W.E.J. (1999). Fear of movement/(re)injury and muscular reactivity in chronic low back pain patients: an experimental investigation. *Pain, 82*, 297-304.
19. Vlaeyen, J.W.S., & Linton, S.J. (2000). Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. *Pain, 85*, 317-332.

20. Hasenbring, M. (1999). Wenn die Seele of die Bandscheiben drückt. *Rubin*, 1/99.
21. Hasenbring, M., Hallner, D., & Klasen, B. (2001). Psychologische Mechanismen im Prozess der Schmerzchronifizierung: Unter- oder überbewertet? *Schmerz, 15*, 442-447.
22. Hasenbring, M.I., Plaas, H., Fischbein, B., & Willburger, R. (2006). The relationship between activity and pain in patients 6 months after lumbar disc surgery: do pain-related coping modes act as moderator variables. *Eur J Pain*, in press.
23. Rachman, S., & Lopatka, C. (1986). Match and mismatch in the prediction of fear-I. *Behav Res Ther, 24*, 387-893.
24. Rachman, S., & Lopatka, C. (1986). Match and mismatch of fear in Gray's theory-II. *Behav Res Ther, 24,* 395-401.
25. Arntz, A., & Hout, A. van den (1988). Generalizability of the match/mismatch model of fear. *Behav Res Ther, 26.* 207-223.
26. Arntz, A., Hout, M.A. van den, Lousberg, R., & Schouten, S. (1990). Case history and shorter communication: Is the match/mismatch model based on a statistical artefact? *Behav Res Ther*, *28*, 249-253
27. Arntz, A., & Lousberg, R. (1990). The effects of underestimated pain and their relationship to habituation. *Behav Res Ther, 28,* 15-28.
28. Arntz, A., Eck, M. van, & Jong, P. de (1990). Avoidance of pain of unpredictable intensity. *Behav Res Ther, 29*, 197-201.
29. Arntz, A., Hout, A, van den, Berg, G. van den, & Meijboom, A. (1991). The effects of incorrect pain expectations on acquired fear and pain responses. *Behav Res Ther, 29,* 547-560.
30. Flor, H., Diers, M., & Birbaumer, N. (2004). Peripheral and electrocortical responses to painful and non-painful stimulation in chronic pain patients, tension headache patients and healthy controls. *Neurosci Lett, 361*, 147-150.
31. Schneider, C., Palomba, D., & Flor, H. (2004). Pavlovian conditioning of muscular responses in chronic pain patients: central and peripheral correlates. *Pain, 112*, 239-247.
32. Läubli, T., Hermens, H.J., & Sjøgaard, G. (2006). Neuromuscular assessment of the elderly worker, NEW: a multidisciplinary European research project. *Eur J Appl Physiol, 96*, 107-109.
33. Schulte, E., Kallenberg, L.A.C., Christensen, H., Disselhorst-Klug, C., Hermens, H.J., Rau, G., & Søgaard, K. (2006). Comparison of the electromyographic activity in the upper trapezius and biceps brachii muscle in subjects with muscular disorders: a pilot study. *Eur J Appl Physiol, 96*, 185-193.
34. Sjøgaard, G., Søgaard, K., Hermens, H.J., Sandsjö, L., Läubli, T., Thorn, S., Vollenbroek-Hutten, M.M.R., Sell, L., Christensen, H., Klipstein, A., Kadefors, R., & Merletti, R. (2006). Neuromuscular assessment in elderly workers with and without work related shoulder/neck trouble: the NEW-study design and physiological findings. *Eur J Appl Physiol, 96*, 110-121.
35. Lousberg, R., Vuurman, E., Lamers, T., Breukelen, G. Van, Jongen, E., Rijnen, H., Maessen, C., & Hermens, H. (2005). Pain report and pain-related evoked potentials operantly conditioned. *Clin J Pain, 21*, 262-271.
36. Seminowicz, D.A., Mikulis, D.J., & Davis, K.D. (2004). Cognitive modulation of pain-related brain responses depends on behavioral strategy. *Pain, 112*, 48-58.
37. Koyama, T., Mchaffie, J.G., Laurienti, P.J., & Coghill, R.C. (2005). The subjective experience of pain: where expectations become reality. *Proc Natl Acad Sci U S A, 102*, 12850-12955.

38. Flor, H., Fydrich, T., & Turk, D.C. (1992). Efficacy of multidisciplinary pain treatment centers: a meta-analytic review. *Pain, 49,* 221-230.
39. Turk, D.C., & Meichenbaum, D. (1994). A cognitive-behavioral approach to pain management. In P.D. Wall & R. Melzack (Eds.). *Textbook of pain*. Edinburgh: Churchill Livingstone.
40. Bradley, L.A. (1996). Cognitive-behavioral therapy for chronic pain. In R.J. Gatchel & D.C. Turk (Eds.). *Psychological approaches to pain management: a practitioner's handbook*. New York: Guilford Press.
41. Kerns, R.D., & Casey, J.M. (1992). Assessment of the psychological context of the experience of chronic pain. In D.C. Turk & R. Melzack (Eds.). *Handbook of pain assessment*. New York: Guilford Press.
42. Turner, J.A. (1982). Comparison of group progressive-relaxation training and cognitive-behavioural group therapy for chronic low back pain. *J Consult Clin Psychol, 50,* 757-765.
43. Turner, J.A., & Clancy, S. (1988). Comparison of operant behavioural and cognitive-behavioural group treatment for chronic low back pain. *J Consult Clin Psychol, 56,* 261-266.
44. Newton-John, T.R., Spence, S.H., & Schotte, D. (1995). Cognitive-behavioural therapy versus EMG Biofeedback in the treatment of chronic low back pain. *Behav Res Ther, 33,* 691-697
45. Nicholas, M.K., Wilson, P.H., & Goyen, J. (1992). Comparison of cognitive-behavioural group treatment and an alternative non-psychological treatment for chronic low back pain. *Pain, 48,* 339-347.
46. Basler, H.D., Jäkle, C., & Kröner-Herwig, B. (1997). Incorporation of cognitive-behavioral treatment into the medical care of chronic low back patients: a controlled randomized study in German pain treatment centers. *Patient Educ Couns, 31,* 113-124.
47. Becker, N., Sjøgren, P., Bech, P., Kornelius Olsen, A., & Eriksen, J. (2000). Treatment outcome of chronic non-malignant pain patient managed in a Danish multidisciplinary pain centre compared to general practice: a randomised controlled trial. *Pain, 84,* 203-211.
48. Spinhoven, P., Kuile, M. ter, Kole-Snijders, A.M., Hutten Mansfeld, M., Den Ouden, D.J., & Vlaeyen, J.W. (2004). Catastrophizing and internal pain control as mediators of outcome in the multidisciplinary treatment of chronic low back pain. *Eur J Pain, 8*, 211-219.
49. Morley, S., Eccleston, C., & Williams, A. (1999). Systematic review and meta-analysis of randomized controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache. *Pain, 80,* 1-13.
50. Winter, F.A.M. (1992). *Pijnrevalidatie voor chronische pijnpatiënten: de pijn de baas*. Enschede: Intermap B.V.
51. Turk, D.C. (1990). Customizing treatment for chronic pain patients: who, what, and why? *Clin J Pain, 6,* 255-270.
52. Lousberg, R. (1994). *Chronic pain: Multiaxial diagnostics and behavioral mechanisms*. Academic thesis. Maastricht: University Press.
53. Chen, H.T. (1990). *Theory-driven evaluations*. London: Sage Publications.
54. Mes, C.A.J., Winter, F.A.M., & Wijck, R. van (1999). Programma-evaluatie in de pijnrevalidatie: analyse van het Pijn de Baas programma. *Gedrag Gezond, 27,* 136-147.
55. Lousberg, R. (1997). Chronische pijn als multidimensioneel fenomeen, en hoe verder? *NTPP, 17,* 40-44.

56. Turk, D.C., Okifuji, A., Sinclair, J.D., & Starz, T.W. (1998). Differential responses by psychosocial subgroups of Fibromyalgia syndrome patients to an interdisciplinary treatment. *Arthritis Care Res, 11,* 397-404.
57. Maruta, T., Swanson, D.W., & Swenson, W.M. (1979). Chronic pain: which patients may a pain-management program help? *Pain, 7,* 321-329
58. Lee, P.W.H., Chow, S.P., Lieh-Mak, F., Chan, K.C., & Wong, S. (1989). Psychosocial factors influencing outcome in patients with low-back pain. *Spine, 14,* 838-843.
59. Milligan, G.W., & Cooper, M.C. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika, 50*, 159-179.
60. Sternbach, R. (1974). *Pain patients: traits and treatments*. New York: Academic Press.
61. Derogatis, L.R. (1977). *SCL-90 administration, scoring and procedures manual*. Baltimore, MD: Johns Hopkins University Press.
62. Bradley, L.A., Prokop, C.K., Margolis, R., & Gentry, W.D. (1978). Multivariate analyses of the MMPI profiles of low back pain patients. *J Behav Med*, *1,* 253-272.
63. Hollon, S.D., & Kendall, P.C. (1980). Cognitive self-statements in depression: development of an Automatic Thoughts Questionnaire. *Cog Ther Res, 4*, 383-395.
64. Armentrout, D.P., Moore, J.E., Parker, J.C., Hewett, J.E., & Felz, C. (1982). Pain patient MMPI subgroups: the psychological dimensions of pain. *J Behav Med, 1,* 201-211.
65. Schwartz, D.P., & DeGood, D.E. (1983). An approach to the psychosocial assessment of the chronic pain patient. *Curr Concepts Pain, 1*, 3-11.
66. Moore, J.E., Armentrout, D.P., Parker, J.C., & Kivlahan, D.R. (1986). Empirically derived pain-patients MMPI subgroups: prediction of treatment outcome. *J Behav Med, 9,* 51-63.
67. Buckelew, S.P., DeGood, D.E., Schwartz, D.P., & Kerler, R.M. (1986). Cognitive and somatic item response pattern of pain patients, psychiatric patients, and hospital employees. *J Clin Psychol, 42,* 852-860.
68. Butterworth, J.C., & Deardorff, W.W. (1987). Psychometric profiles of craniomandibular pain patients: identifying specific subgroups. *J Craniomandib Pract, 5,* 225-232.
69. 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. *J Behav Med, 111,* 159-169.
70. Everitt, B. (1989). *Cluster analysis* (2nd ed.). New York: Halstead-Wiley.
71. Jamison, R.N., Rock, D.L., & Parris, W.C.V. (1988). Empirically derived Symptom Checklist 90 subgroups of chronic pain patients: a cluster analysis. *J Behav Med, 11,* 147-158.
72. Ingram, R.E., Atkinson, J.H., Slater, M.A., Saccuzzo, D.P., & Garfin, S.R. (1990). Negative and positive cognition in depressed and non-depressed chronic pain patients. *Health Psychol, 9,* 300-314.
73. Kerns, R.D., Turk, D.C., & Rudy, T.E. (1985). The West Haven-Yale Multidimensional Pain Inventory (WHYMPI). *Pain, 23,* 345-356.
74. Rudy, T.E., Turk, D.C., Zaki, H.S., & Curtin, H.D. (1989). An empirical taxometric alternative to traditional classification of temporomandibular disorders. *Pain, 36*, 311-320.
75. Rudy, T.E., Turk, D.C., Kubinski, J.A., & Zaki, H.S. (1995). Differential treatment responses of TMD patients as a function of psychological characteristics. *Pain, 61,* 103-112.
76. Turk, D.C., Okifuji, A., Sinclair, J.D., & Starz, T.W. (1996). Pain, disability, and physical functioning in subgroups of patients with Fibromyalgia. *J Rheumatol, 23,* 1255-1262.

77. Dahlstrom, L., Widmark, G., & Carlsson, S.G. (1997). Cognitive-behavioral profiles among different categories of orofacial pain patients: diagnostic and treatment implications. *Eur J Oral Sci, 105*, 377-383.
78. Hutten, M.M., Hermens, H.J., IJzerman, M.J., Lousberg, R., & Zilvold, G. (1999). Distribution of psychological aspects in subgroups of chronic low back pain patients divided on the score of physical performance. *In J Rehabil Res, 22*, 261-268.
79. Bergstrom, G., Jensen, I.B., Bodin, L., Linton, S.J., & Nygren, A.L. (2001). The impact of psychologically different patient groups on outcome after a vocational rehabilitation program for long-term spinal pain patients. *Pain, 93*, 229-237.
80. Hankin, H.A., Spencer, T., Kegerreis, S., Worrell, T., & Rice, J.M. (2001). Analysis of pain behavior profiles and functional disability in outpatient physical therapy clinics. *J Orthop Sports Phys Ther, 31,* 90-95.
81. Vollenbroek-Hutten, M.M., Hermens, H.J., Wever, D., Gorter, M., Rinket, J., & IJzerman, M.J. (2004). Differences in outcome of a multidisciplinary treatment between subgroups of chronic low back pain patient defined using two multiaxial assessment instruments: the multidimensional pain inventory and lumbar dynamometry. *Clin Rehabil, 18*, 566-579.
82. Sijben, N. (1984). Programma-evaluatie…Wat is dat? [Program evaluation…what is it?] In G.M. Schippers, N. Sijben, C. Schaap & J. Mastboom (red.). Conference reports *Programma-evaluatie in de geestelijke gezondheidszorg*. Lisse: Swets & Zeitlinger, 11-27.
83. Arntz, A. (1996). Why do people tend to overpredict pain? On the asymmetries between underpredictions and overpredictions of pain. *Behav Res Ther*, 34, 545-554.
84. Arntz, A., & Hopmans, M. (1998). Underpredicted pain disrupts more than correctly predicted pain, but does not hurt more. *Behav Res Ther*, 36, 1121-1129.

CHAPTER 2

Treatment outcomes in pain rehabilitation: a randomized clinical trial

2.1 Introduction

Programs based on a cognitive-behavioral approach show significant potential when it comes to treating chronic pain. Several randomized controlled trials (RCTs) have been conducted and have incontrovertibly established the effectiveness of multidisciplinary pain management programs. These randomized studies conclude that cognitive-behavioral therapy (CBT) contributes to a decrease in pain, an improvement in physical and psychological functioning, the employment of active coping strategies, an increase of self-efficacy beliefs, more control over pain, and better social role-functioning[1,2,3,4].

At several locations in the Netherlands, cognitive-behavioral programs for the treatment of chronic pain are offered[5]. Although most use a multidisciplinary approach, all programs demonstrate differences in both nature and content. Few pain centers have the capacity to provide inpatient treatment programs. As a result, most centers treat pain patients on an outpatient basis. Since 1984, the Roessingh Center for Rehabilitation (RCR) offers a chronic pain management program based on the cognitive-behavioral approach to chronic pain (CBT-R program)[6]. The aim of the CBT-R program is to improve physical and psychosocial functioning without specifically focusing on decreasing pain[7].

As mentioned briefly in the first chapter, Winter has evaluated the efficacy of the semi-inpatient CBT-R program at the RCR[7]. In this waiting list controlled study, 185 patients at RCR were included between 1990 and 1992. Psychological complaints (as measured with the Symptom Checklist-90 (SCL-90))[8], and pain experience and pain behavior (as measured with the Multidimensional Pain Inventory – Dutch Language Version (MPI-DLV))[9] were measured on four occasions, namely at intake, at the start of the program, after treatment and at one year follow-up. Winter's findings were positive and thus it was concluded that the CBT-R program was successful. Most outcome measures improved and remained improved at follow-up. However, despite positive findings, some critical notes with regard to Winter's study are in order. Firstly, Winter did not use randomization. Consequently, the results of the study should be interpreted with caution. Non-randomized studies can overestimate treatment results, thereby making it difficult to attribute the study results to the treatment itself[10]. A second point is that, although the average results of Winter's study show an overall improvement in functioning of the group between pre and post test, a closer look at the outcome

data shows a relatively large amount of variability in therapy results (Table 2.1). The *average* results do indeed demonstrate a positive effect of the CBT-R program but this does not necessarily mean that all patients showed an improvement. As shown by the mean differences and confidence intervals (CI) in Table 2.1, the extent to which patients improve after treatment differs remarkably. On five MPI-DLV scales, the CIs have negative values at the lower boundary. Unpublished analyses (Mes, 1999) indicate that some patients not only failed to improve following treatment, but also worsened on nearly all MPI-DLV scales after treatment.

Treatment variability is not a phenomenon specific to the CBT-R program. It can also be found in other studies that appraise the effectiveness of chronic pain treatment. For example, in a study by Jensen et al.[11], in which change scores between pre and post multidisciplinary treatment were calculated on measures of patient-reported physical functioning and pain behavior, the mean (M) and standard deviations (SD) of these change scores for physical functioning were M = 7.71 and SD = 9.92. For pain behavior the scores were M = 1.02 and SD = 1.00. The 95% CI of these change scores are respectively 0.814 - 1.206 and 5.664 - 9.756, which indicates a large range of improvement on these variables. In a study by Morley et al., the effect sizes with 95% CI for treatment versus waiting list controls were presented for CBT, behavior therapy (BT) and biofeedback (BFB) on several domains[3]. The CIs for BT on pain experience and mood/affect depression are -0.09 - 0.55 and -0.21 - 0.15, respectively. The CI for CBT for behavior expression was -0.08 - 1.05 and the CI of BFB for behavior activity was -0.03 - 0.80. Again, these data indicate that although there is a group of patients that improve on the measured scales (positive change scores), there are also patients that deteriorate in functioning after treatment (negative change scores).

Table 2.1: Results on the MPI-DLV and SCL-90 by Winter (1992)

Instrument		N	M (SD) pre	M (SD) post	M (SD) $\Delta_{pre-post}$	CI 95%	t value	*P* value
MPI-DLV	Pain severity	156	4.38 (0.87)	3.82 (1.08)	0.55 (1.01)	0.391 - 0.711	6.85	.000
	Interference	154	4.13 (0.91)	3.75 (1.70)	0.39 (0.95)	0.201 - 0.537	5.06	.000
	Life control	156	3.67 (1.35)	4.09 (1.29)	-0.42 (1.55)	-0.670 - -0.174	-3.40	.001
	Affective distress	154	3.25 (1.22)	2.46 (1.22)	0.79 (1.35)	0.575 - 1.007	7.30	.000
	Support	154	3.91 (1.82)	3.66 (1.70)	0.25 (1.35)	0.031 - 0.467	2.29	.023
	Punishing responses	133	1.34 (1.33)	1.31 (1.27)	0.04 (1.19)	-0.170 – 0.242	.35	.726
	Solicitous responses	133	2.87 (1.20)	2.63 (1.17)	0.24 (1.15)	0.038 – 0.438	2.38	.019
	Distracting responses	133	2.14 (1.31)	2.27 (1.23)	-0.12 (0.94)	-0.289 - 0.039	-1.53	.129
	Household chores	154	3.92 (1.53)	3.95 (1.37)	-0.03 (1.50)	-0.217 - 0.156	-.33	.742
	Outdoor work	154	1.30 (1.17)	1.41 (1.28)	-0.11 (1.18)	-0.300 - 0.080	-1.15	.251
	Social activities	154	2.49 (1.07)	2.64 (1.06)	-0.15 (0.94)	-0.303 - 0.001	2.00	.047
	General activities	154	1.81 (0.98)	2.04 (1.04)	-0.23 (0.80)	-0.355 - -0.099	-3.53	.001
SCL-90	Fear	152	21.14 (8.12)	17.80 (7.34)	3.34 (6.58)	2.276 - 4.408	6.27	.000
	Agoraphobia	152	11.23 (5.34)	9.78 (4.77)	1.45 (4.45)	0.732 - 2.181	4.02	.000
	Depressive thoughts	152	35.78 (11.64)	29.61 (10.97)	6.16 (11.64)	4.277 - 8.053	6.53	.000
	Somatic complaints	152	32.43 (8.33)	28.66 (8.46)	3.78 (7.69)	2.530 - 5.022	6.06	.000
	Insufficiency	152	22.01 (6.02)	19.96 (6.23)	2.05 (5.91)	1.095 - 3.011	4.28	.000
	Sensitivity	152	32.97 (10.27)	29.47 (10.74)	3.51 (9.86)	1.907 - 5.107	4.38	.000
	Hostility	152	9.70 (3.70)	8.36 (3.03)	1.34 (2.90)	0.872 - 1.812	5.70	.000
	Sleeping problems	152	9.97 (3.88)	8.61 (3.48)	1.36 (3.75)	0.682 - 1.970	4.47	.000
	Psycho-neuroticism	152	190.01 (48.23)	165.59 (49.64)	24.43 (42.89)	17.470 - 31.386	7.02	.000

2.2 Methods

2.2.1 Subjects

From April 1999 to July 2001, all chronic pain patients referred to the CBT-R program were approached and asked to participate in this study. Criteria for exclusion were: age under 18 years, pain duration of less than six months, pain complaints as a result of a whiplash injury, presence of serious psychopathology, illiteracy or insufficient knowledge of the Dutch language. Patients not included in the study received the usual intake and treatment procedure.

Patients who met the inclusion criteria were asked to participate in the study. After giving written informed consent, they were randomly assigned to one of two groups: an Intervention Group (IG) and a waiting list Control Group (CG). The IG received immediate intake and treatment as soon as possible after informed consent (approximately one month). The CG was assigned to a waiting period of approximately six months followed by the same intake and treatment procedures as the IG. Randomization was performed in blocks of two. The study was designed and executed with permission of the Roessingh Medical Ethical Committee.

Based on the results of Winter‘s study, a power calculation was carried out on the MPI-DLV Interference scale. The same effect size obtained by Winter (δ=.0533) was used to calculate the minimum required sample size. Given an α of .05 and a power of 80% (β = .02), a minimum sample size of n= 60 patients per group was determined[12].

2.2.2 Treatment

The CBT-R program is offered in several variants[13,14]. During the intake procedure, the intake team and the patient decide together which variant is most appropriate for the patient. The following variants are offered:

- *Outpatient group treatment*: This variant is for patients capable of liberating enough time and energy for the treatment program. Participants must possess a physical and mental load capacity that enables them to follow a rather intensive therapy two days per week. The duration of the outpatient program is eight weeks. The number of patients in the out-patient treatment variant is limited to eight.

- *Inpatient group treatment:* This treatment mode is for patients with an extremely low physical load capacity. Many are partly or wholly wheelchair dependent and thus require significant rest during the day. The patients in the inpatient program stay in the RCR for five days a week. The duration of the program is eight weeks. This is spread out over 16 weeks. This means that patients receive one week of therapy followed by a week at home, which is then followed by a week of therapy and so on. Each patient group contains a maximum of seven participants. It is also important to note that the possibility of following the inpatient CBT-R program on an individual basis is available.
- *Semi-inpatient group treatment*: The RCR is the only rehabilitation center in the Netherlands that offers a third treatment mode besides an inpatient and an outpatient mode, namely the semi-inpatient mode. This variant is suited for patients who are rather mobile, are able to manage the movement programs and are mentally capable enough to function in a group. This variant is appropriate for patients who, despite being mobile, are incapable of following the outpatient treatment mode. The semi-inpatient variant, like the outpatient mode, is an eight week program. Every week, the patients attend 2½ days of therapy (and stay overnight) in the RCR pain clinic. The rest of the week is spent at home. This variant provides patients with the opportunity to practice the newly acquired skills at home and give weekly feedback about their experiences. The number of patients in the semi-patient treatment variant is limited to seven.

For the RCT, patients from the outpatient and semi-inpatient treatment group were included in the IG and CG. Patients from the inpatient treatment group were not included because of a different duration of treatment, making comparison with the other two groups impossible.

2.2.3 Procedure

On several occasions before as well as after treatment, questionnaires were sent to the patients at home (t0 – t6, Figure 2.1). The questionnaires used are described in the next paragraph.

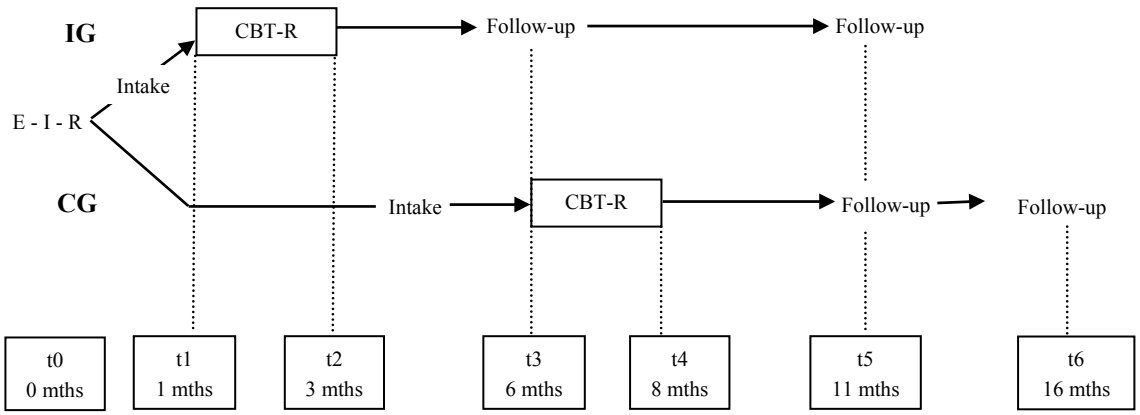


Figure 2.1: Research design (IG = intervention group; CG = control group; E=entrance, I= informed consent, R = randomization, t=time of measurement)

At t3, a maximum contrast was expected between the IG and the CG. The reason for this assumption is that the CBT-R program primarily aims to achieve behavioral change and the expectation is that behavioral change is greater three months after treatment (when compared to immediately after treatment). This study focuses only on the controlled results between t0 and t3.

2.2.4 Outcome measures

In a rehabilitation setting, a patient's functioning is examined from a multidimensional perspective. Additionally, in conjunction with the rehabilitation goal of the CBT-R program, patient disability is expected to decrease and functioning is expected to increase following treatment. As a result, the outcome measures were selected to reflect the level of functioning, activity and role-fulfillment, as rehabilitation goals on this level are usually considered to be the most important[15]. As the CBT-R program is intended at improving functioning despite the pain, it is not the pain itself but rather the *interference* of pain during daily activities that is expected to decrease. For this reason, the MPI-DLV scale Interference was taken as the primary effect parameter of this study. In accordance with Winter's study (and in an effort to replicate the results found by Winter[7]), the following two questionnaires were used to determine the multidimensional functioning of pain patients both before and after treatment:

- *Multidimensional Pain Inventory-Dutch Language Version (MPI-DLV)*[9,16]. The MPI-DLV is a multidimensional self-report questionnaire that is used to measure psychosocial and behavioral aspects of pain. The questionnaire consists of nine scales divided in three parts. The first part is concerned with psychosocial aspects of pain such as pain experience, the influence of pain on different aspects of life, pain control, distress and social support. The second part contains a number of questions related to the patient's perception of the reactions given by the most important significant other (in most cases, the patient's partner) to the patient's pain complaints. The third part measures the patient's activity level. All nine scales of the MPI-DLV have a good internal consistency (Cronbach's alpha is between 0.74 and 0.94) and a satisfying test-retest reliability. In addition, the MPI-DLV is found to be sensitive to treatment changes[16,17].
- *Symptom Checklist-90 (SCL-90)*[8]. The SCL-90 is a multidimensional list of complaints that measures the presence of psychological problems on eight dimensions. The dimensions are: anxiety, agoraphobia, depressive thoughts, somatic complaints, sensitivity, insufficiency, sleeping problems and hostility. The total score gives an indication of a patient's general psychological wellbeing.

Both questionnaires were assessed at all points of measurement. If a questionnaire was not returned within ten days, patients were reminded by mail or telephone.

2.2.5 Statistical analysis

All data were processed using SPSS version 11.5[18]. The primary analysis of the data was focused on determining the outcome of the CBT-R program on MPI-DLV Interference. Secondary, exploratory analyses were performed on the other variables of the MPI-DLV and SCL-90. To compare the data of the IG with the CG, analysis of variance for repeated measures (ANCOVA) was carried out with: a) time as the within-subjects factor (three levels being t1, t2 and t3); and b) group (IG and CG) as a between-subjects factor (two levels). Based on recommendations by Stevens, as well as Tabachnik and Fidell, t0 was added as a covariate[19,20].

Multivariate analyses (MANOVA) were applied using the same factors as described above but then on more than one dependent MPI-DLV variable. The same techniques were used to analyze differences in treatment results between the treatment variants, whereby treatment variant (outpatient versus semi-inpatient)

was used as a between-subjects factor (two levels). Along the lines described by Hair et al., the assumptions of the test were evaluated (multivariate normality of the dependent variables by graphical inspection of pairs of variables and sphericity by the Greenhouse-Geisser adjustment)[21]. Further, the results were checked for the presence and effect of possible outlying cases.

A priori, a between-group as well as a time x group effect was hypothesized for all MPI-DLV and SCL-90 scales in favor of the IG. Based on the results of Winter, the expectation was that, within the IG, the largest effect would be gained between t1 and t2 (the pure treatment period) with an additional, smaller effect between t2 and t3[7]. Throughout the study, an alpha of .05 was applied. All reported *P* values are two-sided.

2.2.6 Missing data

To enlarge the number of analyzable cases in the CG for (M)ANOVA, missing values on a given measurement point were imputed as far as possible by linear approximation (i.e. with the average of the two surrounding valid values), assuming that no differences were to be expected between the times of measurement. In the case of dropout, the Last Observation Carried Forward (LOCF) method was applied[22,23].

For the IG, missing values on t1 were imputed by t0 values assuming that no differences were to be expected between t0 and t1. Further, missing data on t2 were, in so far as it is possible, supplemented by means of linear approximation (mean of t1 and t3). In the case of dropout, the LOCF method was applied.

2.3 Results

2.3.1 Population

Out of the 391 patients referred to the CBT-R program, eight patients did not fulfill the inclusion criteria. Additionally, 86 patients choose not to participate in this study and thus gave no informed consent. No significant differences in age (*P* = .753), sex (*P* = .169) or diagnosis (*P* = .120) were found between patients who did or did not provide informed consent. Consequently, it may be assumed that no selection bias occurred. The remaining 297 patients who did give informed consent were included and randomly assigned to either IG or CG. In Appendix 2A, a flow diagram is presented that reflects the number of patients participating at the various times of measurement. Also, explanations for drop out are mentioned in this

diagram. The flow diagram indicates that the total number of available outpatient (IG 16; CG 19) and semi-inpatient patients (IG 66; CG 45) at baseline was 146. In Table 2.2, the general characteristics of the research population are presented.

The drop-out rate in the CG was slightly larger than the drop-out rate in the IG. This could have been caused by the extended waiting period required for CG participants. However, the difference in drop-out rate between both groups was not significant ($P = .645$). Due to logistic problems, the period between t0 and t1 could not be held exactly stable for both groups. Analyses showed that this period was significantly different between the groups as the period of the CG was a few days shorter than that of the IG ($t_{(144)} = 2.784$; $P = .006$). This difference was controlled for by adding time between t0 and t1 as a covariate in subsequent analyses. Further, the IG and CG were comparable with regard to age, sex, education, diagnosis and pain duration. Additionally, the MPI-DLV and SCL-90 scales showed no significant differences between the two groups. It was thus concluded that the randomization procedure was successful.

Table 2.2: Population characteristics (N = 146)

Variable		**Score**
Age (years)	M(SD), range	43.1 (10.5), 18 - 71
Sex	Male	31 (21.1%)
	Female	115 (78.8%)
Education	Primary school	26 (17.8%)
	Secondary school	95 (65.0%)
	Higher education	17 (11.6%)
Civil status	Married/living together	124 (84.9%)
	Single	15 (10.3%)
Pain duration (years)	M (SD), range	8.4 (8.1), 0.5 - 40
Pain location	Head/face	4 (2.7%)
	Neck/shoulder	4 (2.7%)
	Back	25 (17.1%)
	Hip/pelvis	2 (1.4%)
	Leg/foot	7 (4.8%)
	Arm/hand	3 (2.1%)
	Other	2 (1.4%)
	More than one place	99 (67.8%)
Randomization group	IG	82 (56.2%)
	CG	64 (43.8%)
Drop-out between t0 and t3	IG no drop-out	69 (84.1%)
	IG drop-out	13 (15.9%)
	CG no drop-out	52 (81.3%)
	CG drop-out	12 (18.8%)

2.3.2 Treatment results: IG versus CG

As mentioned in section 2.2.4, the effect of interference of pain with daily life was the main parameter used to decide whether the CBT-R program was effective or not. It was expected that interference would decrease in the IG as a result of the CBT-R and that it would remain the same (or even increase) in the CG. In Figure 2.2, the results of the MPI-DLV Interference scale are presented. This figure shows that the CG remained roughly the same as the IG between t0 and t2, but decreased between t2 and t3. The IG showed a decrease in interference of pain during treatment between t1 and t2. This effect was held to some extent until t3. Looking at Figure 2.2, it becomes evident that the decrease from t2 to t3 in the CG was comparable to the decrease the IG showed between t0 and t1. Anticipation of treatment may have been responsible for the decreases in both groups.

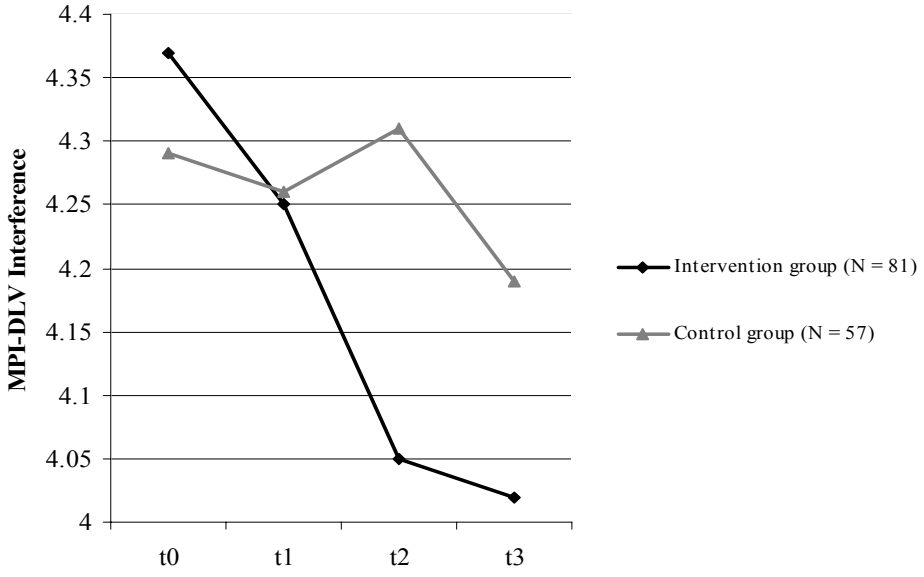


Figure 2.2: Mean scores on the MPI-DLV Interference scale

The results of the ANCOVA over t1, t2 and t3 (with covariates t0 and time between t0 and t1) demonstrated a significant between-group effect ($F_{(1,\ 134)}$ = 5.061; P = .026). In other words, independent of time, there was a difference in interference between the IG and CG. The IG showed lower values than the CG. In

addition to this between-group effect, a non-significant multivariate time x group interaction was observed (P = .112), which could be decomposed into a non-significant linear component ($F_{(1, 134)}$ = 1.193; P = .277) and an almost significant quadratic component ($F_{(1, 134)}$ = 3.842; P = .052). Whereas the CG demonstrated an increase of interference in the time interval t1-t2 (probably representing deterioration due to the six month waiting period prior to intake), the IG showed a large decrease during this interval. This could be interpreted as an initial large treatment effect. Outliers did not influence these effects.

From several studies, it appeared that age, sex and pain duration are related to therapy result[24,25,26]. To correct for a confounding effect of these variables, these variables were simultaneously added in the analysis as between factor (sex) or covariate (age and pain duration). A similar result was obtained, namely a significant between-group effect ($F_{(1, 130)}$ = 4.927; P = .028) and a tendency towards a multivariate time x group interaction ($F_{(1, 130)}$ = 2.631; P = .076), which could be decomposed into a non-significant linear component ($F_{(1, 130)}$ = 1.531; P = .218) and a significant quadratic component ($F_{(1, 130)}$ = 3.982; P = .048). Again, outliers did not influence these effects.

Because the anticipation effects, demonstrated by the IG and CG between t0 and t1 and between t2 and t3 respectively, were quite similar, one may conclude that there has been a genuine ('pure') treatment effect between t1 and t2. Because of this, the MPI-DLV Interference scores at t1 and t2 were compared. The between-group effect was also found to be significant using this procedure ($F_{(1, 135)}$ = 5.017; P = .027). Once again, outliers did not influence this. Controlling simultaneously for age, sex and pain duration in this analysis produced a comparable result ($F_{(1. 131)}$ = 5.219; P = .024).

Based on the results described above, one may conclude that the hypothesis of a between-subjects effect (between IG and CG) can be confirmed for the MPI-DLV scale Interference. This indicates that the IG functioned better than the CG. The hypothesized time x group interaction was not quite significant in the multivariate test. However, the quadratic component was significant which may indeed indicate different interference courses between the two groups. The mechanisms seemed to contradict nature. The CG showed a slight deterioration between t1 and t2 and an improvement between t2 and t3, whereas the IG started off with a relatively large improvement from t1 to t2 followed by a smaller improvement between t2 and t3. The time course of the IG group, however, was completely in line with the a priori hypothesis.

Similar procedures were executed on all other dependent variables. Since this study was only powered on changes with regard to MPI-DLV Interference, significant effects did not necessarily have to emerge on the other dependent variables. In Table 2.3, the mean scores and standard deviations are presented for all variables. In Table 2.4, the results of the ANCOVAs on these variables are presented. Note that t0 was used as a covariate in all analyses.

Table 2.3: Mean (SD) of the investigated variables

Instrument		Group	N	t0	t1	t2	t3
MPI-DLV	Pain severity	IG	80	4.29 (0.94)	4.29 (0.97)	3.83 (1.20)	4.07 (1.18)
		CG	57	4.25 (0.93)	4.14 (1.09)	4.25 (1.01)	4.14 (1.02)
	Interference	IG	81	4.37 (0.76)	4.25 (0.77)	4.05 (0.86)	4.02 (0.90)
		CG	57	4.29 (1.06)	4.24 (1.01)	4.29 (1.01)	4.17 (1.00)
	Life control	IG	80	3.35 (1.15)	3.52 (1.11)	3.48 (1.23)	3.56 (1.30)
		CG	57	3.80 (1.47)	3.93 (1.28)	3.84 (1.38)	3.95 (1.30)
	Affective distress	IG	81	3.31 (1.21)	3.16 (1.14)	2.94 (1.36)	3.11 (1.44)
		CG	57	2.92 (1.45)	2.79 (1.38)	2.87 (1.25)	2.71 (1.23)
	Support	IG	76	4.20 (1.36)	4.15 (1.39)	4.10 (1.28)	4.02 (1.32)
		CG	53	4.52 (1.39)	4.41 (1.38)	4.49 (1.33)	4.48 (1.28)
	Punishing responses	IG	80	1.45 (1.43)	1.51 (1.21)	1.42 (1.27)	1.55 (1.33)
		CG	56	1.34 (1.57)	1.46 (1.38)	1.49 (1.45)	1.49 (1.45)
	Solicitous responses	IG	80	3.14 (1.33)	3.03 (1.35)	2.94 (1.18)	3.04 (1.23)
		CG	56	3.31 (1.38)	3.39 (1.29)	3.38 (1.28)	3.25 (1.26)
	Distracting responses	IG	80	3.16 (1.61)	3.27 (1.44)	3.41 (1.42)	3.29 (1.35)
		CG	55	3.22 (1.57)	3.36 (1.54)	3.41 (1.50)	3.28 (1.61)
	Household chores	IG	81	3.57 (1.37)	3.55 (1.36)	3.65 (1.31)	3.51 (1.26)
		CG	57	3.73 (1.49)	3.60 (1.50)	3.63 (1.53)	3.66 (1.54)
	Outdoor work	IG	81	1.17 (1.19)	1.21 (1.08)	1.22 (1.19)	1.27 (1.24)
		CG	56	1.15 (1.20)	1.18 (1.05)	0.95 (0.98)	1.01 (1.07)
	Social activities	IG	81	2.40 (0.76)	2.44 (0.86)	2.45 (0.85)	2.40 (0.91)
		CG	57	2.41 (1.01)	2.33 (0.99)	2.30 (0.92)	2.27 (0.95)
	General activities	IG	81	2.38 (0.71)	2.40 (0.66)	2.44 (0.66)	2.39 (0.75)
		CG	56	2.41 (0.84)	2.38 (0.83)	2.31 (0.83)	2.32 (0.82)
SCL-90	Fear	IG	80	17.78 (6.40)	17.88 (6.80)	17.90 (6.67)	17.45 (6.21)
		CG	56	16.64 (6.62)	16.59 (6.03)	16.72 (5.90)	16.68 (6.13)
	Agoraphobia	IG	80	9.83 (4.29)	9.82 (3.91)	10.18 (4.62)	9.73 (4.44)
		CG	56	9.25 (3.93)	8.97 (3.89)	8.96 (4.02)	9.19 (4.35)
	Depressive thoughts	IG	78	33.88 (10.84)	32.99 (10.95)	31.79 (11.50)	31.62 (10.96)
		CG	56	30.92 (9.88)	30.56 (9.62)	30.47 (9.79)	30.53 (10.14)
	Somatic complaints	IG	79	30.57 (8.02)	30.61 (7.88)	29.61 (8.12)	29.46 (7.93)
		CG	40	28.58 (7.23)	27.95 (8.08)	28.37 (7.55)	28.99 (7.51)
	Insufficiency	IG	78	21.80 (6.18)	22.17 (6.22)	21.15 (6.30)	21.23 (5.96)
		CG	56	19.73 (5.47)	19.80 (6.24)	19.60 (6.27)	19.74 (6.52)
	Sensitivity	IG	78	30.71 (11.95)	30.38 (11/05)	30.14 (10.01)	29.04 (9.27)
		CG	56	27.40 (10.39)	26.82 (9.45)	27.91 (11.98)	27.86 (12.90)
	Hostility	IG	80	9.41 (3.15)	8.89 (2.86)	8.75 (3.04)	8.84 (3.17)
		CG	56	8.15 (2.31)	8.25 (2.57)	8.01 (2.18)	8.25 (2.96)
	Sleeping problems	IG	80	8.45 (3.03)	8.64 (3.02)	8.33 (3.24)	8.41 (3.30)
		CG	57	8.85 (3.51)	9.11 (3.47)	8.90 (3.21)	9.01 (3.41)
	Psycho-neuroticism	IG	78	176.75 (46.12)	175.30 (46.26)	171.97 (48.34)	169.79 (45.32)
		CG	56	162.81 (42.56)	161.96 (42.19)	162.06 (43.20)	163.42 (47.89)

Table 2.4: Treatment effects on the investigated variables (t1-t2-t3)

Instrument		Time effect	Group effect	Time x group interaction	df	Contrast linear	Contrast quadratic
MPI-DLV	Pain severity	F=.521 *P* =.594	F=.827 *P* =.365	F=5.673 *P* =.004	1, 133	F=2.813 *P* =.096	F=8.009 *P* =.005
	Interference	F=.203 *P* =.806	F=5.061 *P* =.026	F=2.231 *P* =.112	1, 134	F=1.193 *P* =.277	F=3.842 *P* =.052
	Life control	F=.292 *P* =.748	F=.627 *P* =.430	F=.027 *P* =.973	1, 133	F=.012 *P* =.913	F=.044 *P* =.834
	Affective distress	F=1.150 *P* =.318	F=.504 *P* =.479	F=1.002 *P* =.369	1, 134	F=.010 *P* =.920	F=1.918 *P* =.168
	Support[1]	F=.222 *P* =.790	F=.025 *P* =.875	F=1.433 *P* =.241	1, 125	F=2.079 *P* =.152	F=.413 *P* =.522
	Punishing responses	F=1.223 *P* =.296	F=.341 *P* =.560	F=.460 *P* =.631	1, 132	F=.114 *P* =.736	F=.812 *P* =.369
	Solicitous responses	F=2.271 *P* =.112	F=1.585 *P* =.210	F=2.132 *P* =.127	1, 131	F=1.226 *P* =.270	F=3.916 *P* =.050
	Distracting responses	F=.924 *P* =.384	F=.008 *P* =.928	F=.602 *P* =.520	1, 131	F=.805 *P* =.371	F=.283 *P* =.596
	Household chores	F=1.286 *P* =.277	F=.064 *P* =.800	F=.682 *P* =.493	1, 134	F=.404 *P* =.526	F=1.211 *P* =.273
	Outdoor work	F=.936 *P* =.383	F=1.454 *P* =.230	F=1.585 *P* =.210	1, 132	F=.956 *P* =.330	F=2.785 *P* =.098
	Social activities	F=.694 *P* =.493	F=2.032 *P* =.156	F=.123 *P* =.874	1, 134	F=.140 *P* =.708	F=.094 *P* =.759
	General activities	F=.457 *P* =.634	F=.940 *P* =.334	F=1.090 *P* =.339	1, 133	F=.044 *P* =.835	F=2.161 *P* =.144
SCL-90	Fear	F=2.896 *P* =.058	F=.274 *P* =.601	F=.099 *P* =.903	1, 132	F=.055 *P* =.815	F=.146 *P* =.703
	Agoraphobia	F=.551 *P* =.551	F=1.172 *P* =.281	F=.958 *P* =374	1, 131	F=.201 *P* =.655	F=2.139 *P* =.146
	Depressive thoughts	F=.619 *P* =.539	F=.181 *P* =.671	F=.256 *P* =.774	1, 129	F=.460 *P* =.499	F=.022 *P* =.882
	Somatic complaints	F=.187 *P* =.812	F=.022 *P* =.881	F=1.414 *P* =.245	1, 114	F=2.085 *P* =.151	F=.256 *P* =.614
	Insufficiency	F=.811 *P* =.439	F=.122 *P* =.727	F=.291 *P* =.734	1, 130	F=.321 *P* =.572	F=.246 *P* =.621
	Sensitivity	F=1.870 *P* =.161	F=.089 *P* =.765	F=1.093 *P* =.332	1, 130	F=1.690 *P* =.196	F=.059 *P* =.809
	Hostility	F=.185 *P* =.809	F=.821 *P* =.367	F=.858 *P* =.415	1, 131	F=.182 *P* =.671	F=1.914 *P* =.169
	Sleeping problems	F=.486 *P* =.603	F=.313 *P* =.577	F=.254 *P* =.762	1, 133	F=.403 *P* =.527	F=.005 *P* =.945
	Psycho-neuroticism	F=.698 *P* =.484	F=.108 *P* =.742	F=.650 *P* =.508	1, 129	F=.971 *P* =.326	F=.181 *P* =.672

[1] After removing six outliers in the IG and one outlier in the CG a significant time x group interaction ($F_{(1, 118)}$ = 3.151; P = .046) was observed for the variable support.

An additional series of analyses were performed to compare only t1 and t2 (pure treatment effect). In Table 2.5, the results of these ANOVAs are presented.

Table 2.5: Pure treatment (group) effects on the investigated variables (t1-t2)

Instrument		F value	*P* value	df
MPI-DLV	Pain severity	9.893	.002	1, 134
	Interference	5.017	.027	1, 135
	Life control	.597	.441	1, 134
	Affective distress	.494	.483	1, 135
	Support	1.294	.257	1, 129
	Punishing responses	.586	.445	1, 134
	Solicitous responses	1.325	.252	1, 134
	Distracting responses	.083	.774	1, 134
	Household chores	.131	.718	1, 135
	Outdoor work	4.113	.045	1, 135
	Social activities	.164	.686	1, 135
	General activities[1]	2.160	.144	1, 135
SCL-90	Fear	.256	.614	1, 133
	Agoraphobia	1.418	.236	1, 132
	Depressive thoughts	.001	.977	1, 131
	Somatic complaints	.002	.965	1, 115
	Insufficiency	.068	.795	1, 131
	Sensitivity	.546	.461	1, 131
	Hostility	1.610	.207	1, 133
	Sleeping problems	.313	.577	1, 134
	Psycho-neuroticism	.008	.930	1, 131

[1] After removing one outlier in the CG a significant group effect was found for the variable general activities ($F_{(1, 134)} = 4.172$; $P = .043$).

All analyses were checked for the presence of outliers. Only outliers in the MPI-DLV General Activities changed the results. For this variable, no specific reasons could be determined for the outlying cases. The variables marital status, age, sex, pain duration, diagnosis, work status and MPI-DLV classification on t0 were included and did not seem deviated. Further, all analyses were repeated with age, sex and pain duration as covariates to increase power. This resulted in comparable findings.

Finally, a multivariate analysis was carried out over t1, t2 and t3 to investigate the efficacy of treatment on all MPI-DLV variables together. The results of this analysis demonstrated a significant time x group interaction ($F_{(1, 113)} = 5.394$; $P = .006$). This implies a large difference, over time, between IG and CG. In other words, the treatment had succeeded in changing pain-related psychosocial behavior of the IG on a multivariate level as compared to the CG. No significant

between-group effect was found ($F_{(1, 113)} = 1.622$; $P = .205$). Adding the covariates age, sex and pain duration into this analysis showed an almost identical time x group interaction ($F_{(1, 109)} = 5.638$; $P = .005$). On the univariate level, the following variables contributed significantly to this multivariate time x group interaction: MPI-DLV scale for Interference ($F_{(1, 114)} = 10.110$; $P = .002$) and for Life Control ($F_{(1, 114)} = 5.811$; $P = .018$). MPI-DLV scale Pain Severity contributed a relevant multivariate effect but not one that could be considered significant ($F_{(1, 114)} = 3.496$; $P = .064$).

2.3.3 Treatment results: outpatient versus semi-inpatient treatment

All data were reanalyzed with the treatment variant as a between-group factor to see whether there was a difference in treatment outcome between the outpatient and semi-inpatient treatment mode. The means and standard deviations on t0 through t3 are presented in Table 2.6. Only the results of the IG group were used for the analyses since only the IG had followed treatment. Note that the outpatient group (OG) is much smaller than the semi-inpatient group (SG) (respectively n = 15 and n = 66).

The OG and SG differed from each other with regard to education level ($X^2 = 14.095$; $P = .003$), civil status ($X^2 = 12.348$; $P = .030$) and pain location ($X^2 = 11.894$; $P = .018$). The SG was relatively better educated. The OG consisted of relatively more back pain patients but the SG had relatively more patients who experienced pain complaints on more than one location. Because of these differences, subsequent analyses were not only controlled for t0, but also for education level, civil status and pain location. All analyses were checked for the presence of outliers. The results of all analyses indicated that, although patients from the OG seemed to improve somewhat more than patients from the SG, these effects did not reach consistent significance.

Table 2.6: Mean (SD) of the investigated variables(t1-t2-t3)

Instrument		Mode	N	t0	t1	t2	t3
MPI-DLV	Pain severity	OG	15	4.30 (0.88)	4.23 (0.75)	3.53 (1.29)	3.97 (1.10)
		SG	65	4.28 (0.95)	4.30 (1.01)	3.89 (1.18)	4.10 (1.20)
	Interference	OG	15	3.92 (0.87)	3.90 (0.63)	3.47 (0.95)	3.31 (0.99)
		SG	66	4.47 (0.70)	4.33 (0.78)	4.18 (0.79)	4.19 (0.81)
	Life control	OG	15	4.19 (0.91)	4.02 (1.20)	3.69 (1.08)	3.89 (1.04)
		SG	65	3.16 (1.12)	3.41 (1.06)	3.43 (1.27)	3.48 (1.35)
	Affective distress	OG	15	3.36 (1.22)	2.93 (1.08)	2.42 (1.10)	2.98 (1.41)
		SG	66	3.30 (1.22)	3.21 (1.15)	3.06 (1.39)	3.14 (1.46)
	Support	OG	14	4.76 (0.80)	4.31 (1.55)	4.66 (0.94)	4.71 (0.90)
		SG	62	4.07 (1.44)	4.11 (1.36)	3.97 (1.33)	3.85 (1.35)
	Punishing responses	OG	15	1.44 (1.17)	1.07 (1.23)	1.23 (1.05)	1.67 (1.01)
		SG	65	1.45 (1.49)	1.61 (1.01)	1.47 (1.32)	1.52 (1.40)
	Solicitous responses	OG	15	3.18 (1.35)	2.90 (1.22)	3.32 (1.27)	3.53 (1.22)
		SG	65	3.13 (1.34)	3.06 (1.38)	2.86 (1.15)	2.93 (1.21)
	Distracting responses	OG	15	3.13 (1.37)	3.62 (1.25)	3.80 (1.16)	3.82 (1.20)
		SG	65	3.17 (1.67)	3.19 (1.47)	3.32 (1.46)	3.17 (1.37)
	Household chores	OG	15	3.68 (1.42)	3.35 (1.43)	3.65 (1.37)	3.45 (1.24)
		SG	66	3.54 (1.37)	3.60 (1.35)	3.65 (1.31)	3.53 (1.28)
	Outdoor work	OG	15	1.89 (1.39)	1.54 (1.13)	1.95 (1.53)	1.76 (1.34)
		SG	66	1.01 (1.09)	1.13 (1.06)	1.06 (1.05)	1.16 (1.20)
	Social activities	OG	15	2.31 (0.92)	2.45 (1.11)	2.55 (0.90)	2.47 (1.05)
		SG	66	2.43 (0.72)	2.44 (0.80)	2.43 (0.84)	2.38 (0.89)
	General activities	OG	15	2.62 (0.70)	2.45 (0.55)	2.71 (0.53)	2.56 (0.67)
		SG	66	2.33 (0.70)	2.39 (0.69)	2.38 (0.68)	2.35 (0.77)
SCL-90	Fear	OG	14	14.38 (4.05)	15.93 (5.29)	15.82 (4.59)	15.00 (3.57)
		SG	66	18.50 (6.60)	18.30 (7.04)	18.37 (7.00)	18.01 (6.56)
	Agoraphobia	OG	14	8.06 (1.87)	8.14 (1.79)	7.73 (1.66)	7.37 (1.01)
		SG	66	10.20 (4.57)	10.17 (4.15)	10.73 (4.90)	10.27 (4.74)
	Depressive thoughts	OG	13	25.95 (7.35)	28.46 (8.61)	25.81 (7.32)	26.11 (6.08)
		SG	65	35.47 (10.77)	33.89 (11.20)	33.06 (11.86)	32.79 (11.43)
	Somatic complaints	OG	13	25.08 (7.16)	29.62 (9.72)	28.32 (6.84)	28.01 (7.79)
		SG	66	31.66 (7.77)	30.80 (7.54)	29.88 (8.39)	29.77 (7.98)
	Insufficiency	OG	13	17.40 (4.55)	19.69 (6.56)	17.54 (6.22)	17.86 (5.04)
		SG	65	22.68 (6.11)	22.66 (6.08)	21.92 (6.09)	21.94 (5.93)
	Sensitivity	OG	13	24.48 (6.58)	26.97 (8.95)	24.38 (5.71)	24.12 (4.85)
		SG	65	31.96 (12.41)	31.07 (11.36)	31.37 (10.33)	30.09 (9.67)
	Hostility	OG	14	7.71 (1.82)	8.00 (2.57)	8.00 (2.80)	8.20 (3.49)
		SG	66	9.77 (3.26)	9.08 (2.90)	8.93 (3.09)	8.98 (3.12)
	Sleeping problems	OG	14	6.71 (2.16)	7.79 (2.69)	7.70 (3.19)	7.60 (3.31)
		SG	66	8.82 (3.07)	8.82 (3.07)	8.48 (3.26)	8.59 (3.29)
	Psycho-neuroticism	OG	13	142.46 (27.67)	157.20 (39.92)	147.50 (32.09)	147.02 (29.11)
		SG	65	183.60 (46.15)	178.91 (46.87)	177.16 (49.78)	174.62 (46.81)

2.3.4 Comparison RCT results with Winter's study

The primary aim of this study was to replicate the results found by Winter by means of a randomized study. The period of time in which Winter's measurements were conducted (from the start of treatment to eight weeks after treatment) was approximately the same as the period between t1 and t3. The present study's selection criteria were also identical to that of Winter. It is important to note that Winter's research population was comprised only of semi-inpatient patients[7]. Therefore, Winter's results can only be compared with the results of patients from the semi-inpatient treatment in the present study. In order to effectively compare Winter's results with the present study, it was necessary that the two research populations were comparable on as many aspects as possible. In Table 2.7, the general characteristics of both populations are represented.

Table 2.7: Comparison of population characteristics

Variable		**Study Winter (N = 185)**	**Present study (N = 66)**
Sex	Male	44 (23.8%)	11 (16.7%)
	Female	141 (76.2%)	55 (83.3%)
Age	M (SD)	43 (9.4)	42.3 (10.5)
	Range	19-65	19-65
Civil status	Single	11%	9.1%
	Married	81%	86.4%
	Divorced	6%	4.5%
	Widow(er)	2%	-
Education	LO	14%	13.6%
	LBO	42%	27.3%
	MAVO/MBO	32%	48.5%
	HAVO/VWO/HBO/WO	13%	10.6%
Pain duration	< 1 year	5%	12.1%
	2-4 years	25%	25.8%
	5-10 years	31%	28.8%
	> 10 years	40%	33.3%
Pain location	More than 1 place	63.8%	69.7%
	Neck/shoulder	3.8%	1.5%
	Back	24.9%	16.7%
	Other/no classification	5.4%	12.1%
Pain severity at t1 (VAS)	M (SD)	7 (?)	6.3 (2.07)

Despite some small differences, both populations were quite comparable with respect to general characteristics. Consequently, the possibility that differences in population characteristics were responsible for the differences in effect between the two studies can be ruled out.

So far, the results of the present study only demonstrate significant results for the MPI-DLV scales Pain Severity and Interference. Winter's study, however, showed significant results for not only MPI-DLV Pain Severity and Interference, but also for the MPI-DLV scales Affective Distress, Support, Solicitous Responses, Social Activities and General Activities. On the MPI-DLV scales Punishing Responses, Distracting Responses, Household Chores and Outdoor Work no significant results were obtained. Further, Winter demonstrated significant decreases on every scale of the SCL-90, particularly with regard to Fear, Depressive Thoughts, Somatic Complaints, Hostility and Psycho-neuroticism. To make an honest comparison between the two studies, treatment effects were calculated for the present study (in the same way as was done by Winter) by means of t-tests over t1 and t3. The results of these tests are presented in Table 2.8.

Table 2.8 shows that, in the present study, the semi-inpatient treatment produced only small, mainly non-significant effects. This finding is quite different from Winter's findings (Table 2.1). Analyses were carried out to determine whether the mean difference scores of the present study differed from the mean difference scores in Winter's study. For the MPI-DLV variables Support, Punishing Responses, Solicitous Responses, Distracting Responses, Household Chores, and Outdoor Work, no significant differences were found. For all other variables, the present study showed significantly smaller effects compared to Winter's study (all P's< 0.05).

In comparison with the treatment results of the SG, the results of the OG demonstrated a larger improvement (Table 2.9). Except for MPI-DLV Solicitous Responses and SCL-90 Hostility, none of the mean differences from the present study differed significantly from the mean differences found in Winter's study. Therefore, the results of the OG, in contrast to the results from the SG, seem to be more in line with the Winter's findings. However, it is important that these results are interpreted with a great degree of caution as the number of outpatient respondents in the present study was quite small.

Table 2.8: Semi-inpatient treatment effects of present study (t1 - t3)

Instrument		N	M (SD) t1	M (SD) t3	M (SD) Δ_{t1-t3}	CI 95%	t value	*P* value
MPI-DLV	Pain severity	66	4.30 (1.01)	4.10 (1.20)	0.20 (1.05)	-0.05 - 0.46	1.579	.119
	Interference	66	4.33 (0.78)	4.19 (0.81)	0.15 (0.71)	-0.03 - 0.32	1.679	.098
	Life control	66	3.41 (1.06)	3.48 (1.35)	-0.07 (0.18)	-0.36 - 0.22	-.503	.617
	Affective distress	66	3.21 (1.15)	3.14 (1.46)	0.07 (1.10)	-0.20 - 0.34	.521	.604
	Support	62	4.11 (1.36)	3.85 (1.35)	0.26 (0.92)	0.03 - 0.50	2.231	.029
	Punishing responses	66	1.62 (1.23)	1.52 (1.40)	0.10 (1.01)	-0.15 - 0.34	.776	.441
	Solicitous responses	66	3.06 (1.38)	2.93 (1.21)	0.14 (0.85)	-0.07 - 0.35	1.310	.195
	Distracting responses	66	3.19 (1.47)	3.17 (1.37)	0.02 (1.00)	-0.22 - 0.26	.171	.865
	Household chores	66	3.60 (1.35)	3.53 (1.28)	0.07 (1.10)	-0.20 - 0.34	.527	.600
	Outdoor work	66	1.13 (1.06)	1.16 (1.20)	-0.03 (1.03)	-0.28 - 0.23	-.203	.840
	Social activities	66	2.44 (0.80)	2.38 (0.89)	0.06 (0.74)	-0.12 - 0.24	.692	.492
	General activities	66	2.39 (0.69)	2.35 (0.77)	0.03 (0.64)	-0.12 - 0.19	.441	.660
SCL-90	Fear	66	18.30 (7.04)	18.01 (6.56)	0.29 (4.70)	-0.87 - 1.44	.495	.622
	Agoraphobia	66	10.17 (4.15)	10.27 (4.74)	-0.10 (2.99)	-0.84 - 0.64	-.268	.789
	Depressive thoughts	65	33.89 (11.20)	32.89 (11.49)	1.00 (6.97)	-0.72 - 2.73	1.159	.251
	Somatic complaints	66	30.80 (7.54)	29.77 (7.98)	1.03 (5.49)	-0.32 - 2.38	1.530	.131
	Insufficiency	65	22.66 (6.08)	22.00 (5.96)	0.66 (4.42)	-0.43 - 1.76	1.211	.230
	Sensitivity	65	31.07 (11.36)	30.20 (9.70)	0.87 (6.11)	-0.64 - 2.38	1.147	.256
	Hostility	66	9.08 (2.90)	8.98 (3.11)	0.10 (2.14)	-0.42 - 0.63	.386	.701
	Sleeping problems	66	8.82 (3.07)	8.59 (3.29)	0.23 (2.94)	-0.50 - 0.95	.628	.532
	Psycho-neuroticism	65	178.91 (46.87)	175.21 (46.92)	3.70 (26.84)	-2.95 - 10.35	1.112	.270

Table 2.9: Outpatient treatment effects of present study (t1 - t3)

Instrument		N	M (SD) t1	M (SD) t3	M (SD) Δ_{t1-t3}	CI 95%	t value	*P* value
MPI-DLV	Pain severity	15	4.23 (0.75)	3.97 (1.09)	0.27 (1.15)	-0.37 - 0.90	.900	.383
	Interference	15	3.90 (0.63)	3.31 (0.99)	0.59 (0.88)	0.10 - 1.08	2.590	.021
	Life control	15	4.02 (1.20)	3.89 (1.04)	0.13 (1.54)	-0.72 - 0.99	.333	.744
	Affective distress	15	2.93 (1.09)	2.98 (1.41)	-0.05 (1.84)	-1.07 - 0.98	-.095	.925
	Support	15	4.31 (1.55)	4.71 (0.90)	-0.40 (1.41)	-1.18 - 0.38	-1.101	.289
	Punishing responses	15	1.07 (1.01)	1.67 (1.01)	-0.60 (1.28)	-1.31 - 0.11	-1.814	.091
	Solicitous responses	15	2.90 (1.22)	3.53 (1.22)	-0.63 (1.00)	-1.18 - -0.08	-2.453	.028
	Distracting responses	15	3.62 (1.25)	3.82 (1.20)	-0.20 (1.50)	-1.03 - 0.63	-.516	.614
	Household chores	15	3.35 (1.43)	3.45 (1.24)	-0.11 (0.86)	-0.58 - 0.36	-.485	.635
	Outdoor work	15	1.54 (1.13)	1.76 (1.34)	-0.22 (1.15)	-0.86 - 0.41	-.754	.463
	Social activities	15	2.45 (1.11)	2.47 (1.05)	-0.02 (0.68)	-0.39 - 0.36	-.087	.932
	General activities	15	2.45 (0.55)	2.56 (0.67)	-0.12 (0.54)	-0.41 - 0.18	-.834	.418
SCL-90	Fear	14	15.93(5.29)	14.57 (3.27)	1.36 (3.91)	-0.90 - 3.62	1.298	.217
	Agoraphobia	14	8.14 (1.79)	7.39 (1.04)	0.75 (1.37)	-0.04 - 1.54	2.049	.061
	Depressive thoughts	13	28.46 (8.61)	25.50 (5.87)	2.96 (6.21)	-0.79 - 6.71	1.718	.112
	Somatic complaints	13	29.62 (9.72)	26.48 (5.47)	3.14 (6.98)	-1.08 - 7.36	1.623	.131
	Insufficiency	13	19.69 (6.56)	17.46 (5.01)	2.23 (3.30)	0.24 - 4.22	2.441	.031
	Sensitivity	13	26.97 (8.95)	24.35 (4.96)	2.62 (5.35)	-0.62 - 5.85	1.763	.103
	Hostility	14	8.00 (2.57)	7.57 (2.59)	0.43 (1.45)	-0.41 - 1.27	1.104	.290
	Sleeping problems	14	7.79 (2.69)	7.07 (2.70)	0.71 (2.76)	-0.88 - 2.31	.969	.350
	Psycho-neuroticism	13	157.20 (39.92)	142.63 (25.03)	14.57 (23.90)	0.13 - 29.01	2.198	.048

2.3.5 Treatment variability

In this paragraph, treatment variability in the IG is demonstrated. It is important to recognize that, in addition to standard deviation (see Table 2.6), the range of variables also provide an indication of treatment variability. In Table 2.10, the range of variables present during a given time of measurement is presented.

Table 2.10: Range of scores (N = 81)

Instrument		t0	t1	t2	t3
MPI-DLV	Pain severity	2.00-6.00	2.00-6.00	1.00-6.00	1.00-6.00
	Interference	1.50-5.67	2.13-5.78	1.63-6.00	1.67-6.00
	Life control	0.00-5.33	1.00-6.00	0.33-5.67	0.67-6.00
	Affective distress	0.67-6.00	0.33-6.00	0.33-6.00	0.33-6.00
	Support	0.00-6.00	0.00-6.00	0.00-6.00	0.00-6.00
	Punishing responses	0.00-6.00	0.00-5.00	0.00-6.00	0.00-6.00
	Solicitous responses	0.00-5.50	0.00-6.50	0.00-6.00	0.00-5.67
	Distracting responses	0.00-6.00	0.00-6.00	0.00-6.00	0.00-6.00
	Household chores	0.60-5.80	0.40-6.00	0.60-6.00	0.20-6.00
	Outdoor work	0.00-4.75	0.00-4.00	0.00-4.40	0.00-5.80
	Social activities	0.75-4.63	0.38-4.38	0.50-4.63	0.13-5.13
	General activities	0.85-3.93	0.96-3.81	0.78-4.09	0.18-5.02
SCL-90	Fear	10.00-38.00	10.00-40.00	10.00-39.00	10.00-41.00
	Agoraphobia	7.00-26.00	7.00-23.00	7.00-24.00	7.00-27.00
	Depressive thoughts	17.00-57.00	17.00-71.00	16.00-65.00	17.00-71.00
	Somatic complaints	13.00-50.00	13.00-47.00	16.00-48.00	15.00-50.00
	Insufficiency	11.00-37.00	10.00-40.00	11.00-38.00	12.00-38.00
	Sensitivity	18.00-87.00	18.00-66.00	18.00-61.00	18.00-60.00
	Hostility	6.00-21.00	6.00-17.00	6.00-18.00	6.00-20.00
	Sleeping problems	3.00-15.00	3.00-15.00	3.00-15.00	3.00-15.00
	Psycho-neuroticism	98.00-295.00	98.00-318.00	102.00-297.00	101.00-308.00

Another way to demonstrate treatment variability is to look at the variability of difference scores between two occasions upon which measurements were conducted. In Table 2.11, the following is presented: a) the mean differences between t1 and t3; b) the mean differences between t1 and t2; c) standard deviations (SD); and d) the ranges of the mean differences of the IG. Table 2.11 shows that a remarkable range around the value of the mean differences exists.

Table 2.11: Mean differences (Δ(SD)) and ranges of the IG (N = 81)

Instrument		Δ_{t1-t3}	Range	Δ_{t1-t2}	Range
MPI-DLV	Pain severity	0.22 (1.06)	-2.50-2.50	0.44 (1.09)	-3.50-3.00
	Interference	0.23 (0.76)	-1.56-2.33	0.20 (0.67)	-1.56-2.37
	Life control	-0.04 (1.25)	-3.00-3.00	0.04 (1.13)	-3.33-2.67
	Affective distress	0.05 (1.26)	-5.00-2.67	0.22 (1.38)	-2.67-3.66
	Support	0.13 (1.06)	-5.00-2.67	0.05 (0.96)	-5.00-2.34
	Punishing response	-0.03 (1.09)	-2.67-2.34	0.09 (1.11)	-3.33-3.34
	Solicitous response	0.01 (0.93)	-2.67-2.83	0.09 (0.88)	-2.50-2.50
	Distracting response	-0.02 (1.10)	-3.33-2.67	-0.14 (1.06)	-4.33-2.33
	Household chores	0.04 (1.05)	-2.80-3.60	-0.09 (0.91)	-5.00-2.40
	Outdoor work	-0.06 (1.05)	-3.80-4.00	-0.01 (0.92)	-2.66-4.00
	Social activities	0.05 (0.73)	-1.75-1.75	-0.07 (0.61)	-1.38-1.75
	General activities	0.07 (0.62)	-2.19-1.70	-0.04 (0.50)	-1.46-1.35
SCL-90	Fear	0.47 (4.57)	-17.00-22.00	-0.05 (4.54)	-19.00-14.00
	Agoraphobia	0.05 (2.79)	-19.00-6.00	-0.40 (2.38)	-9.00-5.00
	Depressive thoughts	1.33 (6.85)	-19.00-23.00	1.00 (6.83)	-23.00-17.00
	Somatic complaints	1.38 (5.77)	-19.00-16.00	1.06 (5.49)	-22.00-15.00
	Insufficiency	0.93 (4.28)	-10.00-13.00	0.93 (4.20)	-11.00-13.00
	Sensitivity	1.16 (5.99)	-9.00-31.00	0.06 (5.66)	-24.00-14.12
	Hostility	0.16 (2.03)	-5.00-5.00	0.11 (2.07)	-10.00-5.00
	Sleeping problems	0.31 (2.90)	-6.00-8.00	0.39 (2.49)	-5.00-10.00
	Psycho-neuroticism	5.51 (26.54)	-76.00-122.00	2.66 (26.22)	-88.02-70.00

A final way to illustrate treatment variability is by means of a histogram. Figure 2.3 presents the IG difference scores for the MPI-DLV scale Interference. Figure 2.4 presents the IG difference scores for the SCL-90 scale Somatic Complaints. Note that, on a mean group level, both variables demonstrate a significant mean difference, namely 0.23 for the MPI-DLV Interference scale ($t_{(1, 80)} = 2.714$; $P = .008$) and 1.38 for the SCL-90 Somatic Complaints scale ($t_{(1, 78)} = 2.129$; $P = .036$). Please note that negative values imply, for the Interference scale, an improvement, but for the Somatic Complaints scale, a worsening and vice versa. Based on the histograms and the statistics, one may conclude that a significant improvement for both variables can be seen on a *mean* level. However, on an *absolute* level, it becomes evident that a considerable number of patients did not improve but had actually worsened.

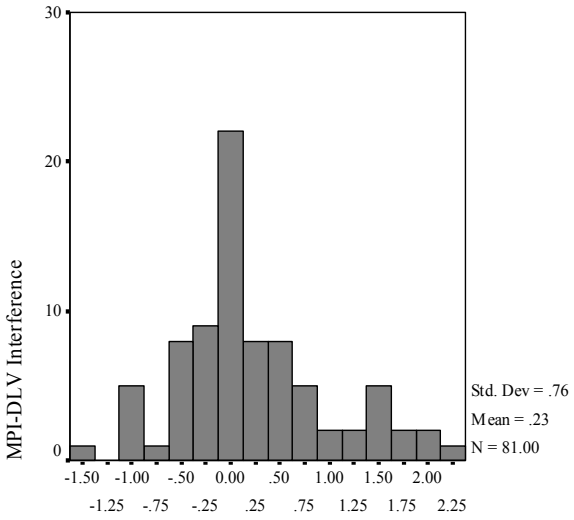


Figure 2.3: Histogram of the difference scores between t1 and t3 on MPI-DLV Interference

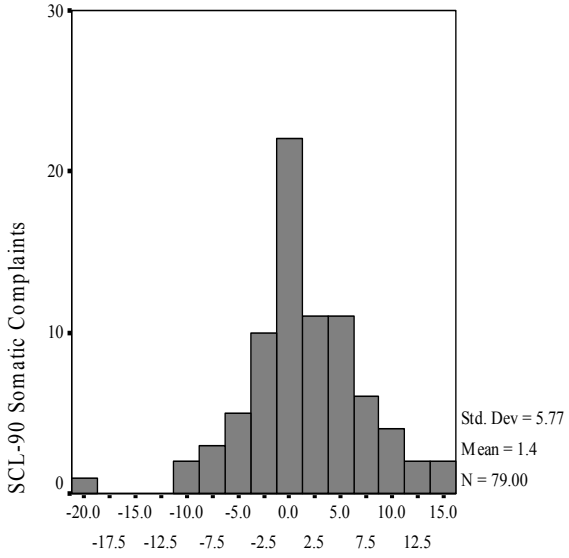


Figure 2.4: Histogram of the difference scores between t1 and t3 on SCL-90 Somatic Complaints

2.4 Discussion

The primary aim of the study described in this chapter was to replicate the findings of Winter (1992). With the exception of the MPI-DLV variables Punishing Responses, Distracting Responses, Household Chores and Outdoor Work, Winter found significant improvement on all other variables of the MPI-DLV and SCL-90. The findings of the present study diverge significantly from Winter's findings. The present study found, when analyzing t1, t2 and t3 by means of ANCOVAs, only one significant group effect for the MPI-DLV Interference and one significant time x group interaction for the MPI-DLV Pain Severity. With respect to MPI-DLV Interference, the effect demonstrated in the present study was approximately half the effect found by Winter. When only t1 and t2 were compared, significant group effects were found for the MPI-DLV variables Pain Severity, Interference and Outdoor Work. Using multivariate analyses, a clear significant difference between the IG and the CG was found indicating that the treatment provided to the IG had succeeded in changing pain-related characteristics and psychosocial behavior on a multivariate level. However, since the absolute treatment effects were small and, moreover, almost half the size of the effects found by Winter[7], the clinical relevance of these effects is questionable. This issue will be explored further in Chapter 3.

Because only small effects were observed, a series of post hoc analyses were performed in order to obtain more insight into the predictive variables for the outcome of the CBT-R program (Schreurs, personal communication). The results of these post hoc analyses showed that the outcome of the treatment may have been influenced by work status, age, pain severity and interference at baseline, in addition to changes in psychological functioning and activity level during treatment. This would imply that the effectiveness of the CBT-R program could be improved by paying more attention to these variables both during the intake procedure and during treatment.

In analysis of only the SG, which is the best means of comparing the present study with Winter's study, one significant effect was found, namely the effect on MPI-DLV Support. Additionally, an effect which was not quite significant was found for the MPI-DLV Interference. With the exception of the MPI-DLV variables Support, Punishing Responses, Solicitous Responses, Distracting Responses, Household Chores and Outdoor Work, the current study's SG demonstrated that the effects for the remainder of the variables were

significantly smaller than the effects found in Winter's study. The results of the OG showed greater improvements than the SG results. In addition to the above, a comparison of the mean differences found by Winter with the results of the OG in the present study showed that, with the exception of MPI-DLV Solicitous Responses and SCL-90 Hostility, the mean differences did not significantly differ from each other. Therefore, the results found for the OG seem to be more in line with the Winter's findings than the SG results. This may be an indication that inclusion criteria for the different treatment forms have been changed since Winter conducted his study. However, due to the limited number of OG patients, a more extensive comparison between study populations is needed to determine the extent to which inclusion criteria have been changed.

In short, it may be concluded that the results found by Winter could not be replicated by means of the semi-inpatient data in the present study**.** Differences in population characteristics were ruled out as a possible explanation for the inability to replicate Winter's results as the population characteristics for both Winter's study and the present study were almost identical. Several other factors may have been responsible for this failure to replicate Winter's results. Firstly, Winter's research population consisted of 185 patients while the present study used data from only 66 patients. It may be possible that the population from the present study was too small to be able to produce significant results. This would mean that the present study lacked sufficient power. If the effect sizes of both studies were comparable and Winter found significant results while the RCT did not, then it is possible that the RCT had a power problem. However, the effect sizes were not comparable. Consequently, insufficient power is not the cause for the failure to replicate Winter's findings.

A second potential explanation for the present study's failure to replicate Winter's findings is that the process of pain patient referral to RCR has changed since Winter's study. Post hoc consultations with therapists from the CBT-R program revealed that there may have been changes in the past few years with respect to the kinds of patients that are referred to the RCR. Although most institutions in the Netherlands use the same workbook as Winter as a starting point for treatment and also offer both outpatient or inpatient treatment modules (mainly on an individual basis)[27], the RCR is the only rehabilitation facility that offers semi-inpatient treatment on a group basis. For this reason, it may be possible that pain patients with more serious and severe pain complaints are increasingly being referred to the RCR. Because these patients suffer from more severe pain

complaints than patients at other institutions, it may very well take more time to show improvement in functioning for patients at the RCR

A third possible explanation for the inability to replicate Winter's results is that there were differences between the two groups with respect to both the content of treatment programs and the frequency of treatment sessions. However, a closer investigation of this possibility revealed that not many changes had occurred in the content and frequency of treatment at the RCR in the years between the two studies. The only significant difference was that occupational therapy changed from an individual treatment component on a consultation basis to a weekly recurring treatment component on a group basis. Given that significant changes in content and frequency did not occur, one may conclude that the treatment components in the present study are indeed comparable to the treatment components applied in Winter's study.

An alternative explanation to the above three is linked to differences in the ways in which the two studies were conducted. Winter's study had only one intervention group. Winter used the period between t0 and t1 as the control period. In addition, in Winter's study, no randomization took place. One disadvantage of non-randomization is that results may be overestimated. This may have been the case in Winter's study. Further, non-randomization makes it difficult to attribute findings to the actual treatment program. Factors such as the natural course of complaints could have contributed to the positive results. However, the patients who participate in the CBT-R program are usually patients with long-term pain problems. Therefore, it is unlikely that Winter's significant results can solely be attributed to spontaneous recovery.

A fifth explanation for the divergent findings of the two studies is related to the way in which the questionnaires were handled. In Winter's study, patients filled in the questionnaires for the pre and post measurement while they were present at the pain clinic. In the present study, for all times of measurement patients were requested to fill in the questionnaires at home and to return them by mail anonymously. Unfortunately, many patients in the present study were not motivated to return all the questionnaires they were sent. The option of returning the questionnaires anonymously may have motivated patients of the present study to be more honest in their responses, thereby decreasing social desirability bias. Additionally, it is possible that, given the absence of environmental treatment influences, measurements done by mail are less biased.

A sixth explanation for the difference between Winter's findings and the findings of the present RCT is that Winter's t3 data were measured during a one day follow-up treatment, which took place approximately four weeks earlier than the t3 measurement in the present study. This may have had an impact on the results generated. It could also be hypothesized that the one day follow-up treatment may have positively biased patients' pain judgments as, on this day, patients were in a positive atmosphere together with their fellow group members and under the supervision of their therapists.

Lastly, the process by which treatment occurred may provide an explanation for the non-replication of study results. It is important to investigate the extent to which the CBT-R program was carried out according to protocol. If program contents, execution or goals were not conducted according to the treatment directions, this may have had consequences for the achievement of treatment objectives. In Chapter 4, the results of a process evaluation of the CBT-R program are presented. The extent to which process variables may have been responsible for the lack of results in the present study will be discussed in this chapter.

When compared to results from other studies evaluating the efficacy of CBT for chronic pain, the results from the present RCT appear meager. Morley et al. concluded, in their systematic review, that active psychological treatment based on the principles of CBT was effective relative to waiting list control conditions[3]. They reported that CBT produced significant changes in measures of pain experience, mood/affect, cognitive coping and appraisal, pain behavior and activity level, and social role functioning. The present study generated different results. With the exception of pain experience and interference, no significant results were found in the present study. An obvious explanation for the present study's failure to generate findings comparable to internationally demonstrated treatment effects could not be found. It is likely that the causes for this discrepancy lie in the treatment process. This will be investigated in Chapter 4.

The validity of the RCT results may be questioned, since the data were overanalyzed by performing innumerable analyses without correcting for multiplicity. The primary analysis of the data was focused on determining the effect of the program on MPI-DLV Interference. All other analyses were secondary. Since the RCT was powered on MPI-DLV Interference and the outcome of the primary analysis was already very small, a large treatment effect on the other variables was not to be expected. For the other variables, when analyzed

separately, only very small and mostly insignificant effects were found. Correcting the analysis for multiplicity, for example by using Bonferroni correction procedures, would even have decreased the number of significant results. Therefore, the secondary analyses were performed in a more explorative manner and not corrected for multiplicity.

The second aim of this study was to demonstrate treatment variability using the results of the present randomized study of the CBT-R program. It was clearly demonstrated that indeed a large variability exists with respect to IG group treatment results. Even in the case of a significant effect on the main parameter (MPI-DLV Interference), it was evident that there were still some patients showing deterioration on this variable. It is imperative, if treatment results are to be improved, that causes of, not only variability, but also the deterioration in functioning, be investigated.

Despite the fact that no definite explanations can be given for the discrepancies in results between Winter's study and the present study, the utility of the present semi-inpatient mode of treatment must be appraised as SG patients failed to show significant improvements. It is possible that this failure to improve is because SG patients' pain behavior is more resistant to changes or because SG patients experience more pain problems than OG patients. Another possible reason is that patients were incorrectly categorized. The referral of patients to either SG or OG may not always have happened strictly according to protocol or on the basis of functional level. Occasionally, patients were assigned to either the SG or the OG on the basis of more practical reasons such as the distance required to travel from home to the rehabilitation center or on the requirement of the patients to stay home to care for children. A stricter referral protocol and a clearer definition of referral criteria may help to create more clear-cut, homogeneous groups.

One of the most probable explanations for treatment variability is the multidimensional nature of chronic pain. In pain patients, multiple factors may contribute to the development and maintenance of chronic pain. This supports the biopsychosocial model and treatment approach to chronic pain. In addition, it is well known that when all chronic pain patients are put together as one group, the group is tremendously heterogeneous. It is also known that this heterogeneous group is comprised of relatively homogeneous subgroups. The CBT-R program may be relatively aspecific for these homogeneous subgroups. This can be best illustrated by looking at variables concerning activity. Some patients are

overloaded and thus should decrease the amount of energy expended on daily activities. For others living a sedentary life, an increase in activity is necessary to show improvement. This potential aspecificity of the CBT-R may be a reason why the present study did not succeed in demonstrating mean significant effects on household chores, outdoor work, social activities and general activities. It is thus possible that the results of active and passive patients cancelled each other out. This would mean that there was no match between the heterogeneous aspecific standard treatment program, on the one hand, and the relatively homogeneous specific patient subgroups, on the other hand. With this in mind, we turn to the role of treatment variants. Participants in the CBT-R program were divided in three variants, namely inpatient, semi-inpatient and outpatient. A further division of the chronic pain population into subgroups and an adaptation of the CBT-R program, wherein the specific needs and problems of each subgroup are taken into account, may help reduce treatment variability and thereby enlarge the overall effects of cognitive behavioral treatment of chronic pain. Chapter 3 focuses on this by comparing the results of the CBT-R program based upon the different MPI-DLV classification types.

Acknowledgement

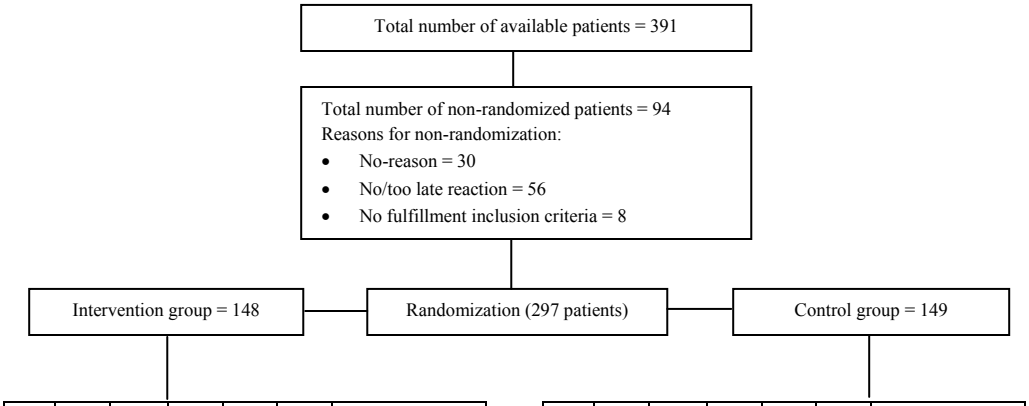
With special thanks to Jitske Gulmans for the collection and analysis of data during her internship.

References

1. Nicholas, M.K., Wilson, P.H., & Goyen, J. (1992). Comparison of cognitive-behavioural group treatment and an alternative non-psychological treatment for chronic low back pain. *Pain, 48,* 339-347.
2. Basler, H.D., Jäkle, C., & Kröner-Herwig, B. (1997). Incorporation of cognitive-behavioral treatment into the medical care of chronic low back patients: a controlled randomized study in German pain treatment centers. *Patient Educ Couns, 31,* 113-124.
3. Morley, S., Eccleston, C., & Williams, A. (1999). Systematic review and meta-analysis of randomized controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache. *Pain, 80,* 1-13.
4. Becker, N., Sjøgren, P., Bech, P., Kornelius Olsen, A., & Eriksen, J. (2000). Treatment outcome of chronic non-malignant pain patient managed in a Danish multidisciplinary pain centre compared to general practice: a randomised controlled trial. *Pain, 84,* 203-211.
5. Nederlands Instituut van Psychologen (NIP) (2000). Sectie Revalidatie. *Nieuwsbrief, 30*, 5-7.
6. Turk, D.C., Meichenbaum, D., & Genest, M. (1983). *Pain and behavioral medicine, a cognitive-behavioral perspective*. New York: Guilford Press.
7. Winter, F.A.M. (1992). *Pijnrevalidatie voor chronische pijnpatiënten: De pijn de baas*. Enschede: Intermap
8. Arrindell, W.A., & Ettema, J.H.M. (1986). *SCL-90: Handleiding bij een multidimensionele psychopathologie-indicator*. Lisse: Swets & Zeitlinger.
9. Lousberg, R., & Groenman, N.H. (1994). *Multidimensionele Pijnvragenlijst - Dutch Language Version (MPI-DLV)*. *Handleiding*. Maastricht: Uitgeverij Stichting Klinisch Pijnonderzoek.
10. Pocock, S.J. (1983). *Clinical trials: A practical approach*. Chichester: John Wiley & Sons Ltd.
11. Jensen, M.P., Romano, J.M., Turner, J.A., Good, A.B., & Wald, L.H. (1999). Patient beliefs predic patient functioning: further support for a cognitive-behavioral model of chronic pain. *Pain, 81,* 95-104.
12. Knapp, R.G., & Miller, M.C. (1992). *Clinical epidemiology and biostatistics*. Baltimore: Williams & Wilkins.
13. Revalidatiecentrum Het Roessingh (RCR) (1996). *De pijn de baas: revalidatieprogramma's voor mensen met chronische pijnklachten. Informatiebrochure*. Enschede: Het Roessingh.
14. Divisie pijnrevalidatie Het Roessingh (1999). *Beslissingsondersteunend en verrichtingenprotocol van het Pijn de Baas programma*. Enschede: Het Roessingh.
15. Bussmann, J.B.J, & Stam, H.J. (1998). Techniques for measurement and assessment of mobility in rehabilitation: a theoretical approach. *Clin Rehabil, 12,* 455-464.
16. Lousberg, R., Breukelen, G.J.P. van, Groenman, N.H., Schmidt, A.J.M., Arntz, A., & Winter, F.A.M. (1999). Psychometric properties of the Multidimensional Pain Inventory, Dutch language version (MPI-DLV). *Behav Res Ther 37*: 167-182.
17. Lousberg, R. (1994). *Chronic pain: multiaxial diagnostics and behavioral mechanisms*. Maastricht: University Press.
18. SPSS (2002). *SPSS Base 11.5 User's guide*. USA: SPSS Inc.
19. Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3 ed.). New Jersey: Lawrence Erlbaum Associates.

20. Tabachnik, B.G., & Fidell, L.S. (1996). *Using multivariate statistics* (3 ed.). New York: Harper Collins College Publishers.
21. Hair, J.F., Anderson, R.E., Tatham, R.L., & Black, W.C. (1998). *Multivariate data analysis*. London: Prentice-Hall International.
22. Unnebrink, K., & Windeler, J. (2001). Intention-to-treat: methods for dealing with missing values in clinical trials of progressively deteriorating diseases. *Stat Med, 20,* 3931-3946.
23. Streiner, D.L. (2002). The case of missing data: methods of dealing with dropouts and other research vagaries. *Can J Psychiatry, 47,* 68.75.
24. Aronoff, G.M., & Evans, W.O. (1982). The prediction of treatment outcome at a multidisciplinary pain center. *Pain, 14,* 67-73.
25. Block, A.R., Kremer, E., & Gaylor, M. (1980). Behavioral treatment of chronic pain: variables affecting treatment efficacy. *Pain, 8,* 367-375.
26. Guck, T.P., Meilman, P.W., Skultety, F.M., & Dowd, E.T. (1986). Prediction of long-term outcome of multidisciplinary pain treatment. *Arch Phys Med Rehab, 67,* 293-296.
27. Winter, F.A.M. (1988). *De pijn de baas: werkboek*. Enschede: Intermap.

Appendix 2A: Flow diagram of patient numbers during the trial



B	SI 66	Ou 16	In 25	Ot 7	Nt 34	Drop-out
t0	66	15	25	7	33	Total 2 • 1 language • 1 motivation
t1	63	15	20	5	25	Total 18 • 10 no CBT-R • 5 ?? • 2 motivation • 1 other
t2	58	15	17	4	22	Total 12 • 4 no CBT-R • 1 ?? • 5 motivation • 2 other
t3	55	14	16	4	19	Total 8 • 4 ?? • 2 motivation • 2 other
t5	44	11	14	2	6	Total 31 • 17 no CBT-R • 10 ?? • 1 motivation • 3 other
Total number of patients who completed study up to t3 = 108 (73.0%)						
Total number of patients who completed study up to t5 = 77 (52.0%)						

B	SI 45	Ou 19	In 22	Ot 11	Nt 52	Drop-out
t0	42	18	22	9	48	Total 10 • 12 language • 1 private • 4 ?? • 2 motivation • 1 other
t1	37	17	21	8	36	Total 20 • 6 no CBT-R • 1 private • 1 ?? • 9 motivation • 3 other
t2	37	16	20	8	29	Total 9 • 1 no CBT-R • 1 ?? • 7 motivation
t3	36	16	18	7	16	Total 17 • 11 no CBT-R • 1 private • 5 other
t4	34	15	15	3	2	Total 24 • 19 no CBT-R • 1 ?? • 2 motivation • 2 other
t5	32	14	12	3	2	Total 6 • 3 ?? • 2 motivation • 1 other
t6	27	11	7	3	1	Total 14 • 1 no CBT-R • 10 ?? • 1 motivation • 2 other
Total number of patients who completed study up to t3 = 93 (62.4%)						
Total number of patients who completed study up to t6 = 49 (32.9%)						

(B=Baseline; SI = Semi-Inpatient treatment; Ou = Outpatient treatment; In = Inpatient treatment; Ot = Other treatment; Nt = No treatment (because of refusal or renouncement)

CHAPTER 3

Cognitive-behavioral treatment of chronic pain: differential effects of MPI-DLV subgroups

3.1 Introduction

Treatment efficacy is usually determined by looking at the mean results of a relatively large sample of patients. However, within a group with a positive mean treatment result, a group of patients who failed to respond to treatment or even deteriorated after treatment can usually be found. Nowadays, one of the hypothesized causes for this treatment variability is that the group of chronic pain patients as a whole is in fact a very heterogeneous group[1]. One way to deal with this heterogeneity is to divide chronic pain patients into subgroups. Subgroups can be created using the specific characteristics certain patients have in common[1,2]. It has been argued that the identification of homogeneous subgroups makes it possible to tailor treatment programs to the specific needs of the subgroups. Doing this would likely lead to better treatment effects.

Many individual differences and demographic variables of chronic pain patients have been examined to determine who benefits most from treatment. The predictive value of variables such as age, pain history, number of operations, and psychological distress has been investigated[3,4]. Unfortunately, no consistent evidence has been found that supports the contention that these variables have either a positive or negative influence on treatment outcome[5]. Several attempts have been made to identify subgroups on the basis of psychological characteristics, psychopathology and cognitive factors. This has been done with the Minnesota Multiphasic Personality Inventory (MMPI)[,7,8,9,10,11], the Symptom Checklist 90 (SCL-90)[12,13,14,15,16] and the Automatic Thoughts Questionnaire[17,18]. To date, the results of studies using these questionnaires to identify patient subgroups are both contradictory and inconsistent[1,19]. Additionally, adaptation of treatment programs based on the identification of subgroups (using the MMPI, SCL-90 or the Automatic Thoughts Questionnaire) has not resulted in better treatment outcomes.

However, the identification of subgroups of patients on the basis of the West Haven-Yale Multidimensional Pain Inventory (MPI)[20] has proven to be much more successful. An advantage of the MPI is that pain-related psychosocial and behavioral factors are included in the classification. The MPI is advantageous in that it, unlike the aforementioned questionnaires, focuses on multiple pain-related factors. The MMPI, SCL-90 and the Automatic Thoughts Questionnaire focus only on one single pain-related factor. Evidently, chronic pain is a multidimensional problem and thus requires a multidimensional classification system. The MPI

offers such a system. Several studies have demonstrated the usefulness of the MPI in identifying differential responses to the same standardized treatment, for example in patients with temporomandibular disorders[21,22,23], fibromyalgia[2,24], and other chronic pain conditions[25,26,27,28]. The results of these studies have provided significant support for the hypothesis that customizing treatment and outcome measures using patient characteristics, psychosocial needs and somatic needs, may indeed improve treatment efficacy as well as the evaluation of treatment outcomes.

This chapter describes part of the findings derived from a randomized clinical trial (RCT) of a multidisciplinary cognitive-behavioral program for chronic pain treatment (CBT-R program) at the Roessingh Center for Rehabilitation (RCR) in the Netherlands. The initial findings of the RCT indicated that the program did not result in meaningful overall treatment effects (Chapter 2). However, closer inspection of the data revealed significant variability in the treatment outcome measures, a finding which poses some serious questions. As a result, the aim of the present chapter is to investigate the degree to which this variability can be explained by the differential effects of patient subgroups.

3.2 Methods

In this section, the methods applied to investigate the role of patient subgroups in explaining the variability found in the RCT on the CBT-R program are presented. A more detailed description of the methods applied for the RCT study in general can be found in Chapter 2.

3.2.1 Design

From April 1999 to July 2001, 391 chronic pain patients were referred to the CBT-R program. Of these 391 patients, 297 participated in the RCT. After giving written informed consent, the 297 subjects were randomly assigned to one of two groups: an Intervention Group (IG) and a waiting list Control Group (CG). The IG received immediate intake and treatment as soon as possible after informed consent (approximately one month). The CG was assigned to a waiting period of approximately six months followed by the same intake and treatment procedures as the IG. Randomization was performed in blocks of two. The study was designed and executed with permission of the Roessingh Medical Ethical Committee.

Patients received questionnaires by mail at home on numerous occasions, both prior to and following treatment. The occasions upon which these questionnaires were sent can be found in Figure 3.1.

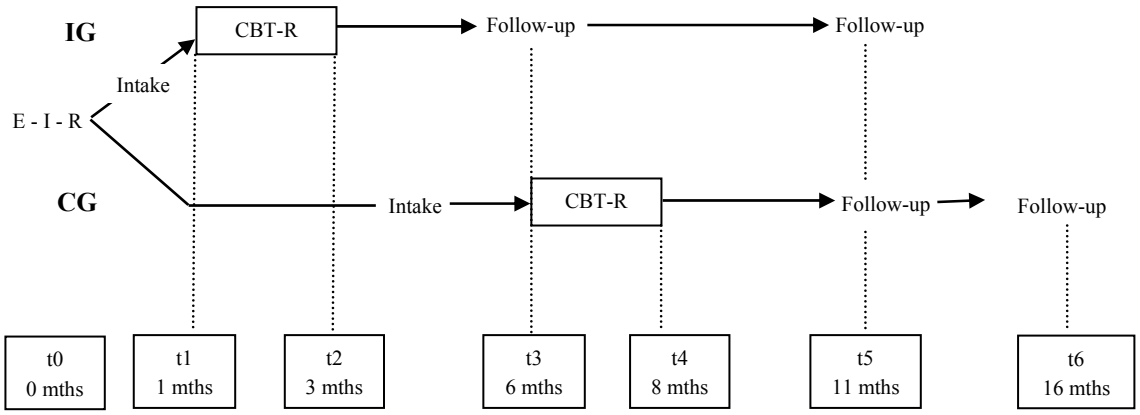


Figure 3.1: Research design (IG = intervention group; CG = control group; E=entrance, I= informed consent, R = randomization, t=time of measurement)

3.2.2 Outcome measures

Multiple outcome measures assessed by the questionnaires were used to evaluate the effect of treatment. Psychosocial outcome measures included:

1. All scales of the Multidimensional Pain Inventory-Dutch Language Version (MPI-DLV). These scales were used to measure cognitive-behavioral and psychological reactions to pain. The MPI-DLV scales were also used to identify subgroups of patients[31,32].
2. The visual analogue scale (VAS) of the McGill Pain Questionnaire-Dutch Language Version (MPQ-DLV). This scale was used to measure the intensity of experienced pain[34].
3. The total score (psychoneuroticism) of the Dutch version of the Symptom Checklist (SCL-90). This score was used as an indication of the presence of psychological difficulties[35,36].

Behavioral outcome measures included seven items of the RAND 36-item Health Survey (RAND-36). The RAND-36 is designed to evaluate patients' experienced disability and health status[37].

With the MPI-DLV, four pain patients profiles can be identified, namely a dysfunctional type (DYS), an interpersonally distressed type (ID), an adaptive coping type (AC) and an average type (AV)[32,38]. The DYS type tends to report high levels of pain severity, greater pain-related interference in daily life, high levels of affective distress, low levels of life control, low activity levels and a highly supportive environment (as reflected by a high level of solicitous responses). Compared to the other profiles, the DYS type tends to use analgesic medication more than the other types. The ID type is characterized by a high degree of pain severity, high levels of affective distress and low levels of life control. Further, a low level of environmental support (as reflected by a high level of punishing responses) characterizes this type. The AC type tends to report less pain severity, lower levels of affective distress, higher levels of activity and less pain-related interference in daily life than the DYS and ID types. The environmental support for the AC type is somewhat lower than of the DYS type but considerably higher than that of the ID type. The AC type's use of analgesic medication tends also to be lower than others. The AV type shares characteristics with the DYS, ID and AC profiles. In general, the AV type experiences less suffering compared to the DYS and ID type, namely less pain severity and interference, less negative distress, more pain control and a higher activity level[31]. Also, the AV type is characterized by a relatively high intelligence quotient[32,33]. In addition to the abovementioned profiles, an additional category exists for patients who cannot be classified into one of the mentioned profile types. These patients are classified as Anomalous. Patients with missing values on the MPI-DLV scale for Support and on the scales that deal with responses of significant others to the patient could not be assigned to a subgroup.

3.2.3 Data analysis

As described in Chapter 2, missing values were imputed by means of linear approximation and/or the Last Observation Carried Forward (LOCF) method. All data were analyzed using SPSS 11.5[39].

Multivariate analyses (MAN(C)OVA) were applied using a) posttest or follow-up scores as dependent (within-subjects) variables; and b) the MPI-DLV cluster type at t0 (three levels) as well as treatment mode (semi-inpatient group (SG) or outpatient group (OG)) as between-subjects factors. Based on recommendations by Stevens as well as Tabachnik and Fidell, pretest scores were added as covariates to correct for initial differences between the IG and the CG[40,41]. To determine the contribution of separate variables to the multivariate effects,

univariate analyses (AN(C)OVA) were performed. Along the lines described by Hair et al., the assumptions of the test were evaluated[42]. Further, the results were checked for the presence and effect of possible outlying cases. Since the distribution of the MPI-DLV cluster types on t0 did not differ between the IG and the CG, the variable MPI-DLV cluster at t0 was applied in the analyses to determine differential cluster effects.

Since the number of patients in each MPI-DLV cluster was too small to analyze the IG and the CG separately, the groups were combined. A short-term treatment effect was defined as the difference between pretest and post test scores (IG t1-t2; CG t3-t4) and a long-term effect was seen as the difference between pretest scores and follow-up scores (IG t1-t3; CG t3-t5). Although the assumption was that the treatment group would have no effect on potential subgroup differences, the variable treatment group was added as a between-subjects factor. The same was done with treatment mode (OG versus SG). Patients of both the SG and the OG were included in the analyses and treatment mode was used as between-subjects factor.

In this investigation of the role of subgroups in explaining variability, a two-sided *P* value equal to or lower than .05 was considered a reflection of statistical significance. Because of the explorative nature of the analyses, alpha's between .05 and .15 were interpreted in terms of trends. At first, results were interpreted on the basis of their statistical significance. Since statistical significance does not in itself provide concise information about a given intervention's clinically meaningful effects, results were also judged on the basis of their clinical relevance[43,44]. To date, there is little consensus on standards for determining clinical relevance. Therefore, clinical relevance was determined in two ways. First, a distinction was made on the basis of the RAND-36 scale for Health Change between patients who claim to have experienced an improvement in functioning (score > 50), patients who claim to have experienced a deterioration in functioning (score < 50) and patients who claim to have experienced minimal or no change after treatment (score = 50). Chi-square tests were used to test the differences between clusters. The second criterion for the determination of clinical relevance was a difference score between pretest and posttest/follow-up scores of at least ½ SD from the mean pretest score of the research population[45].

3.3 Results

3.3.1 Population

In Table 3.1, the general characteristics of the research population based on the MPI-DLV profile at t0 are presented. In this chapter, the group of AC patients at t0 could not be analyzed separately as the size of this group was too small (N = 8). Additionally, patients classified at t0 as anomalous (N = 6) and patients for whom no classification could be made due to a high degree of missing data at t0 (N = 18) were excluded from the analyses. Significant differences with regard to the general characteristics were found neither between IG and CG, nor between clusters (all *P*'s > .330).

Table 3.1: Population characteristics by MPI-DLV profile on t0

Variable		**DYS (N = 47)**	**ID (N = 28)**	**AV (N = 39)**
Age (years)	M(SD)	42.0 (10.3)	41.9 (8.5)	44.7 (11.1)
	Range	19-65	23-57	21-71
Sex	Male	10 (21.3%)	6 (21.4%)	10 (25.6%)
	Female	37 (78.7%)	22 (78.6%)	29 (74.4%)
Partner	No partner	2 (4.3%)	3 (10.7%)	0 (0.0%)
	Partner	45 (95.7%)	25 (89.3%)	39 (100.0%)
Education	Primary	8 (17.0%)	6 (21.4%)	8 (20.5%)
	Secondary	36 (76.6%)	20 (71.5%)	25 (64.1%)
	Higher	3 (6.4%)	2 (7.1%)	6 (15.4%)
Pain duration (years)	M(SD)	9.0 (9.6)	9.7 (8.2)	6.9 (6.4)
	Range	0.5-40	0.5-32	0.5-28
Treatment mode	SG	36 (76.6%)	24 (85.7%)	29 (74.4%)
	OG	11 (23.4%)	4 (14.3%)	10 (25.6%)
Drop-out rate	No drop-out	42 (89.4%)	25 (89.3%)	34 (87.2%)
	Drop-out	5 (10.6%)	3 (10.7%)	5 (12.8%)
Randomization group	IG	27 (57.4%)	20 (71.4%)	22 (56.4%)
	CG	20 (52.6%)	8 (28.6%)	17 (43.6%)

3.3.2 Changes with regard to MPI-DLV subgroups as a result of treatment

A shift in MPI-DLV clusters was expected following treatment. The most likely shift to occur was a shift from the DYS or ID clusters towards the AV or AC clusters. In Figure 3.2 (IG) and 3.3 (CG), changes in MPI-DLV classifications between t1 and t2 are presented. These figures show that, for the DYS group, 42.9% of the IG patients made a shift towards an AV or AC cluster, as compared to 18.8% patients of the CG. For the ID group, 34.7% of the IG patients made a shift towards an AV or AC cluster. In the CG, all ID patients remained in the same

cluster at t2. These results demonstrate that more positive results were obtained for the IG than for the CG. Some of the AV patients in the IG and CG shifted towards DYS and ID patients, which may be interpreted as a decline in functioning. However, because the n of the clusters was small, results must be interpreted carefully.

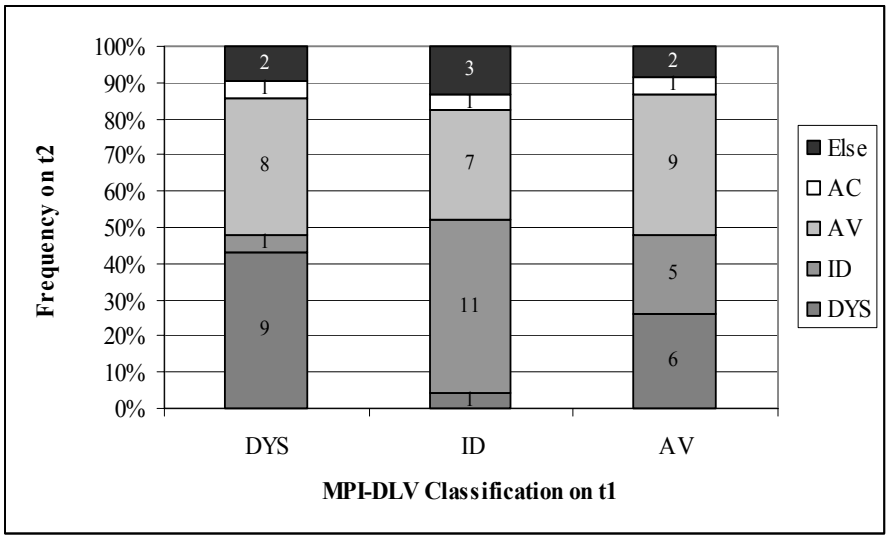


Figure 3.2: Changes in MPI-DLV classification of the IG from t1 to t2

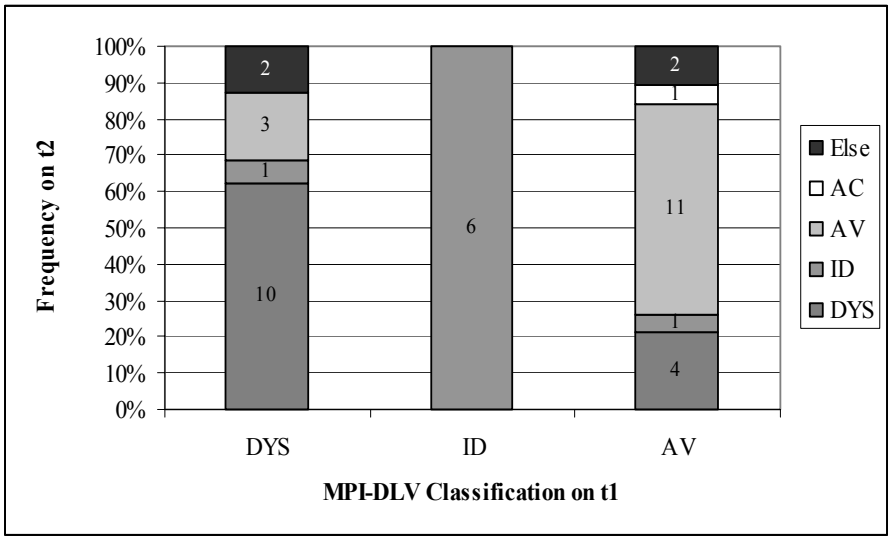


Figure 3.3: Changes in MPI-DLV classification of the CG from t1 to t2

3.3.3 Differential treatment effect

In Figure 3.4, the proportional mean difference scores between pretest and posttest scores of the IG and CG together are presented by MPI-DLV cluster at t0 for the variables of the MPI-DLV, SCL-90, MPQ-DLV and RAND-36. The proportional mean difference score was calculated by dividing the mean difference between pretest and posttest score by the pretest score, multiplied by 100%. Figure 3.4 indicates that there are differential cluster effects. The interpretation of this result requires further explanation. A negative proportional mean difference for the MPI-DLV variables Life Control, Distracting Responses and General Activity, and all RAND-36 scales indicate an improvement in functioning. In the case of the MPI-DLV Support and Solicitous Responses scales, the results must be interpreted with caution. For DYS patients, a decrease (as indicated by a proportional mean difference score above zero) on the Support and Solicitous Responses scales may be considered positive, whereas for an ID type this would be considered a negative result. For all other variables, a positive proportional mean difference can be interpreted as a sign of improvement. On most variables, the DYS patients show the largest improvement and thus appear to have benefited most from the CBT-R program. For DYS patients, the differential cluster effects for the MPI-DLV scales Affective Distress (χ^2 = 8.517; P = .014), Solicitous Responses (χ^2 = 8.749; P =

.013), and the RAND-36 variable Pain (χ^2 = 9.158; *P* = .010) are significant. A trend towards a significant cluster effect was also found for the MPI-DLV scale Interference (χ^2 = 4.311; *P* = .116) and the MPQ-DLV VAS (χ^2 = 4.126; *P* = .127).

Additionally, the proportional mean difference scores between pretest and follow-up scores for the IG and CG together indicate a trend towards a significant cluster effect for the MPI-DLV variable Solicitous Responses (χ^2 = 5.008; *P* = .082) and the RAND-36 scale Mental Health (χ^2 = 4.334; *P* = .114).

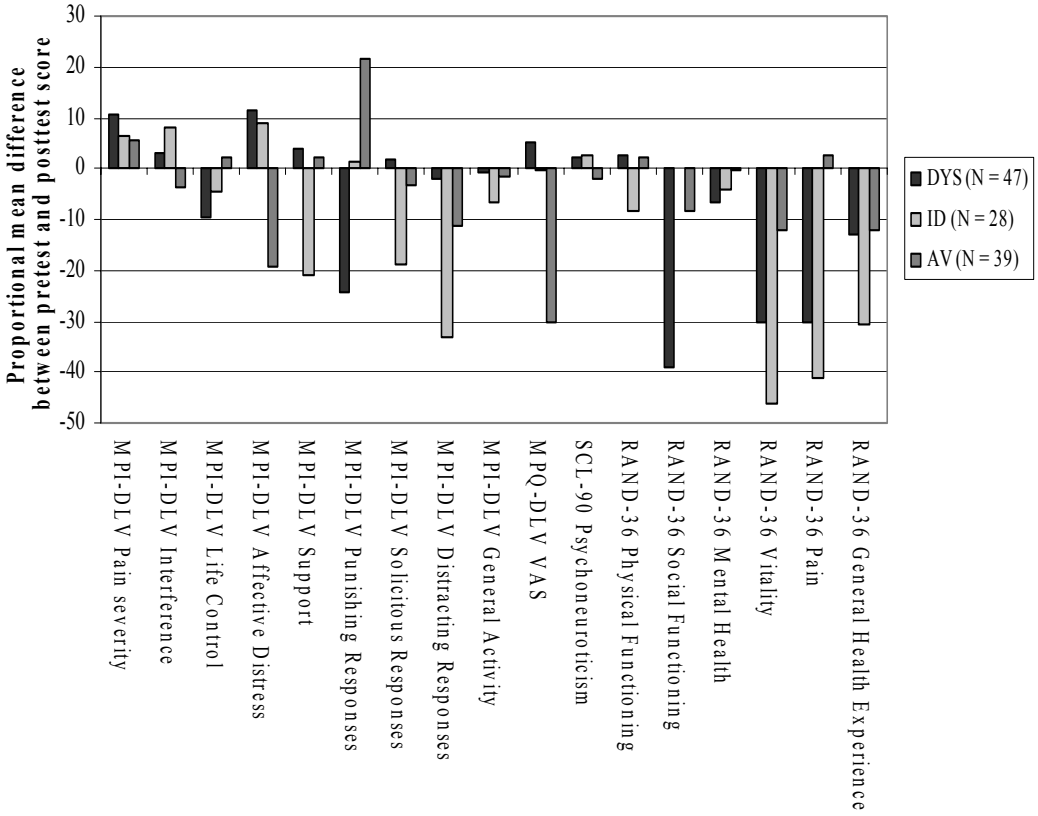


Figure 3.4: Proportional mean difference scores between pretest and posttest scores per MPI-DLV cluster on t0

A multivariate analysis with the posttest scores of all the MPI-DLV variables as dependent variables and all the MPI-DLV pretest scores as covariates demonstrated a significant between-subjects cluster x treatment mode interaction effect ($F_{(2, 98)}$ =

4.980; *P* = .009).This means that treatment mode may impact the kind of differential treatment effect found between the different MPI-DLV clusters. However, multivariate analyses that included the treatment group (IG versus CG) as a between-subjects factor resulted in a non-significant interaction effect between cluster and treatment group (*P* = .996). Further, the multivariate tests demonstrated a significant cluster x MPI-DLV scale x treatment mode interaction effect ($F_{(16, 98)}$ = 1.977; *P* = .039). This would imply that the cluster x treatment mode interaction effect was dependent of the MPI-DLV scale used.

Multivariate analysis on the follow-up scores of all MPI-DLV variables demonstrated a stronger significant between-subjects cluster x treatment mode interaction effect ($F_{(2, 99)}$ = 6.287; *P* = .003). However, in doing so, the significance of the cluster x MPI-DLV scale x treatment mode interaction effect disappeared (*P* = .205).

In an effort to determine the specific MPI-DLV variables that significantly contribute to the multivariate effects described above, univariate analysis were performed. In Table 3.2, the between-subjects effects of the MPI-DLV cluster as well as the MPI-DLV cluster x treatment mode interaction effects are presented. The results indicate that, directly after treatment, cluster effects from the MPI-DLV variables Interference and Solicitous Responses, as well as the MPI-DLV cluster x treatment mode interaction effects of the MPI-DLV scales Pain Severity and Punishing Responses, significantly contributed to the multivariate treatment effect. At follow-up, the multivariate effects were mostly derived from cluster effects in the MPI-DLV variables Solicitous Responses and Interference.

An additional multivariate analysis was performed with the scales of the other questionnaires as dependent variables. This analysis revealed that, directly after treatment, a trend could be distinguished with respect to a cluster x scale x treatment mode interaction effect ($F_{(14, 665)}$ = 2.079; *P* = .059). The RAND-36 Social Functioning scale contributed the most on a univariate level with a cluster x treatment mode interaction effect ($F_{(2, 113)}$ = 3.391; *P* = .037). However, at follow-up, no significant effects were found on a multivariate or univariate level.

Table 3.2: Results of univariate analysis over MPI-DLV variables (N = 114)

MPI-DLV variable	Directly after treatment		At three months follow-up	
	Between-subjects effect MPI-DLV cluster (F, *P*)	Between-subjects interaction effect MPI-DLV cluster x treatment mode (F, *P*)	Between-subjects effect MPI-DLV cluster (F, *P*)	Between-subjects interaction effect MPI-DLV cluster x treatment mode (F, *P*)
Pain Severity	.974 (.381)	5.826 (.004)	.787 (.458)	1.723 (.183)
Interference	4.291 (.016)	2.093 (.128)	2.936 (.057)	2.366 (.099)
Life Control	.956 (.388)	.598 (.552)	.232 (.793)	.727 (.486)
Affective Distress	2.346 (.101)	1.477 (.233)	2.218 (.114)	3.332 (.039)
Support	.734 (.482)	.313 (.732)	1.826 (.166)	2.723 (.070)
Punishing Responses	.468 (.628)	6.104 (.003)	.074 (.928)	1.849 (.162)
Solicitous Responses	3.170 (.046)	.648 (.525)	4.187 (.018)	.044 (.957)
Distracting Responses	.295 (.745)	.280 (.756)	.455 (.636)	1.550 (.217)
General Activity	.458 (.634)	.705 (.497)	.227 (.797)	.073 (.929)

In order to interpret the results of the preceding analyses, mean difference scores between pretest and posttest/follow-up were examined (Appendix 3A and 3B). The following conclusions could be drawn:

- ID patients from the OG seemed to improve the most on the MPI-DLV Punishing and Solicitous Responses scales, while ID patients from the SG improved most on the MPI-DLV variable Distracting Responses. These effects were found immediately after treatment and were maintained at follow-up.
- ID patients in the OG appeared to improve more than others on scales related to the experience and impact of pain, such as the MPI-DLV Pain Severity and Interference scales, the MPQ-DLV VAS and the RAND-36 scale for Pain.
- Both DYS and ID patients showed improvement with respect to psychosocial aspects of pain, as reflected by the MPI-DLV Life Control and Affective Distress scales. Again, the ID patients from the OG seemed to improve most on these scales. The results on the SCL-90 Psychoneuroticism scale and the RAND-36 scales for Social Functioning, Mental Health and Vitality (partially) supported these findings. ID patients also scored best on the RAND-36 variable General Health Experience.
- DYS patients seemed to benefit the most from the SG treatment, especially on the MPI-DLV variables Pain Severity, Affective Distress, Solicitous Responses and the RAND-36 variable Social Functioning. This trend was not only found immediately after treatment but also at follow-up.
- Patients in the OG showed more progress than the SG on the MPI-DLV variable General Activity. However, no clear distinction between clusters could me made.

3.3.4 Clinical relevance

In Table 3.3, the distribution of patients across the MPI-DLV clusters at t0 is presented. This distribution is combined with an analysis of the extent to which patients experienced improvement, as determined by the RAND-36 variable Health Change (score > 50) directly after treatment and at three months follow-up. For all the MPI-DLV clusters, both directly after treatment and at follow-up, approximately half of the patients experienced deterioration in functioning. 15-35 % of the patients experienced no change in health status and approximately ¼ of the patients appeared to have improved health status. The differences between the MPI-DLV clusters were not significant ($P > .546$).

Table 3.3: Experienced improvement directly after treatment and at three months follow-up

MPI-DLV cluster on t0	Directly after treatment			At three months follow-up		
	worsened	no change	improved	worsened	no change	improved
DYS (n=47)	28 (59.6%)	7 (14.9%)	12 (25.5%)	25 (53.2%)	14 (29.8%)	8 (17.0%)
ID (n=28)	13 (46.4%)	6 (21.4%)	9 (32.1%)	14 (50.0%)	6 (21.4%)	8 (28.6%)
AV (n=39)	19 (48.7%)	11 (28.2%)	9 (23.1%)	17 (43.6%)	14 (35.9%)	8 (20.5%)

Tables 3.4 and 3.5 present the number of clinically relevant treatment effects per MPI-DLV cluster at t0 (N = 114) directly after treatment and at three months follow-up. A ½ SD from the mean pretest scores was used as the criterion for a clinically relevant treatment effect. In the tables, an increase of MPI-DLV Pain Severity, Interference, Affective Distress, and Punishing Responses, SCL-90 Psychoneuroticism and MPQ-DLV VAS should be interpreted as a worsening. An increase of MPI-DLV Life Control, Distracting Responses and General Activity, and all RAND-36 scales should be interpreted as an improvement. An increase on MPI-DLV Support should be interpreted as an improvement for the ID and AV patients; for the DYS patients this should be interpreted as a worsening. Finally, an increase on MPI-DLV Solicitous Responses should be interpreted as an improvement for the ID patients and as a worsening for the DYS patients. Whether this means an improvement or worsening for the AV patients depends on the specific score on this scale.

The results in Table 3.4 and 3.5 indicate that, on most scales, the percentage of patients showing either no clinically relevant improvement or deterioration in functioning is larger than the percentage of patients who actually improved. However, for the MPI-DLV variables Affective Distress and Punishing Responses, the percentage of improved ID patients was larger than the percentage of patients who either showed no clinically relevant improvement or deteriorated in functioning. Comparable results were found for the percentage of DYS and AV patients showing an improvement on the RAND-36 variable Social Functioning directly after treatment, as well as for the percentage of DYS patients showing an improvement on the MPI-DLV variable Life Control at follow-up. Additionally, ID patients showed the largest percentage of clinically relevant improvements on all variables directly after treatment. At follow-up, the results were less pronounced. However, ID patients still showed the largest percentage of clinically relevant improvements on most of the variables.

Table 3.4: Number (%) of clinically relevant changes directly after treatment

Instrument		Cluster	Directly after treatment		
			decrease	no change	increase
MPI-DLV	Pain severity	DYS	15 (31.9%)	29 (61.7%)	3 (6.4%)
		ID	9 (32.1%)	17 (60.7%)	2 (7.1%)
		AV	11 (28.2%)	22 (56.4%)	6 (15.4%)
	Interference	DYS	14 (29.8%)	27 (57.4%)	6 (12.8%)
		ID	10 (35.7%)	15 (53.6%)	3 (10.7%)
		AV	7 (17.9%)	24 (61.5%)	8 (20.5%)
	Life Control	DYS	11 (23.4%)	24 (51.1%)	12 (25.5%)
		ID	6 (21.4%)	12 (42.9%)	10 (35.7%)
		AV	13 (33.3%)	19 (48.7%)	7 (17.9%)
	Affective Distress	DYS	20 (42.6%)	20 (42.6%)	7 (14.9%)
		ID	13 (46.4%)	10 (35.7%)	5 (17.9%)
		AV	10 (25.6%)	10 (25.6%)	19 (48.7%)
	Support	DYS	14 (29.8%)	29 (61.7%)	4 (8.5%)
		ID	5 (17.9%)	18 (64.3%)	5 (17.9%)
		AV	6 (15.4%)	25 (64.1%)	8 (20.5%)
	Punishing Responses	DYS	9 (19.1%)	27 (57.4%)	11 (23.4%)
		ID	12 (42.9%)	7 (25.0%)	9 (32.1%)
		AV	16 (41.0%)	19 (48.7%)	4 (10.3%)
	Solicitous Responses	DYS	10 (21.3%)	30 (63.8%)	7 (14.9%)
		ID	3 (10.7%)	16 (57.1%)	9 (32.1%)
		AV	8 (20.5%)	28 (71.8%)	3 (7.7%)
	Distracting Responses	DYS	6 (12.8%)	34 (72.3%)	7 (14.9%)
		ID	3 (10.7%)	18 (64.3%)	7 (25.0%)
		AV	5 (12.8%)	25 (64.1%)	9 (23.1%)
	General Activity	DYS	7 (14.9%)	33 (70.2%)	7 (14.9%)
		ID	6 (21.4%)	18 (64.3%)	4 (14.3%)
		AV	7 (18.4%)	22 (57.9%)	9 (23.7%)
SCL-90	Psychoneuroticism	DYS	8 (18.2%)	32 (72.7%)	4 (9.1%)
		ID	7 (25.0%)	19 (67.9%)	2 (7.1%)
		AV	4 (10.5%)	29 (76.3%)	5 (13.2%)
MPQ-DLV	VAS	DYS	17 (37.0%)	23 (50.0%)	6 (13.0%)
		ID	10 (35.7%)	12 (42.9%)	6 (21.4%)
		AV	9 (23.1%)	22 (56.4%)	8 (20.5%)
RAND-36	Physical Functioning	DYS	12 (25.5%)	28 (59.6%)	7 (14.9%)
		ID	4 (14.3%)	16 (57.1%)	8 (28.6%)
		AV	6 (15.4%)	27 (69.2%)	6 (15.4%)
	Social Functioning	DYS	12 (25.5%)	13 (27.7%)	22 (46.8%)
		ID	9 (32.1%)	12 (42.9%)	7 (25.0%)
		AV	12 (30.8%)	12 (30.8%)	15 (38.5%)
	Mental Health	DYS	13 (27.7%)	23 (48.9%)	11 (23.4%)
		ID	8 (28.6%)	12 (42.9%)	8 (28.6%)
		AV	12 (31.6%)	15 (39.5%)	11 (28.9%)
	Vitality	DYS	6 (12.8%)	25 (53.2%)	16 (34.0%)
		ID	4 (14.3%)	12 (42.9%)	12 (42.9%)
		AV	8 (21.1%)	18 (47.4%)	12 (31.6%)
	Pain	DYS	10 (21.3%)	19 (40.4%)	18 (38.3%)
		ID	2 (7.1%)	13 (46.4%)	13 (46.4%)
		AV	12 (30.8%)	17 (43.6%)	10 (25.6%)
	General Health Experience	DYS	14 (29.8%)	18 (38.3%)	15 (31.9%)
		ID	2 (7.1%)	18 (64.3%)	8 (28.6%)
		AV	6 (15.4%)	19 (48.7%)	14 (35.9%)

Table 3.5: Number (%) of clinically relevant changes at three months follow-up

Instrument		Cluster	Three months follow-up		
			decrease	no change	increase
MPI-DLV	Pain severity	DYS	10 (21.3%)	34 (72.3%)	3 (6.4%)
		ID	8 (28.6%)	15 (53.6%)	5 (17.9%)
		AV	9 (23.1%)	23 (59.0%)	7 (17.9%)
	Interference	DYS	11 (23.4%)	29 (61.7%)	7 (14.9%)
		ID	9 (32.1%)	17 (60.7%)	2 (7.1%)
		AV	8 (20.5%)	23 (59.0%)	8 (20.5%)
	Life Control	DYS	12 (25.5%)	17 (36.2%)	18 (38.3%)
		ID	8 (28.6%)	10 (35.7%)	10 (35.7%)
		AV	11 (28.2%)	20 (51.3%)	8 (20.5%)
	Affective Distress	DYS	18 (38.3%)	18 (38.3%)	11 (23.4%)
		ID	12 (42.9%)	7 (25.0%)	9 (32.1%)
		AV	12 (30.8%)	16 (41.0%)	11 (28.2%)
	Support	DYS	10 (21.3%)	33 (70.2%)	4 (8.5%)
		ID	7 (25.0%)	16 (57.1%)	5 (17.9%)
		AV	13 (33.3%)	20 (51.3%)	6 (15.4%)
	Punishing Responses	DYS	8 (17.0%)	30 (63.8%)	9 (19.1%)
		ID	11 (39.3%)	9 (32.1%)	8 (28.6%)
		AV	12 (30.8%)	20 (51.3%)	7 (17.9%)
	Solicitous Responses	DYS	15 (31.9%)	23 (48.9%)	9 (19.1%)
		ID	3 (10.7%)	13 (46.4%)	12 (42.9%)
		AV	9 (23.1%)	24 (61.5%)	6 (15.4%)
	Distracting Responses	DYS	13 (27.7%)	29 (61.7%)	5 (10.6%)
		ID	3 (10.7%)	18 (64.3%)	7 (25.0%)
		AV	7 (17.9%)	25 (64.1%)	7 (17.9%)
	General Activity	DYS	7 (14.9%)	26 (55.3%)	14 (29.8%)
		ID	7 (25.0%)	16 (57.1%)	5 (17.9%)
		AV	8 (20.5%)	26 (66.7%)	5 (12.8%)
SCL-90	Psychoneuroticism	DYS	10 (22.7%)	31 (70.5%)	3 (6.8%)
		ID	7 (25.0%)	20 (71.4%)	1 (3.6%)
		AV	5 (13.2%)	31 (81.6%)	2 (5.3%)
MPQ-DLV	VAS	DYS	17 (37.0%)	22 (47.8%)	7 (15.2%)
		ID	9 (32.1%)	13 (46.4%)	6 (21.4%)
		AV	11 (28.2%)	18 (46.2%)	10 (25.6%)
RAND-36	Physical Functioning	DYS	15 (31.9%)	23 (48.9%)	9 (19.1%)
		ID	10 (35.7%)	11 (39.3%)	7 (25.0%)
		AV	7 (17.9%)	22 (56.4%)	10 (25.6%)
	Social Functioning	DYS	14 (29.8%)	16 (34.0%)	17 (36.2%)
		ID	6 (21.4%)	11 (39.3%)	11 (39.3%)
		AV	11 (28.2%)	15 (38.5%)	13 (33.3%)
	Mental Health	DYS	15 (31.9%)	18 (38.3%)	14 (29.8%)
		ID	4 (14.3%)	13 (46.4%)	11 (39.3%)
		AV	12 (31.6%)	15 (39.5%)	11 (28.9%)
	Vitality	DYS	8 (17.0%)	20 (42.6%)	19 (40.4%)
		ID	5 (17.0%)	17 (60.7%)	6 (21.4%)
		AV	8 (21.1%)	22 (56.4%)	8 (21.2%)
	Pain	DYS	12 (25.5%)	19 (40.4%)	16 (34.0%)
		ID	6 (21.4%)	12 (42.9%)	10 (35.7%)
		AV	9 (23.1%)	20 (51.3%)	10 (25.6%)
	General Health Experience	DYS	11 (23.4%)	25 (53.2%)	11 (23.4%)
		ID	4 (14.3%)	16 (57.1%)	8 (28.6%)
		AV	7 (17.9%)	20 (51.3%)	12 (30.8%)

When comparing pre-test with post-test scores, a significant difference between clusters with regard to the number of clinically relevant changes was found for the MPI-DLV variables Affective Distress ($\chi^2_{(4)} = 14.261$; $P = .007$) and Punishing Responses ($\chi^2_{(4)} = 12.335$; $P = .015$). Additionally, a trend towards significance was found for the MPI-DLV variable Solicitous Responses ($\chi^2_{(4)} = 7.754$; $P = .101$) and the RAND-36 variable General Health Experience ($\chi^2_{(4)} = 7.869$; $P = .097$). When comparing pre-test with follow-up scores, significant differences between clusters with regard to the number of clinically relevant changes were found for the MPI-DLV Solicitous Responses scale ($\chi^2_{(4)} = 10.142$; $P = .038$). A trend indicating difference between clusters was also found for MPI-DLV Punishing Responses ($\chi^2_{(4)} = 7.895$; $P = .095$).

3.4 Discussion

The aim of the present chapter was to determine whether the variability in general treatment outcome of the CBT-R program could be (partially) explained by differential treatment effects of MPI-DLV clusters. Based on the results of the multivariate analyses, as well as the mean difference scores and number of clinically relevant changes, it can be concluded that indeed a differential treatment response between clusters of patients exists. In Figures 3.2 and 3.3, it becomes evident that, with respect to MPI-DLV clustering, positive shifts are greater in the IG than in the CG. Given that patients were randomly assigned to either the IG or the CG, it can be assumed that the differential cluster responses found in the present study were actually the result of the CBT-R program. The results are in line with studies conducted by Turk et al. and Gatchel et al., wherein positive shifts in MPI-DLV clustering as a result of treatment have been demonstrated[2,46]. The fact that the present study demonstrated smaller effects than the studies conducted by Turk en al. and Gatchel et al. is fairly logical as the overall effect of the CBT-R program has already been shown to be reasonably small.

In addition to a general cluster effect, the results from the multivariate analyses demonstrate a general treatment mode effect. This coincides with the finding in Chapter 2 that, in general, patients from the OG obtained a better treatment effect than patients from the SG. It is thus apparent that, in this study, the differential cluster response depends on the mode of treatment patients received. Because the OG sample was relatively small, no definite conclusions can be drawn. However, the finding that treatment mode partially determines the differential

cluster response is important as it supports the differential treatment of patients in different treatment modes (OG versus SG). For a more detailed description of OG and SG treatment, please see Chapter 2.

ID patients from the OG improved most on scales related to the experience and impact of pain. This finding contradicts the findings found in a study by Turk et al. in which patients with Fibromyalgia Syndrome, classified as DYS, significantly improved on pain severity and interference of pain[2]. However, a study by Rudy et al. demonstrated significant pain reduction for both ID and DYS patients with temporomandibular disorders[22]. Differences in the sample size of the clusters are unlikely to be the cause of the differences in study results. The most likely explanation is that, in the current study, SG and OG patients were analyzed together. Support for this explanation can be found by analyzing the SG patients separately, which demonstrated that DYS patients in the SG improved most on the MPI-DLV variable Pain Severity.

Results from the present study demonstrated that, while ID patients appear to show greater improvement in the OG treatment, DYS patients seem to benefit more from the SG treatment. The difference between these two clusters becomes apparent when looking at the MPI-DLV response scales. The expectation was that, after treatment, ID patients would report improvement on the MPI-DVL variable Distracting Responses and a decreased score for the MPI-DLV variables Punishing Responses and Solicitous Responses. For DYS patients, the expectation was an improved score for the Distracting Responses scale and a poorer score on the Solicitous Responses scale. The results show that ID patients from the OG did improve the most on the MPI-DLV Punishing and Solicitous Responses scales, while ID patients from the SG improved most on the MPI-DLV variable Distracting Responses. DYS patients from the SG showed the greatest improvement on the Solicitous Responses scale and a small but noteworthy improvement on the Punishing Responses scale. These results are more or less in line with the expectations. The only exception is found in the Distracting Responses scale. The expectation was that scores for Distracting responses would increase in the DYS cluster. This was not the case. However, it is important to note that robust effects could not be expected as patients did not receive a cluster specific treatment. In order to obtain a larger differential treatment response, a cluster specific (tailor made) treatment is essential. This is particularly relevant with respect to the role of significant others. For DYS patients, the involvement of significant others should be such that solicitous responses decrease and stimulating,

distracting responses increase. For ID patients, significant others should be encouraged to show understanding instead of punishing pain behavior.

The results from Figure 3.4 as well as Table 3.4 and Appendix 3A demonstrate that the ID patients improved most on the MPI-DLV Life Control and Affective Distress scales. This is remarkable. One would expect that the AV patients would improve more on the Life Control scale and, as a logical consequence, the Affective Distress scale (as the more control one has usually results in less distress). AV types are, in general, better adjusted in terms of cognitive rationality than other types. The expectation is even stronger in the present study as the CBT-R program is specifically aimed at changing and learning pain-coping cognitions. However, the results from Appendix 3A demonstrate that AV patients actual worsened on both the Life Control and the Affective Distress scales. However, it is important to note that the AV patient group, as a whole, and particularly AV patients from the SG, did show improvement on the SCL-90 Psychoneuroticism scale and the RAND-36 variables Social Functioning and Mental Health. It is possible that the CBT-R program does not sufficiently succeed in connecting with the specific cognitions of AV patients. This may possibly explain the lack of positive change on the Life Control and Affective Distress scales. Assuming that the underlying theoretical, cognitive mechanisms of the AV patients are indeed correct, it is imperative that we investigate ways in which the CBT-R program can be adapted to create a better fit with the cognitions of AV patients.

DYS patients are characterized by high levels of pain severity, greater pain-related interference in daily life, and low activity levels. These are factors that are, in many ways, more ‘physical’ than the other variables investigated. The expectation was that DYS patients would improve most on the scales relating to these factors. Although DYS patients as a group did obtain positive changes with respect to the MPI-DVL variables Pain Severity and Interference, the MPQ-DLV VAS, and the RAND-36 variables Pain and Vitality, results for the MPI-DLV variable General Activity and the RAND-36 variable Physical Functioning showed no change and, in some cases, physical deterioration. It is possible that the MPI-DLV variable General Activity was not specific enough to measure physical activity and that, for this reason, negative results were obtained. The MPI-DLV General Activity scale thus appears to be more suitable for the assessment of social activity. Further, a large part of the CBT-R program focuses on increasing the physical load capacity of patients. From a realistic baseline, this load capacity is

increased by means of graded activity. Another possible reason why an improvement with regard to physical functioning could not be demonstrated may be that the established increase of physical load capacity was not enough to make DYS patients feel that they had actually improved. Perhaps longer term measurements would produce better results as patients would have more time and more opportunities to increase their load capacity and to gain a sense of improvement in this area.

It is possible that, with the passing of time, differential treatment effects between the MPI-DLV clusters may become more apparent. The CBT-R program aims to establish behavioral change but behavioral change can be an incredibly long process. An eight week treatment program may then be too short to obtain significant changes in behavior. As a result, a differential treatment response may very well be unobtainable if measured only directly after treatment. It is more plausible that changes would be evident three months following therapy. However, the present study has demonstrated that the results were less significant or even worse at follow-up compared to the results found directly after treatment. As the CBT-R program is not adapted to the specific characteristics of the different MPI-DLV clusters (meaning that there was no differential treatment for every cluster), it is unrealistic to expect a large differential treatment effect both directly after treatment and at follow-up. The fact that the differential cluster effects diminished at follow-up, particularly the effects on the MPI-DLV responses scales, may indicate that it is difficult to maintain newly acquired behaviors in the home environment, especially when other factors, such as the influence of significant others, is present.

It may be argued that the subgroup specific results should have been corrected for multiple testing. However, the number of respondents from which data could be analyzed was relatively small due to a high drop-out rate. Because the number of patients included was too small to conduct subgroup specific analyses, no correction for multiple testing could be done. One could also contend that the subgroup specific results should have been analyzed using multilevel regression techniques. Indeed, from a technical point of view, these advanced techniques are the preferred method of analyzing the present data. However, it is questionable whether these techniques would actually result in more significant effects in view of the size of the already demonstrated effects. Furthermore, given that the results

of the present study are marginal, it is preferable to judge the data on their clinical relevance.

Although differential treatment effects between MPI-DLV clusters were found, these differences were relatively small. A potential explanation for these marginal results was found during a process evaluation of the CBT-R program (see Chapter 4). The semi-inpatient treatment takes place in two groups, the so-called 'Voorhuis' group (VH) and the 'Deel' group (D). The differences between these groups ought to be limited to the location of treatment in the rehabilitation center and the therapists assigned to these locations. However, despite the fact that the CBT-R program was mainly conducted according to protocol (insofar as was possible), significant differences between the VH and D group with regard to treatment effect were found for the MPI-DLV variable Pain Severity ($t = 2.034$; $P = .046$). A trend towards a significant difference on the MPI-DLV variable Interference ($t = 1.939$; $P = .056$) was also found. Further, an opposite pattern in treatment effect was found between VH and D group on several variables. Unfortunately, due to the limited number of patients in the D group, a VH-D between-group variable could not be included in the present chapter's analyses. Nevertheless, it may be hypothesized that the results from VH and D group cancelled each other out and, in doing so, no large significant differential cluster effects could be demonstrated.

In conclusion, this investigation has produced some important clinically relevant results. The present study has demonstrated that a differential treatment response of MPI-DLV clusters may have been responsible for the variability in treatment outcome results of the CBT-R program. The fact that differential effects were found supports the idea that customizing the CBT-R program to this differential response might improve overall treatment efficacy. Therefore, it is important to take into account a patient's cluster type when determining which treatment mode is most appropriate Lastly, it could be argued that other variables relating to the treatment process or the underlying theory of the CBT-R program may influence (confound and/or modify) the expected cluster specific effects. This will be investigated in the chapters to come.

Acknowledgement

With special thanks to Jitske Gulmans for the collection and analysis of data during her internship.

References

1. Turk, D.C. (1990). Customizing treatment for chronic pain patients: who, what, and why? *Clin J Pain, 6,* 255-270.
2. Turk, D.C., Okifuji, A., Sinclair, J.D., & Starz, T.W. (1998). Differential responses by psychosocial subgroups of Fibromyalgia syndrome patients to an interdisciplinary treatment. *Arthritis Care Res, 11,* 397-404.
3. Maruta, T., Swanson, D.W., & Swenson, W.M. (1979). Chronic pain: which patients may a pain-management program help? *Pain, 7,* 321-329
4. Lee, P.W.H., Chow, S.P., Lieh-Mak, F., Chan, K.C., & Wong, S. (1989). Psychosocial factors influencing outcome in patients with low-back pain. *Spine, 14,* 838-843.
5. Turk, D.C., & Rudy, T.E. (1990). Neglected factors in chronic pain treatment outcome studies: referral patterns, failure to enter treatment, and attrition. *Pain, 43,* 7-25.
6. Sternbach, R. (1974). *Pain patients: traits and treatments*. New York: Academic Press.
7. Bradley, L.A., Prokop, C.K., Margolis, R., & Gentry, W.D. (1978). Multivariate analyses of the MMPI profiles of low back pain patients. *J Behav Med, 1,* 253-272.
8. Armentrout, D.P., Moore, J.E., Parker, J.C., Hewett, J.E., & Felz, C. (1982). Pain patient MMPI subgroups: the psychological dimensions of pain. *J Behav Med, 1,* 201-211.
9. Moore, J.E., Armentrout, D.P., Parker, J.C., & Kivlahan, D.R. (1986). Empirically derived pain-patients MMPI subgroups: prediction of treatment outcome. *J Behav Med, 9,* 51-63.
10. 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. *J Behav Med, 111,* 159-169.
11. Everitt, B. (1989). *Cluster analysis* (2nd ed.). New York: Halstead-Wiley.
12. Derogatis, L.R. (1977). *SCL-90 administration, scoring and procedures manual*. Baltimore, MD: Johns Hopkins University Press.
13. Schwartz, D.P., & DeGood, D.E. (1983). An approach to the psychosocial assessment of the chronic pain patient. *Curr Concepts Pain, 1,* 3-11.
14. Buckelew, S.P., DeGood, D.E., Schwartz, D.P., & Kerler, R.M. (1986). Cognitive and somatic item response pattern of pain patients, psychiatric patients, and hospital employees. *J Clin Psychol, 42,* 852-860.
15. Butterworth, J.C., & Deardorff, W.W. (1987). Psychometric profiles of craniomandibular pain patients: identifying specific subgroups. *J Craniomandib Pract, 5,* 225-232.
16. Jamison, R.N., Rock, D.L., & Parris, W.C.V. (1988). Empirically derived Symptom Checklist 90 subgroups of chronic pain patients: a cluster analysis. *J Behav Med, 11,* 147-158.
17. Hollon, S.D., & Kendall, P.C. (1980). Cognitive self-statements in depression: development of an Automatic Thoughts Questionnaire. *Cog Ther Res, 4,* 383-395.
18. Ingram, R.E., Atkinson, J.H., Slater, M.A., Saccuzzo, D.P., & Garfin, S.R. (1990). Negative and positive cognition in depressed and non-depressed chronic pain patients. *Health Psychol, 9,* 300-314.
19. Turk, D.C. (2005). The potential treatment matching for subgroups of patients with chronic pain: lumping versus splitting. *Clin J Pain, 21*, 44-55.

20. Kerns, R.D., Turk, D.C., & Rudy, T.E. (1985). The West Haven-Yale Multidimensional Pain Inventory (WHYMPI). *Pain, 23,* 345-356.
21. Rudy, T.E., Turk, D., Zaki, H., & Curtin, H. (1989). An emperical taxometric alternative to traditional classsification of temporomandibular disorders. *Pain, 36,* 311-320.
22. Rudy, T.E., Turk, D.C., Kubinski, J.A., & Zaki, H.S. (1995). Differential treatment responses of TMD patients as a function of psychological characteristics. *Pain, 61,* 103-112.
23. Dahlström, L., Widmark, G., & Carlsson, S.G. (1997). Cognitive-behavioral profiles among different categories of orofacial pain patients: diagnostic and treatment implications. *Eur J Oral Sci, 105,* 377-383.
24. Turk, D.C., Okifuji, A., Sinclair, J.D., & Starz, T.W. (1996). Pain, disability, and physical functioning in subgroups of patients with Fibromyalgia. *J Rheumatol, 23,* 1255-1262.
25. Hutten, M.M.R., Hermens, H.J., IJzerman, M.J., Lousberg, R., & Zilvold, G. (1999). Distribution of psychological aspects in subgroups of chronic low back pain patients divided on the score of physical performance. *Int J Rehabil Res, 22,* 261-268.
26. Bergström, G., Bodin, L., Jensen, I.B., Linton, S.J., & Nygren, A.L. (2001). Long-term, non-specific spinal pain: reliable and valid subgroups of patients. *Beh Res Ther, 39,* 75-87.
27. Hankin, H.A., Spencer, T., Kegerreis, S., Worrell, T., & Rice, J.M. (2001). Analysis of pain behavior profiles and functional disability in outpatient physical therapy clinics. *J Orthop Sports Phys Ther, 31,* 90-95.
28. Vollenbroek-Hutten, M.M., Hermens, H.J., Wever, D., Gorter, M., Rinket, J., & IJzerman, M.J. (2004). Differences in outcome of a multidisciplinary treatment between subgroups of chronic low back pain patient defined using two multiaxial assessment instruments: the multidimensional pain inventory and lumbar dynamometry. *Clin Rehabil, 18*, 566-579.
29. Revalidatiecentrum Het Roessingh (RCR) (1996). *De pijn de baas: revalidatieprogramma's voor mensen met chronische pijnklachten. Informatiebrochure*. Enschede: Het Roessingh.
30. Divisie Pijnrevalidatie Het Roessingh (1999). *Beslissingsondersteunend en verrichtingenprotocol van het Pijn de Baas programma*. Enschede: Het Roessingh.
31. Lousberg, R., & Groenman, N.H. (1994). *Multidimensionele Pijnvragenlijst - Dutch Language Version (MPI-DLV): Handleiding*. Maastricht: Uitgeverij Stichting Klinisch Pijnonderzoek.
32. Lousberg, R., Breukelen, G.J.P. van, Groenman, N., Schmidt, A.J.M., Arntz, A., & Winter, F.A.M. (1999). Psychometric properties of the Multidimensional Pain Inventory, Dutch language version (MPI-DLV). *Behav Res Ther, 37,* 167-182.
33. Lousberg, R., Groenman, N., & Schmidt, A. (1996). Profile characteristics of the MPI-DLV clusters of pain patients. *J Clin Psychol, 52*, 161-167.
34. Verkes, R.J., Vanderiet, K., Vertommen, H., Kloot, W.A. van der, & Meij, J. van der (1989). De MPQ-DLV: Een standaard Nederlandstalige versie van de McGill Pain Questionnaire voor België en Nederland. In: W.A. van der Kloot, & H. Vertommen (red.). *De MPQ-DLV: Een standaard nederlandstalige versie van de McGill Pain Questionnaire; achtergronden en handleiding.* Lisse: Swets & Zeitlinger, p. 57-69.
35. Derogatis, L.R. (1983). SCL-90R. *Manual II. Clinical Psychometric Research*. Maryland.
36. Arrindell, W.A., & Ettema, J.H.M. (1986). *SCL-90: Handleiding bij een multidimensionele psychopathologie-indicator*. Lisse: Swets & Zeitlinger.

37. Zee, K.I. van der, & Sanderman, R. (1993). *Het meten van de algemene gezondheidstoestand met de RAND-36: Een handleiding*. Groningen: Noordelijk Centrum voor Gezondheidsvraagstukken.
38. Turk, D.C., & Rudy, T.E. (1988). Toward an empirically derived taxonomy of chronic pain patients: integration of psychological assessment data. *J Consult Clin Psychol, 56,* 233-238.
39. SPSS (2002). *SPSS Base 11.5 User's guide*. USA: SPSS Inc.
40. Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3 ed.). New Jersey: Lawrence Erlbaum Associates.
41. Tabachnik, B.G., & Fidell, L.S. (1996). *Using multivariate statistics* (3 ed.). New York: Harper Collins College Publishers.
42. Hair, J.F., Anderson, R.E., Tatham, R.L., & Black, W.C. (1998). *Multivariate data analysis*. London: Prentice-Hall International.
43. Dyer, I. (1998). The significance of statistical significance. *Accid Emerg Nurs, 6,* 92-98.
44. Kendall, P.C. (1999). Clinical significance. *J Consult Clin Psychol, 67,* 283-284.
45. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.
46. Gatchel, R.J., Noe, C.E., Pulliam, C., Robbins, H., Deschner, M., Gajraj, N.M., & Vakharia, A.S. (2002). A preliminary study of Multidimensional Pain Inventory profile differences in predicting treatment outcome in a heterogeneous cohort of patients with chronic pain. *Clin J Pain*, *18*, 139-143.

Appendix 3A: Mean differences (MD(SD)) between pretest and posttest/follow-up scores on MPI-DLV *(DYS SG N=36, OG N=11; ID SG N=24, OG N=4; AV SG N=29, OG N=10)*

Instrument	MPI-DLV cluster on t0	MD(SD) pretest-posttest		MD(SD) pretest-follow-up	
		SG	OG	SG	OG
Pain Severity	DYS	0.62 (0.97)	0.09 (0.70)	0.30 (0.77)	0.23 (0.82)
	ID	0.22 (0.76)	1.13 (2.59)	0.06 (1.17)	0.88 (1.49)
	AV	0.00 (0.80)	0.99 (1.15)	0.22 (1.04)	-0.05 (1.59)
Interference	DYS	0.18 (0.61)	0.26 (0.62)	0.03 (0.66)	0.38 (0.81)
	ID	0.25 (0.56)	0.92 (1.09)	0.28 (0.59)	0.88 (1.06)
	AV	-0.05 (0.64)	-0.14 (0.66)	0.05 (0.61)	-0.10 (0.75)
Life Control	DYS	-0.11 (1.16)	-0.03 (1.15)	-0.13 (1.23)	0.12 (1.49)
	ID	-0.06 (1.01)	-0.17 (1.97)	-0.06 (1.39)	-0.50 (1.60)
	AV	0.17 (1.26)	0.67 (1.04)	-0.02 (1.03)	0.10 (0.88)
Affective Distress	DYS	0.48 (1.30)	0.42 (0.82)	0.24 (1.04)	0.21 (1.37)
	ID	0.26 (1.38)	1.34 (1.56)	0.10 (1.41)	1.59 (0.83)
	AV	-0.25 (1.06)	-0.20 (1.44)	-0.01 (0.97)	-0.40 (1.88)
Support	DYS	0.26 (0.47)	0.05 (0.69)	0.24 (0.59)	0.09 (0.52)
	ID	-0.06 (1.01)	0.17 (0.80)	-0.06 (1.46)	0.34 (0.98)
	AV	0.01 (0.65)	0.33 (1.63)	0.46 (0.88)	-0.13 (0.57)
Punishing Responses	DYS	-0.01 (1.04)	-0.23 (1.04)	-0.04 (0.96)	-0.30 (1.00)
	ID	0.11 (1.56)	2.09 (1.13)	0.15 (1.54)	0.92 (0.69)
	AV	0.65 (1.03	-0.73 (1.94)	0.49 (1.00)	-0.60 (1.01)
Solicitous Responses	DYS	0.42 (0.79)	-0.47 (0.89)	0.31 (0.95)	-0.61 (1.11)
	ID	-0.22 (0.64)	-0.54 (0.63)	-0.40 (1.09)	-0.71 (0.48)
	AV	0.21 (0.78)	0.00 (0.79)	0.31 (0.73)	-0.15 (0.66)
Distracting Responses	DYS	0.05 (0.86)	-0.12 (0.99)	0.28 (0.87)	-0.24 (1.34)
	ID	-0.60 (1.53)	0.33 (0.67)	-0.43 (1.26)	0.75 (0.57)
	AV	-0.18 (0.96)	0.01 (1.35)	0.00 (1.00)	0.07 (1.30)
General Activity	DYS	-0.01 (0.37)	0.06 (0.35)	-0.05 (0.57)	-0.11 (0.36)
	ID	0.07 (0.52)	-0.17 (0.76)	0.07 (0.72)	-0.04 (0.71)
	AV	0.08 (0.43)	-0.16 (0.47)	0.06 (0.45)	-0.01 (0.41)

Negative scores on Pain Severity, Interference, Affective Distress, and Punishing Responses should be interpreted as a worsening. Negative scores on Life Control, Distracting Responses, and General Activity should be interpreted as an improvement. A negative score on Support means an improvement for ID and AV patients, but a worsening for DYS patients. A negative score on Solicitous Responses means an improvement for ID patients, but a worsening for DYS patients. Whether this means an improvement or worsening for AV patients depends on the specific score on this scale.

Appendix 3B: Mean differences (MD(SD)) between pretest and posttest/follow-up scores on SCL-90, MPQ-DLV, RAND-36 ***(DYS SG N = 36, OG N = 11; ID SG N = 24, OG N = 4; AV SG N = 29, OG N = 10)***

Instrument		MPI-DLV cluster on t0	MD(SD) pretest-posttest		MD(SD) pretest-follow-up	
			SG	OG	SG	OG
SCL-90	Psycho-neuroticism	DYS	3.16 (28.27)	5.93 (7.88)	5.29 (22.04)	12.04 (11.40)
		ID	1.88 (27.49)	33.04 (30.43)	9.98 (31.39)	34.05 (25.55)
		AV	0.44 (20.27)	-6.20 (9.70)	3.02 (22.95)	0.84 (11.99)
MPQ-DLV	VAS	DYS	0.76 (1.83)	0.45 (2.25)	0.70 (1.91)	0.64 (2.04)
		ID	0.41 (1.46)	0.25 (4.10)	0.53 (2.21)	1.58 (3.76)
		AV	-0.21 (1.70)	0.26 (1.70)	0.28 (2.06)	-0.28 (2.76)
RAND-36	Physical Functioning	DYS	0.79 (12.97)	2.05 (12.69)	0.51 (13.91)	5.91 (14.29)
		ID	-1.04 (11.44)	-5.00 (16.33)	-0.42 (14.44)	-6.25 (26.26)
		AV	1.90 (11.60)	-1.25 (11.13)	-0.69 (10.58)	-1.50 (7.09)
	Social Functioning	DYS	-9.20 (23.41)	-0.57 (16.41)	-3.96 (24.09)	1.14 (19.73)
		ID	7.81 (20.46)	-18.75 (12.50)	-1.04 (19.48)	-18.75 (7.22)
		AV	0.00 (20.04)	-4.38 (27.33)	-2.16 (16.39)	1.25 (18.11)
	Mental Health	DYS	-2.89 (14.23)	0.36 (9.54)	-0.44 (13.55)	0.36 (11.38)
		ID	-0.08 (14.09)	-1.00 (13.22)	-5.50 (14.29)	-4.00 (9.80)
		AV	-2.07 (13.34)	13.78 (23.07)	-1.24 (12.00)	13.33 (27.06)
	Vitality	DYS	-7.43 (15.56)	0.00 (7.42)	-5.21 (13.39)	-3.18 (11.24)
		ID	-5.83 (15.98)	-12.50 (10.41)	-2.92 (14.44)	-11.25 (12.50)
		AV	-2.41 (12.15)	7.22 (23.06)	-2.24 (11.39)	3.89 (17.28)
	Pain	DYS	-4.93 (16.97)	2.23 (15.66)	-3.00 (18.40)	0.93 (10.97)
		ID	-5.83 (13.81)	-13.78 (10.19)	-2.47 (20.89)	-13.78 (10.19)
		AV	1.13 (13.84)	1.22 (14.37)	-0.14 (15.72)	0.82 (15.55)
	General Health Experience	DYS	-1.81 (12.77)	1.14 (17.15)	-0.56 (11.26)	4.09 (20.71)
		ID	-3.54 (11.58)	-5.00 (10.80)	-3.13 (15.09)	-2.50 (15.00)
		AV	-3.28 (9.84)	-4.50 (19.07)	-3.62 (11.01)	2.50 (15.14)

A negative score on SCL-90 Psychoneuroticism and MPQ-DLV VAS should be interpreted as a worsening.
A negative score on all RAND-36 scales should be interpreted as an improvement.

CHAPTER 4
Evaluation in pain rehabilitation: explaining treatment variability using process variables

4.1 Introduction

One of the most accepted methods for investigating treatment efficacy is to determine the outcome of a treatment program by means of a (double blind) randomized controlled trial (RCT). However, by focusing primarily on treatment outcomes, other factors that influence the treatment *process* may be overlooked. In doing so, it is possible that the very factors responsible for treatment success remain unrecognized (black box evaluation, Figure 4.1).

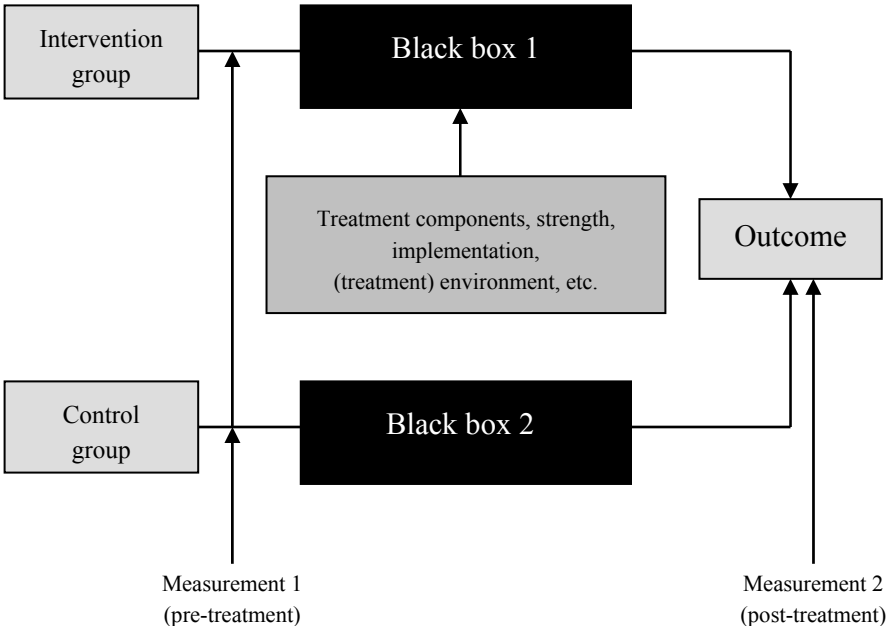


Figure 4.1: Black box evaluation

Although we are often able to establish that treatment results are positive, it is often times difficult to determine exactly which elements are responsible for the positive results. Likewise, when studies that focus on outcomes show no significant improvements or even negative results, it is impossible to determine whether these results were caused by an inadequate implementation of treatment, by the inclusion or exclusion of certain treatment components or by other factors affecting treatment outcome[1]. Evidently, ignoring the process of treatment makes it difficult

to generalize treatment results to other settings. Therefore, not only the outcome of a treatment program, but also the process of treatment, should be considered if we are to effectively evaluate treatment success. As a result, it is imperative that we include process variables in the evaluation of a program.

Several definitions of program evaluation have been proposed. An overview of definitions is presented by Dekker and Leeuw[2]. It appears that differences with respect to the definition of program evaluation are linked to the disciplinary background of the evaluators. These differences concern a) the realization of a priori identified goals[3]; b) criteria for success or failure[4]; c) the predominant focus of the evaluation (process or outcome)[5]; and d) the goal of the evaluation (many evaluations only focus on determining the effect of a program)[6,7,8]. Despite the differences, there is consensus with respect to the idea that program evaluation is a form of empirical scientific research that may serve to attend practical actions[9].

The present study utilizes a definition of program evaluation proposed by Chen[5]. According to Chen, program evaluation is defined as the systematic collection of empirical evidence for the purpose of: a) determining the congruence between the planned program and the program that is actually implemented; and b) verifying program-impact, underlying causal mechanisms, and the level of generalizability of the program. The theory-driven method proposed by Chen states "that every intervention program operates under a certain program theory, which, often implicit or asystematic, provides in general guidelines for the design of a program and explains how a program is supposed to work" (p. 39). Further, Chen defines program theory as "a specification of what must be done to achieve the desired goals, what other impacts may also be anticipated, and how these goals and impacts would be generated" (p. 43). Every program theory is comprised of both normative theory and causative theory. Normative theory has a prescriptive nature and concerns the structure of a program, guidelines for pursuing the outcome, and indications with respect to how the program should be implemented. Normative theory can be divided in three domains, each with its own theory:

- The *treatment domain* describes the nature of the treatment components (activities), as well as the treatment strength in terms of frequency, duration and intensity;
- The *implementation-environment domain* specifies the nature of the contextual situation for the treatment to be implemented; and
- The *outcome domain* specifies the nature of the treatment outcome or goals.

Causative theory has a descriptive nature. It describes the underlying mechanisms between program actions, implementation processes and program outcome. Causative theory is important for understanding program efficacy.

In Chapter 2, the general results of a randomized controlled trial (RCT) conducted to determine the effectiveness of a pain management program (the CBT-R program) provided at the Roessingh Rehabilitation Center (RCR) in the Netherlands were presented. The RCT results demonstrated that, although many patients showed an improvement after treatment, some patients did not improve. Some patients even deteriorated following treatment. This chapter endeavors to evaluate whether the treatment variability found in Chapter 2 is caused by factors in the treatment process. It is possible that differences between the prescribed CBT-R program and the CBT-R program that was actually implemented may explain this treatment variability and/or the negative outcome results. One could contend that when treatment goals are not attained, either the treatment actions were based on false theoretical assumptions (a theory failure) or the actions themselves were not conducted according to predetermined assumptions (a program failure). Assuming that the theoretical background of the CBT-R program is indeed correct, the treatment program was evaluated on the basis of the program's normative theory. As a result, the present evaluation investigates the prescriptions for treatment components, goals and implementation in the CBT-R program.

4.2 Methods

First and foremost, a critical inquiry of the CBT-R treatment protocol was performed. The protocol was drawn up in 1999 by all members of the multidisciplinary pain rehabilitation team of the CBT-R program[10]. Following this inquiry, the congruence between the prescribed specifications in the protocol and the treatment program that was actually implemented was determined. To do this, the following methods were used:

a. Weekly evaluation forms were filled in by CBT-R therapists. The evaluation forms were comprised of questions about therapy contents, the attendance of patients, the pursued treatment goals, whether patients met the selection criteria for the program, problems experienced during treatment and other factors that may have influenced treatment outcomes. An example of the evaluation forms can be found in Appendix 4A. For

practical and organizational reasons, four patient groups were included in the evaluation. One group was an outpatient group. The other three were semi-inpatient groups. The semi-inpatients groups were labeled as follows: a) Deel Monday (DM); b) Deel Wednesday (DW); and c) Voorhuis Wednesday (VW). In short, the therapists of each group filled in and submitted weekly evaluation forms during eight weeks of treatment in 2000.

b. Multidisciplinary patient consultations between therapists were observed. Twice a month, all semi-inpatient and outpatient group therapists got together for a plenary meeting. Additionally, the groups of therapists met separately twice a month. During these meetings, the individual patient progress (or lack thereof) was evaluated. Factors that impact a patient's lack of progress were discussed. Such factors included both behavioral factors, such as the patient's characteristics, behavior and attitude, and environmental factors. The observation of these meetings was done in 1999.

c. Following completion of the RCT, an evaluative interview with five (approximately one fifth) of the therapists involved in the CBT-R program was held. In this interview, personal observations and judgments of the treatment and group processes, as experienced by the therapists, were discussed.

d. Lastly, patients that participated in the RCT were asked to evaluate their CBT-R program experience. This was done using a Dutch translation of the Treatment Helpfulness Questionnaire (THQ)[11]. On a 10 point scale (-5 being very damaging and +5 being very helpful), patients were asked to indicate how they had experienced the various treatment components of the CBT-R program. The THQ is a reliable and valid measure for assessing patients' perceptions on how helpful treatment modalities offered at multidisciplinary pain centers are (Cronbach Alpha's ranging from 0.84 to 0.89; Pearson test-retest reliability $r = 0.92$)[11]. Patients completed the THQ at the end of their treatment period. For a description of the patients that participated in the RCT, please see Chapter 2.

4.3 Results

4.3.1 Critical inquiry of the CBT-R treatment protocol

The critical inquiry of the treatment protocol revealed that several unclear and incomplete statements were included in the protocol. Evidently, this lack of clarity can make it difficult to determine the congruence between the actual treatment provided and the treatment protocol. One difficulty concerned the protocol's lack of congruence with Chen's format and classification system of normative theory domains[5]. As a result, not all domains of the normative theory could be identified.

Other difficulties that arose were related to the content of the protocol. Firstly, the protocol states that, in order to participate in the CBT-R program, an indication for rehabilitation should be present. However, the protocol does not describe how this kind of indication can be identified or determined. Secondly, although selection criteria for participation in the CBT-R program are described, the protocol fails to operationalize these criteria. It also fails to indicate how selection criteria can be measured. For example, one criterion for inclusion is that the patient 'shows an obvious decrease in social functioning as a result of the pain'. Unfortunately, the protocol fails to indicate how this can be measured and to what degree social functioning has to be diminished to be included in the CBT-R program. Another example relates to exclusion criteria. The protocol states that patients cannot be included if they have extensive cognitive disturbances or if they cannot sufficiently be instructed. However, the protocol fails to operationalize what 'extensive' or 'sufficiently' is or to give an indication of how these criteria can be measured. Evidently, it is impossible to determine, from the protocol, exactly who can be included or excluded from the CBT-R program. One final difficulty with the protocol must be mentioned: Although the protocol does describe, for every treatment discipline, which activities should be performed to attain specific sub-goals, it remains unclear exactly which specific activity must be performed to attain a certain goal. It is further unclear how that activity would lead to a given goal (the underlying theory), how the sub-goals are operationalized, and how the attainment of goals can be measured. This lack of definition and operationalization of activities, goals and attainment of goals can be illustrated by the following: For the psychology component, the protocol does not define which sub-goals should be attained each week. Instead, the weekly goals are almost always identical and include such goals as increased psychosocial load capacity, enhanced insight with respect to one's psychological functioning, improvement in the restoration of

balance between relaxation, diversion and strain, enhanced assertiveness and communication skills, boosted self-image, and so on. Further, in the third week of treatment, patients are expected to learn about relaxation. In their treatment workbook, they are taught relaxation exercises, visualization exercises and a so-called 'bank' exercise. Unfortunately, the protocol fails to describe how a 'bank' exercise leads, for example, to the development of a positive self-image or to an increase of psychosocial load capacity. It also fails to indicate when a patient has reached a certain goal and how this can be determined.

The above mentioned problems with the protocol relate not only to the normative theory of the treatment domain, but also to the normative theory of the outcome domain. Difficulties were also experienced with respect to determining the normative theory of the implementation-environment domain. The normative theory of the implementation-environment domain should describe the contextual environment in which the treatment is implemented. This domain comprises of seven dimensions. These dimensions are: participant, implementor, delivery mode, implementing organization, interorganizational relationship, micro-context and macro-context. In the CBT-R treatment protocol, guidelines are provided for the participant dimension only and, on that dimension, are limited to inclusion and exclusion criteria. Information on the other dimensions is not described in the protocol. Although there is a general acceptance that the therapy is based on and should be executed according to the cognitive-behavioral approach to chronic pain, the protocol is not explicit about this and thus fails to provide a precise description of this approach.

4.3.2 Results of evaluation forms, multidisciplinary consultations and the evaluative interview

For every normative domain, the main inconsistencies between the planned treatment and actual treatment provided are described.

4.3.2.1 Treatment domain

The therapy strength of the CBT-R program, as defined by duration, frequency and intensity, is mainly based on years of experience with the treatment of chronic pain patients. A clear theoretical basis for the treatment strength of the CBT-R program is missing.

According to the protocol, the official treatment duration for every group is eight weeks. During the summer holidays, some groups had a summer break of two

or three weeks. However, aside from summer breaks, the time schedule for the treatment components generally lined up with the protocol. The frequency of the components and the time spent on each component also took place according to protocol. However, a few aspects of treatment did not. Often, therapy sessions were cut short by five to ten minutes as therapists and rehabilitants were frequently required to move from one location to another in order to attend the next session. In addition, responses from anonymous interviews with patients at the end of their treatment revealed that many patients found the frequency of some treatment components too low. Specifically, patients found that too little time was spent on consultations with the rehabilitation physician, on meetings with the occupational therapist and on vocational rehabilitation. Consultations with the rehabilitation physician were experienced as 'poor' or 'insufficient' by almost a third of the patients.

According to the therapists, most patients experience the CBT-R program as very intense. This was also confirmed in the anonymous interviews with patients. In these interviews, almost a fifth of the patients claimed that the variation between exertion and relaxation was indeed insufficient.

In several cases, due to vacation or illness of therapists, no occupational therapy or social work was available for the three semi-inpatient groups. Further, no replacement therapists were available due to significant staff shortages. Interviews with patients confirmed that indeed almost half of the patients missed some parts of the treatment as a result of cancellations by therapists and double-bookings with other treatment components. Frequently, individual sessions with a social worker, rehabilitation physician or occupational therapist were booked during other sessions such as physiotherapy or sports. For these reasons, rehabilitants missed out on parts of the CBT-R program.

Although most of the actual CBT-R program *content* corresponded with the protocol, there was a relatively large number of minor protocol violations. These violations are summarized in Table 4.1. A more detailed description of the evaluation results categorized according to discipline can be found in Appendix 4B.

Table 4.1: Inconsistencies between planned and actual treatment contents per discipline

Treatment discipline	**Inconsistencies**
Psychology	Specific requests, needs and requirements of the rehabilitants made it such that the treatment offered occasionally diverged from the protocol.
Social work	Not all topics were covered. The topics covered were determined based on the situation and needs of the rehabilitant. Additionally, members of the VW group were unable to have a consultation with the social worker during the first week as no regular social worker was available.
Physiotherapy	▪ In the DM group, one session was cancelled in the last week of treatment due to a follow-up treatment day planned for another group. Further, no opportunities to practice basic skills were provided and no specific forms of relaxation were taught during the third week. However, in the last week, some time was created to work on extra relaxation exercises. It is unknown whether or not specific attention was given to certain treatment aids and individual help on specific components of the program. ▪ In the DW group, specific daily activities were practiced only in the last two weeks of the program and this should have occurred earlier. Additionally, no specific relaxation exercises were practiced in the course of the treatment. Lastly, explanation of the principles of physiotherapy and formulation of practical goals was still being done in the fifth and sixth week of treatment, which, according to protocol, should have been completed earlier. ▪ Like the DW group, in the VW group, an explanation of the principles of physiotherapy and the formulation of practical goals continued to occur quite late, namely in the fifth and sixth week of treatment. Additionally, one rehabilitant received four individual consultations for a specific issue.
Fitness, swimming & sports	▪ On an individual basis, rehabilitants sometimes engaged in activities that were not described in the protocol as a result of patients' pain severity at that time. ▪ The DM did not do any aqua-jogging despite it being prescribed by the protocol. ▪ It is unclear whether the rehabilitants received advice for sport stimulation in the last week of treatment.

Table 4.1: Inconsistencies between planned and actual treatment contents per discipline (continued)

Treatment discipline	Inconsistencies
Occupational therapy	For all groups, it is unclear whether the rehabilitants had an individual intake during the first week of the program. Additionally, the following can be noted: ▪ In the VW group, no standing or sitting activities were done in the third week. ▪ The DM group's third week was cancelled. As a result, the program was postponed for one week. As a result, the facultative program started one week later. ▪ In the DW group, the first week was cancelled and the program was thus postponed for one week. However, the facultative program still started in the fifth week as would have been the case if the group had started on time. Further, in this group, several components were not executed. These components include: learning to get in and out of bed and filling in lists of standing and sitting activities. Additionally, this group focused on computer work more than prescribed in accordance with the requests of the rehabilitants.
Rehabilitation physician	Two rehabilitants of the DM group had an extra consultation outside the planned hours during the third week. Two rehabilitants of the DW group had an extra consultation during the second week.
Vocational rehabilitation	None of the therapists filled in the evaluation forms as hardly any patients participated in this discipline.

Therapists are expected to attend a half hour multidisciplinary team meeting (MDO) every week to discuss progress of every patient. However, not every therapist was present or on time due to overlap with other therapy sessions, vacations, or because they were required to move from one location to another. Additionally, half an hour was often too short to discuss every patient. Consequently, not every therapist received complete information on every patient. This means that some therapists were unaware of the patients' progress in treatment provided by other disciplines. As a result, it may have been difficult for therapists to effectively adjust their treatment to the needs of the patients as many therapists were unable to receive input from other therapists with respect to where individual accents should be placed. Although the results of the weekly MDOs were described in a report, it is unclear whether these reports were actually read by all therapists.

4.3.2.2 Implementation-environment domain

The implementation-environment was evaluated on two dimensions only, namely the participants and implementors (therapists). Other dimensions were not evaluated as they were not described in the protocol.

Participants

In general, the rehabilitants were both cooperative and active and the atmosphere in all groups was positive. Rehabilitants also almost always supported the aims of treatment. However, in one semi-inpatient group, there was some conflict due to the behavior of one of the rehabilitants. Regrettably, the extent to which this conflict impacted the group process is unknown. In addition, two members of the outpatient group appeared to lack sufficient mental capability to participate. Unfortunately, this was not recognized prior to treatment and only became apparent during the treatment. In another case, severe psychiatric problems emerged during treatment. Once again, it is unclear how these situations impacted the group process.

During treatment, it became apparent that two rehabilitants in the DW group had a much higher load capacity than the rest of the group. For one participant in the DM group, the program proved to be unsuitable as a result of her age. She was too young. Additionally, one rehabilitant in the DM group dropped out in the first week because he preferred a more somatic oriented treatment. Another rehabilitant in the DM group was overly occupied with psychological problems. In the VW group, there was a rehabilitant with a very low mental load capacity and a rehabilitant with noteworthy underlying psychopathology. In short, while most of the patients did meet the inclusion criteria and did not possess exclusion criteria, some participants were clearly not suited for treatment in the CBT-R program. Despite their lack of suitability, these patients usually continued to participate until the end of treatment.

On several occasions, rehabilitants were absent during sessions of psychology, physiotherapy, occupation therapy or sports. In the majority of these cases, the reason for absence was linked to a double-booking in which rehabilitants had planned sessions with a social worker or rehabilitation physician at the same time as other treatment components.

Implementors

All therapists had completed the required education and possessed the skills necessary to treat patients. However, for all the groups investigated, the absence of therapists was significant. Additionally, many patients experienced a switch in therapists. Both absence and therapist change were caused by staff shortages, vacancies, and illness. Illness played a significant role in the case of the social worker and occupational therapists. In the outpatient group, a psychologist took over the social work sessions as there was no social worker available. Additionally, the DM group received only a half hour of psychological therapy on Mondays. This was because the psychologists were too busy. In the VW group, occupational therapy was organized by both an occupational therapist and a physiotherapist. In addition, there was no regular social worker available for the VW group. It is important to note that the normative evaluation took place during the summer period. The results of this evaluation may thus fail to properly represent the actual absence of or changes in therapists.

With regard to the attitude of the therapists, anonymous interviews with patients at the end of treatment indicated that most patients experienced contact with the therapists and group supervisors as very good. Participants also indicated that the therapists treated the patients with sufficient respect, expertise and understanding. More than half of the patients claimed that therapists took their needs, requests and desires into consideration.

4.3.2.3 Outcome domain

For every discipline, specific aims of treatment were formulated in the protocol. These aims generally correspond with the general aims of the CBT-R program. Section 4.3.1 has already pointed out that the protocol fails to state exactly how these aims are operationalized and how these aims can be measured. Additionally, the discipline psychology introduces, in the protocol, the term ‘psychosocial load capacity’. Once again, a clear definition or operationalization of this term is not provided by either the protocol or by the Dutch dictionary. However, it is important to note that the results of the process evaluation indicate that most of the actual treatment aims do correspond with the aims described in the protocol. There are, however, a few inconsistencies, namely:

- Social work was not strictly limited to social work as described in the protocol. It appears that, in practice, social work also aimed to promote and create stimulating conditions so that patients would be better able to apply the newly

learned behaviors in daily life. This aim is not mentioned anywhere in the protocol.

- The protocol indicated that rehabilitation physicians are limited to providing advice with respect to provisions, aids and adaptations. One rehabilitation physician not only provided advice but also prescribed certain provisions, aids and adaptations to participants. This prescription is not mentioned in the protocol
- The protocol states that, during the last week of physiotherapy, rehabilitants are supposed to work out some practical goals for the period between end of the program and follow-up. Given that a DM group session in the last week was cancelled, it is impossible to establish whether or not this was actually done with the DM group.

4.3.3 Results of the THQ

In Table 4.2, the mean (standard deviation) and range scores on the THQ are presented. The experienced helpfulness of vocational rehabilitation and creativity is not presented here, as too many patients failed to rate these components. Table 4.2 shows that psychology lessons, physiotherapy and social work were considered as the most useful components of the CBT-R program. Responses from anonymous interviews with patients at the end of treatment confirmed the findings in Table 4.2. These interviews were conducted during the same period of investigation. Most of the patients experienced the CBT-R program as very helpful for both themselves as well as for other chronic pain patients. More than half of the patients indicated an improvement in functioning. Functioning was considered to be good or sufficiently acceptable after treatment. Only a few patients (< 5%, N = 98) indicated that their pain problem had worsened. Additionally, more than half of the patients claimed that the information provided on the content, duration and aims of the CBT-R program was either good or sufficient. About a quarter of patients considered the provision of information to be poor. These patients wanted more information about how daily life activities could be better adapted (occupational therapy), about pain control and about specific pain complaints. With respect to the information and explanations on treatment components and treatment goals, almost all patients claimed that this information was either good or at least sufficient.

Table 4.2: Experienced helpfulness of the CBT-R treatment components as measured with the THQ (N = 85)

Treatment component	Mean	SD	Minimum	Maximum
Rehabilitation physician	1.85	1.88	-4.00	5.00
Psychology lessons	3.96	1.07	-0.30	5.00
Psychology exercises	3.41	1.48	-1.50	5.00
Psychology visualizations	3.31	1.45	-0.70	5.00
Physiotherapy (relaxation)	3.91	1.38	-2.50	5.00
Fitness	2.26	2.23	-4.50	4.90
Swimming	2.61	2.24	-4.50	5.00
Sports	1.50	2.39	-4.80	5.00
Social Work	3.99	1.20	-0.50	5.00
Occupational therapy	2.06	1.73	-1.40	5.00
Activity Therapy	2.59	1.67	-1.80	5.00

Using the THQ, an effort was made to determine whether the treatment effect of the CBT-R program, as measured during the RCT (Chapter 2), was related to the degree to which patients considered the treatment components helpful. In Chapter 2, only the results of the Multidimensional Pain Inventory – Dutch Language Version (MPI-DLV) and the Symptom Checklist-90 (SCL-90) were discussed. Other instruments of the RCT included seven items of the RAND 36-item Health Survey (RAND-36)[12] and the visual analogue scale (VAS) of the McGill Pain Questionnaire-Dutch Language Version (MPQ-DLV)[13]. Treatment effect on these items was also analyzed for the present study and calculated for every patient by subtracting the pre treatment scores from the post treatment scores (26 variables). No distinction was made between the intervention and the control group.

Because of the relatively small N available (81) and the large number of comparisons (regressions) that had to be performed (286), a decision had to be made to reduce the number of treatment components (predictors) included in the analyses. A factor analysis revealed that the set of 11 treatment components could be reduced to three factors, namely a 'psychic' factor (with high loadings of > 0.5 for the psychology components and occupational therapy), a 'physical' factor (with high loadings for rehabilitation physician, fitness, sports, swimming, and activity therapy) and a third factor with only a high loading (0.88) for the relaxation component of physiotherapy. 26 regressions were performed with each aspect of

treatment effect as the dependent variable and the three abovementioned factors as predictors (Table 4.3).

Table 4.3: Regression analyses between treatment effect and helpfulness factors

Instrument		**t value (*P* value)**			**F-value (*P*)**	**R^2**
		Psychic	Physical	Physiotherapy		
MPI-DLV	Pain severity	-.525 (.601)	-.832 (.409)	-.213 (.832)	.759 (.522)	.040
	Interference	.917 (.363)	-.821 (.415)	.057 (.955)	.357 (.784)	.019
	Life control	-.807 (.423)	1.696 (.096)	-.045 (.964)	.993 (.403)	.051
	Negative distress	.176 (.861)	-1.025 (.310)	.704 (.484)	.439 (.726)	.023
	Support	-.168 (.867)	.662 (.511)	-2.461 (.017)	2.186 (.101)	.112
	Disctracting responses	-.679 (.500)	1.598 (.116)	-1.393 (.169)	1.291 (.287)	.066
	Punishing responses	-.247 (.806)	-.781 (.438)	.473 (.638)	.347 (.791)	.019
	Solicitous responses	-.740 (.462)	1.071 (.289)	-.713 (.479)	.553 (.648)	.029
	General activity	-.636 (.528)	1.346 (.184)	2.454 (.017)	3.533 (.021)	.162
SCL-90	Fear	-.784 (.436)	.330 (.743)	1.134 (.262)	.572 (.636)	.030
	Agoraphobia	-.376 (.708)	.092 (.927)	.699 (.487)	.188 (.904)	.010
	Depressive thoughts	.138 (.891)	-1.461 (.150)	-.132 (.895)	.977 (.410)	.052
	Somatic complaints	.870 (.388)	-.345 (.731)	-.405 (.687)	.269 (.847)	.014
	Insufficiency	.151 (.880)	-.696 (.490)	-1.104 (.274)	.849 (.473)	.045
	Sensitivity	.000 (1.000)	-.869 (.389)	-.383 (.703)	.508 (.678)	.027
	Hostility	.465 (.644)	-.122 (.903)	-.310 (.758)	.088 (.967)	.005
	Sleeping problems	.000 (1.000)	-.869 (.839)	-.383 (.703)	.508 (.678)	.027
	Psycho-neuroticism	.423 (.674)	-.960 (.341)	.097 (.923)	.311 (.817)	.017
MPQ-DLV	VAS	-1.321 (.192)	-.653 (.517)	-.968 (.337)	2.365 (.081)	.116
RAND-36	Physical functioning	1.385 (.172)	-.265 (.792)	1.769 (.082)	2.592 (.062)	.124
	Social functioning	-.555 (.581)	1.317 (.193)	.785 (.436)	1.031 (.386)	.053
	Mental health	-.933 (.355)	1.662 (.102)	-.330 (.743)	.941 (.427)	.049
	Vitality	-.226 (.822)	2.153 (.036)	-.092 (.927)	1.907 (.139)	.094
	Pain	.387 (.700)	1.899 (.063)	1.649 (.105)	4.288 (.009)	.190
	General health experience	1.564 (.124)	-.822 (.414)	2.469 (.017)	3.828 (.015)	.173
	Health change	.770 (.445)	.507 (.614)	.207 (.837)	.675 (.571)	.036

Evidently, three models were found to be significant at the 0.05 level. The *P* values of the three predictor variables (factors) indicated that the relaxation component of physiotherapy was a significant predictor in three cases.

4.3.4 Post hoc analyses of the results of the CBT-R program

In a post hoc analysis of the RCT results, an effort was made to explain treatment variability in the CBT-R program by means of process variables. For these analyses, only results of the semi-inpatient groups were used. No distinction was made between the intervention and the control group. In the CBT-R program, the semi-inpatient treatment takes place in two groups, namely the so-called 'Voorhuis' group (VH) and the 'Deel' group (D). A comparison was made between results of patients from the VH group (N = 67) and patients from the D group (N = 28). First and foremost, differences between these groups with regard to general characteristics were determined. No differences were found with regard to age, sex, pain location, pain duration, education, drop out rate, and MPI-DLV classification at baseline (all *P*'s > 0.05). Following this, the treatment effect for every variable measured in the RCT was calculated by subtracting the pre treatment scores from the post treatment scores. The differences in treatment effect between the two groups were determined by means of t-tests (Table 4.4). Although, in both groups, the CBT-R program was mostly conducted according to protocol, the results of the t-tests indicate that there was a significant difference between the VH and the D group in treatment effect on the MPI-DLV Pain Severity ($t_{(88)} = -2.095$; $P = .039$) and Interference scales ($t_{(88)} = -2.367$; $P = .020$). The VH group demonstrated an improvement with regard to pain severity and interference, while the D group showed deterioration on these items. Additionally, between the two groups, several other variables showed an opposite pattern with respect to treatment effect. Although the patterns were not statistically significant, these patterns should not be left unconsidered.

Table 4.4: Difference in treatment effect between VH and D groups

Instrument		Mean (SD)		T value (df)	*P* value
		VH	D		
MPI	Pain severity	-0.39 (1.10)	0.10 (0.79)	-2.095 (88)	0.039
	Interference	-0.27 (0.69)	0.10 (0.62)	-2.367 (88)	0.020
	Life control	0.02 (1.12)	0.15 (1.21)	-0.526 (88)	0.600
	Negative distress	-0.23 (1.03)	0.10 (1.12)	-1.344 (88)	0.182
	Support	-0.29 (0.86)	-0.44 (0.62)	0.803 (83)	0.424
	Punishing responses	-0.35 (1.01)	-0.05 (0.76)	-1.368 (87)	0.175
	Solicitous responses	-0.21 (0.86)	-0.05 (0.82)	-0.769 (87)	0.444
	Distracting responses	-0.03 (0.98)	-0.05 (0.82)	0.074 (87)	0.941
	General activity	0.00 (0.61)	-0.12 (0.41)	0.908 (88)	0.366
SCL	Psychoneuroticism	-10.69 (30.83)	-3.62 (20.76)	-1.071 (86)	0.287
MPQ	VAS	-0.61 (2.09)	-0.28 (1.96)	-0.711 (88)	0.479
RAND-36	Physical functioning	-0.2 (11.88)	1.94 (13.77)	-0.684 (88)	0.496
	Social functioning	2.58 (20.10)	2.04 (19.28)	0.119 (88)	0.906
	Mental health	1.33 (12.44)	1.33 (10.41)	0.000 (88)	1.000
	Vitality	3.02 (12.78)	3.98 (13.63)	-0.322 (88)	0.748
	Pain	3.34 (18.17)	0.30 (13.44)	0.780 (88)	0.437
	General health experience	3.65 (13.48)	0.00 (10.56)	1.251 (88)	0.214
	Change in health	13.10 (27.26)	9.72 (28.66)	0.530 (88)	0.598

Figure 4.2 also shows that the VH and D group differ with respect to results on the MPI-DLV variable Interference. For every measurement, the results shown are categorized according to control and intervention group (for a detailed description of when measurements occurred, please see Figure 2.1.). Figure 4.2 indicates that, for the VH control and the VH intervention group, the results resemble the general effect on MPI-DLV Interference that was demonstrated in Figure 2.2. In essence, the VH control group roughly remained the same between t0 and t2 but showed a small decrease on the Interference scale between t2 and t3, while the VH intervention group showed a decrease between t1 and t2 that continued slightly until t3. Both the D intervention and the D control group showed, respectively, a varying pattern and no change on the Interference scale.

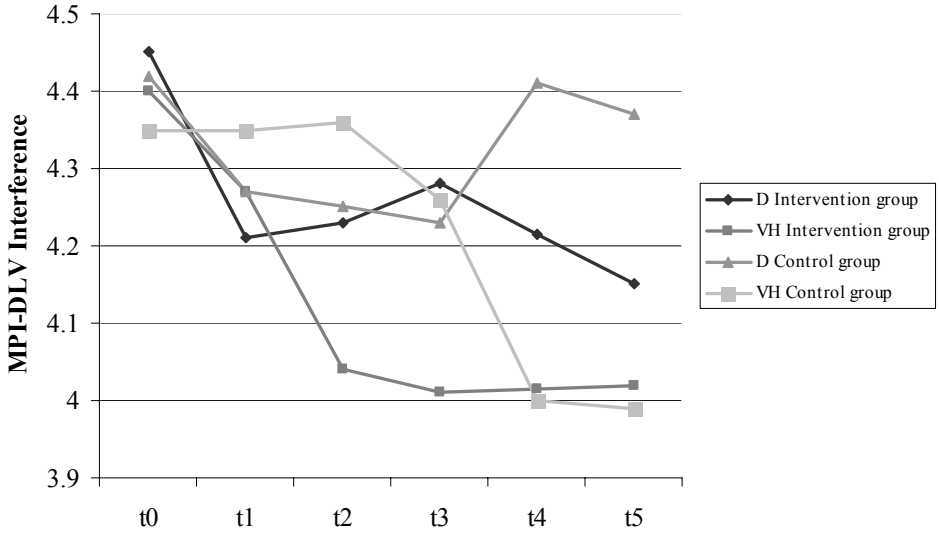


Figure 4.2: MPI-DLV Interference over time for the VH and the D groups

4.4 Discussion

In this study, the CBT-R program was evaluated by looking at the treatment process. An attempt was made to establish the congruency of the planned CBT-R program (the protocol) and the treatment that was actually implemented in daily practice. Unfortunately, the protocol failed to clearly describe and operationalize the treatment, outcome and implementation-environment domains. The protocol itself may very well be a useful tool for the therapists in their daily practice but, for an extensive evaluation of the treatment process, it has proven to be unsatisfactory for the following reasons: 1) the theories for treatment *strength* appear to be insufficient; 2) the selection criteria are not operationalized, thereby making it difficult to determine which patients are actually suitable for participation; 3) for many treatment activities, the protocol fails to describe the theoretical background and specific aims of these activities making it thus unclear how, theoretically, the activities should lead to the attainment of goals; 4) both the general and discipline-specific aims of the program are described in vague terms; 5) most of the aims of the CBT-R program are not operationalized; and 6) no description is given with respect to how treatment outcome can be measured. As a result, it was difficult and

basically impossible to adequately determine the congruence between the planned treatment and the actual treatment. Unfortunately, a poorly operationalized protocol for treatment almost inevitably leads to substantial treatment variability.

One may argue that the instruments applied in the present study were insufficient and thus unable to determine what actually occurred during the implementation of the CBT-R program. However, the choice of evaluation instruments was severely restricted as the available protocol was unclear and, in many ways, incomplete. The evaluation could have been strengthened by including additional treatment groups in the process evaluation or by having independent observers monitor the entire treatment process (all eight weeks). However, both time and budgetary restraints made these options impossible.

The results of the THQ demonstrated that, in general, patients were satisfied with the CBT-R program. The experienced helpfulness of the physiotherapy sessions (specifically, the relaxation component) turned out to be a significant predictor for treatment effect. The smaller experienced helpfulness of the other treatment components may have contributed to the variability in treatment results of the CBT-R program found in Chapter 2.

The results of the evaluation done by therapists indicated that a few incongruencies between the planned treatment and the actual treatment exist. Most of the actual contents and goals of the CBT-R program did however correspond with the protocol. Differences between the protocol and the treatment that was implemented were found with respect to cancellations of therapy sessions (due to vacations, illness and staff shortages), double-booking of therapy sessions, and shorter therapy sessions (due to the fact that time was required to move from one location to another). In addition, some program components were adapted to fit the specific needs and requests of the rehabilitants. It is possible that this very incongruency may even have a positive impact on treatment outcome. Further, differences in treatment effect were found between the VH and the D group. The only differences between these groups ought to be the location of treatment in the rehabilitation center and the involved therapists. In principle, these differences should not affect treatment outcome. However, a significant difference in outcome between the VH group and the D group was found with regard to the variables Pain Severity and Interference. So far, only speculations can be made about the causes of this difference. The interviewed therapists did however indicate that the differences between groups may have been related to the fact that, in the D group, patients experienced high therapist turnover as the therapists assigned to this group

were often ill. Another potential explanation is that the organization in the D group was more troublesome than in the other group during the period of investigation. Conclusions about the differences in treatment effect between the VH group and the D group can only be made when it is entirely clear, for both groups, how and which treatment components were exactly implemented. Nevertheless, the fact that treatment variability could be demonstrated on the basis of these groups implies that differences, and thus program failures in the treatment process, do exist. Obviously, these program failures and the incongruencies already described need to be identified and resolved before the assumption can be made that treatment variability is caused by theory failures.

Before any conclusions can be drawn with regard to program or theory failures, the treatment process of the CBT-R program has to be extensively described and operationalized through protocol. It is also imperative that the congruence between the planned and actual treatment situation be further investigated before conclusions can be drawn. Once the process is extensively described and the congruence better investigated, the actual impact of the treatment process on the treatment effect and variability can be better understood. This is particularly relevant for the CBT-R program as the program offers a multidisciplinary treatment environment. When several disciplines are involved in the treatment process, it is difficult to link the overall treatment outcome to discipline-specific activities.

In the existing scientific literature, numerous studies on the effectiveness of multidisciplinary cognitive-behavioral therapy (CBT) have been published. Most conclude that multidisciplinary CBT is indeed an effective method for treating chronic pain[14,15]. If one takes a good look at academic publications on CBT and the several CBT handbooks that exist, it becomes clear that a broad range of cognitive and behavioral techniques can be used to accomplish the primary goals of CBT. The primary goals of CBT include improving patients' physical and emotional functioning and health-related quality of live. According to Turk[16] (personal communication), the fact that the techniques are highly diverse does not invalidate the general model of the cognitive-behavioral approach of pain. Whether treatment programs follow the general cognitive-behavioral perspective is more important and thus of greater interest in establishing the effectiveness of CBT. On the other hand, it may be argued that the contents of the applied components of CBT should be unequivocal. Failing this, attributing the outcome of treatment to certain

components, and thus establishing which components are most effective, becomes difficult.

Lastly, it is important to note that program evaluation is always an ongoing process. To date, there are no perfect treatment programs for chronic pain. In every program, treatment variability can be found. Numerous factors influence the treatment process. As a result, treatment outcomes may vary along with these factors. If we truly want to improve treatment programs, we have to continue evaluating them and this can best be done by alternating evaluations of the normative and causative theories of a program. However, in most cases, time and financial restrictions limit the possibilities for evaluation. In the case of the present study, the number of influencing factors that was evaluated was also limited. As a result, the present study should be seen as a first step towards a more effective and more optimal pain management program.

References

1. Turk, D.C., Rudy, T.E., & Sorkin, B.A. (1993). Neglected topics in chronic pain outcome studies: determination of success. *Pain, 53,* 3-16.
2. Dekker, P.J., & Leeuw, F.L. (1989). *Beleids- en programma-evaluaties: een stand van zaken in de sociaal-wetenschappelijke literatuur*. Leiden: DSWO Press.
3. James, G. (1961). *Evaluation and planning of health programs. Administration of community health services*. Chicago: Int. City Managers Association.
4. Suchman, E.A. (1967). *Evaluative research: principles and practive in public service and socail action programs*. New York: Russel, Sage Foundation.
5. Chen, H.T. (1990). *Theory-driven evaluations*. London: Sage Publications.
6. Sijben, N. (1984). Programma-evaluatie…wat is dat? In: G.M. Schippers, N. Sijben, C. Schaap, & J. Mastboom (Eds.) (1984). *Programma-evaluatie in de geestelijke gezondheidszorg.* Lisse: Swets & Zeitlinger, p. 11-27.
7. Sijben, N. (1987). *Omzien naar weldoen. Programma-evaluatie in theorie en praktijk*. Nijmegen: B.V. Boekdrukkerij F.E. MacDonald.
8. Donker, M. (1990). Principes en praktijk van programma-evaluatie. Utrecht: NcGv.
9. Wijck, R. van (1997). *Programma evaluatie en kwaliteit van zorg. Onderzoek naar verbetering van de zorg voor mensen met meervoudige beperkingen*. Groningen: Rijksuniversiteit, bewegingswetenschappen.
10. Roessingh (1999). Beslissingsondersteunend en verrichtingenprotocol. De pijn de baas: semiklinisch revalidatieprogramma voor chronische pijnpatiënten. Enschede: Divisie Pijnrevalidatie.
11. Chapman, S.L., Jamison, R.N., & Sanders, S.H. (1996). Treatment Helpfulness Questionnaire: A measure of patient satisfaction with treatment modalities provided in chronic pain management programs. *Pain, 68*, 349-361.
12. Zee, K.I. van der, & Sanderman, R. (1993). Het meten van de algemene gezondheidstoestand met de RAND-36: Een handleiding. Groningen: Noordelijk Centrum voor Gezondheidsvraagstukken (NCG).
13. Verkes, R.J., Vanderiet, K., Vertommen, H., Kloot, W.A. van der, & Meij, J. van der (1989). De MPQ-DLV: Een standaard Nederlandstalige versie van de McGill Pain Questionnaire voor België en Nederland. In: W.A. van der Kloot, & H. Vertommen (red.) - *De MPQ-DLV: Een standaard nederlandstalige versie van de McGill Pain Questionnaire; achtergronden en handleiding.* Lisse: Swets & Zeitlinger, p. 57-69.
14. Flor, H., Fydrich, T., & Turk, D.C. (1992). Efficacy of multidisciplinary pain treatment centers: a meta-analytic review. *Pain, 49*, 221-230.
15. Morley, S., Eccleston, C., & Williams, A. (1999). Systematic review and meta-analysis of randomized controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache. *Pain,* 80, 1-13.
16. Turk, D.C., & Okifuji, A. (2003). Cognitive-behavioral approach to pain management. In: R. Melzack, & P.D. Wall (Eds.). *Handbook of pain management*. London: Churchill Livingstone, p. 533-541.

17. Winter, F.A.M. (1988). Werkboek: De pijn de baas. Enschede: Intermap.

Appendix 4A : Weekly observation form for psychologists

Datum: .... Behandelweek: ....

Naam psycholoog: ....

Pijngroep: Deel-Ma / Deel-Wo / Voorhuis-Ma / Voorhuis-Wo / poliklinisch
*(*streep door wat niet van toepassing is)*

---

1. Aan welk thema(s) is deze week gewerkt? ....

2. Kruis in onderstaande tabel aan wat je de afgelopen week gedaan hebt tijdens de behandeling (meerdere antwoorden mogelijk).

	Maandag/Woensdag	Dinsdag/Donderdag	Woensdag/Vrijdag
Ochtend	❒ uitleg doel en werkwijze ❒ theorie behandeld uit werkboek ❒ vragen behandeld uit werkboek ❒ ontspanningsoefeningen/ ❒ visualisaties ❒ bankoefening ❒ valoefening ❒ vertrouwensoefening ❒ assertiviteitsoefening (communicatie, presentatie) ❒ formuleren persoonlijke doelstellingen ❒ anders, namelijk	❒ uitleg doel en werkwijze ❒ theorie behandeld uit werkboek ❒ vragen behandeld uit werkboek ❒ ontspanningsoefeningen/ ❒ visualisaties ❒ bankoefening ❒ valoefening ❒ vertrouwensoefening ❒ assertiviteitsoefening (communicatie, presentatie) ❒ formuleren persoonlijke doelstellingen ❒ anders, namelijk	❒ uitleg doel en werkwijze ❒ theorie behandeld uit werkboek ❒ vragen behandeld uit werkboek ❒ ontspanningsoefeningen/ ❒ visualisaties ❒ bankoefening ❒ valoefening ❒ vertrouwensoefening ❒ assertiviteitsoefening (communicatie, presentatie) ❒ formuleren persoonlijke ❒ doelstellingen ❒ anders, namelijk ....
Middag	❒ uitleg doel en werkwijze ❒ theorie behandeld uit werkboek ❒ vragen behandeld uit werkboek ❒ ontspanningsoefeningen/ ❒ visualisaties ❒ bankoefening ❒ valoefening ❒ vertrouwensoefening ❒ assertiviteitsoefening (communicatie, presentatie) ❒ formuleren persoonlijke doelstellingen ❒ anders, namelijk	❒ uitleg doel en werkwijze ❒ theorie behandeld uit werkboek ❒ vragen behandeld uit werkboek ❒ ontspanningsoefeningen/ ❒ visualisaties ❒ bankoefening ❒ valoefening ❒ vertrouwensoefening ❒ assertiviteitsoefening (communicatie, presentatie) ❒ formuleren persoonlijke doelstellingen ❒ anders, namelijk	❒ uitleg doel en werkwijze ❒ theorie behandeld uit werkboek ❒ vragen behandeld uit werkboek ❒ ontspanningsoefeningen/ ❒ visualisaties ❒ bankoefening ❒ valoefening ❒ vertrouwensoefening ❒ assertiviteitsoefening (communicatie, presentatie) ❒ formuleren persoonlijke doelstellingen ❒ anders, namelijk ....

3. Heb je deze week (bij één of meerdere revalidanten) iets anders gedaan dan je vooraf gepland had? Zo ja, wat en waarom?
   Gepland: .....
   .....
   .....
   Daadwerkelijk: .....
   .....
   .....
   Reden: .....
   .....

4. Geef hieronder aan welk huiswerk de revalidanten deze week hebben meegekregen.
   .....

5. Is het huiswerk van vorige week behandeld?
   - ❒ ja
   - ❒ nee, reden: .....
     .....

6. Kruis hieronder aan welke doelstellingen je deze week hebt nagestreefd (meerdere antwoorden mogelijk).
   - ❒ Vergroten psychosociale belastbaarheid
   - ❒ Onderkennen en hanteren eigen grenzen t.a.v. psychosociale belastbaarheid
   - ❒ Herstellen balans ontspanning, inspanning en afleiding
   - ❒ Vergroten inzicht in eigen functioneren
   - ❒ Vergroten inzicht in pijn en factoren die daarop van invloed zijn
   - ❒ Kennisopbouw omtrent theorie over pijnmanagement-principes
   - ❒ Leren toepassen pijnmanagement-principes in praktijk
   - ❒ Leren toepassen van ontspanningsoefeningen
   - ❒ Vergroten assertieve vaardigheden
   - ❒ Vergroten communicatieve vaardigheden
   - ❒ Leren omgaan met eigen beperkingen en mogelijkheden
   - ❒ Ontwikkelen positief zelfbeeld
   - ❒ Anders, namelijk .....

7. Zijn er revalidanten bij wie de behandeling niet goed aansluit?
   - ❒ nee
   - ❒ ja, bij ..... revalidanten (vul in aantal)
     Reden(en):
     - ❒ te hoge mentale belastbaarheid
     - ❒ te lage mentale belastbaarheid
     - ❒ te hoge fysieke belastbaarheid
     - ❒ te lage fysieke belastbaarheid
     - ❒ taalprobleem
     - ❒ revalidant staat niet achter doelstelling van behandeling
     - ❒ anders, namelijk: .....

8. Zijn er revalidanten afwezig geweest tijdens één van de therapie-uren of hebben revalidanten niet meegedaan?
   - ❒ nee
   - ❒ ja, aantal afwezig ....., aantal niet (volledig) meegedaan: .....
     Reden(en):
     - ❒ programma te zwaar
     - ❒ oorzaak in thuissituatie revalidant
     - ❒ ziekte
     - ❒ dubbelplanning
     - ❒ anders, ...
     - ❒ weet niet

9. Deze vraag alleen invullen aan het eind van week 8!!!
   Kruis hieronder aan welke doelstellingen **niet** bereikt zijn bij de revalidanten bij wie je daar wel naar gestreefd hebt. Geef ook het aantal revalidanten bij wie die doelstelling niet is bereikt.
   - ❒ Vergroten psychosociale belastbaarheid (aantal: .....)
   - ❒ Onderkennen en hanteren eigen grenzen t.a.v. psychosociale belastbaarheid (aantal: .....)
   - ❒ Herstellen balans ontspanning, inspanning en afleiding (aantal: .....)
   - ❒ Vergroten inzicht in eigen functioneren (aantal: .....)
   - ❒ Vergroten inzicht in pijn en factoren die daarop van invloed zijn (aantal: .....)
   - ❒ Kennisopbouw omtrent theorie over pijnmanagement-principes (aantal: .....)
   - ❒ Leren toepassen pijnmanagement-principes in praktijk (aantal: .....)
   - ❒ Leren toepassen van ontspanningsoefeningen (aantal: .....)
   - ❒ Vergroten assertieve vaardigheden (aantal: .....)
   - ❒ Vergroten communicatieve vaardigheden (aantal: .....)
   - ❒ Leren omgaan met eigen beperkingen en mogelijkheden (aantal: .....)
   - ❒ Ontwikkelen positief zelfbeeld (aantal: .....)
   - ❒ Op een adequate manier met de pijn omgaan (aantal: .....)
   - ❒ Anders, namelijk ..... (aantal: .....)

10. Ruimte voor opmerkingen
    .....
    .....
    .....
    .....
    .....
    .....
    .....
    .....
    .....
    .....
    .....
    .....
    .....
    .....
    .....

Appendix 4B: Congruence between planned and actual treatment contents

In this section, the main inconsistencies between the planned and actual treatment contents are presented by discipline.

Psychology
The norm is to cover one chapter of the workbook every week. As a result, only the first eight chapters of the workbook are included in the treatment program itself[17]. The rehabilitant is thus required to learn the last two chapters on his or her own time. The patients did receive homework assignments every week. These assignments included reading the chapter for the following week and completing the corresponding questions. As a general rule, every planned activity was executed according to the sequence described in the protocol. However, analyses of the therapists' evaluation forms revealed that divergence from the protocol did happen on occasion. This was done in accordance with the requests, needs and requirements of the rehabilitants.

Social work
No evaluation forms were received from one of the semi-inpatient groups (VW) as no regular social worker was available. For the same reason, this group also did not have the expected consultation in the first week of the program. However, in general, the protocol was followed. The topics that were covered were determined by the needs and situation of the rehabilitant. In every rehabilitant, different factors may influence the pain complaints. A definite sequence of topics to be treated is thus uncommon. With the problems of the rehabilitant in mind, a plan of action is made and then followed. When situations and factors that influence a patient's functioning either at home or in the treatment situation appear, the plan of action is changed. The plan can thus be adjusted according to the needs of the patient.

Physiotherapy
The rehabilitants were assigned homework every week. It appears that, as a general rule, every rehabilitant actually completed the homework assigned on all occasions. In the final week of the program, one physiotherapy session planned for the DM group was cancelled due to a follow-up treatment day planned for another group. During this session, the intention was to have rehabilitants work out some practical

goals for the period between the end of the program and the follow-up treatment day. It is unclear whether this was actually done during the other session of the same week. Further, the protocol states that physiotherapy in the first four weeks of the program should aim to teach patients functional body positions as well as basic principles and skills with regard to relaxation and daily activity. In the final four weeks of therapy, more specific forms of relaxation and activity are practiced. For the DM group, physiotherapy did not follow strict protocol. In the fourth week of therapy, the DM group was already focusing on specific forms of relaxation and daily activities. Additionally, no basic skills were practiced and no specific forms of relaxation were done in the third week. However, time was spent on extra relaxation exercises in the final week when the group had some time leftover. In the end, every planned component was executed. Although an evaluation did take place in the final week of treatment, it is unclear whether attention was paid to preparing for the follow-up treatment day. Whether advice was given, whether exercises were assigned and whether goals were created remains unclear. Additionally, the evaluation failed to indicate whether patient's received personal attention and advice with respect to specific program components and certain aids. Further, for both the DW and VW group, graded activity was mentioned as an activity. However, the protocol indicates that graded activity should be a goal. It is also important to note that in the DW group, practicing specific daily activities was only done in the final two weeks of treatment, while the protocol indicates that this should be done from the fifth week on. Besides, no specific relaxation exercises were practiced in the DW group. The protocol states also that, in the first four weeks, the principles of physiotherapy should be explained and practical goals should be formulated. In the fifth and sixth week of the DW and VW group, the principles of physiotherapy were still being explained. Lastly, in the VW group, one rehabilitant had four individual consultations for a specific concern. This diverges from the protocol as most physiotherapy is done in groups.

Fitness, swimming & sports

Since no evaluation forms for swimming or sports were completed during the first weeks of treatment in the VM and VW groups, no definite inconsistencies can be established. A post-hoc inquiry with therapists however indicates that these activities were conducted in accordance with the protocol. Further analyses of the therapists' evaluation forms for the DM group revealed that this group did not participate in aqua-jogging. Additionally, some rehabilitants engaged in activities

that diverge from the protocol because of the severity of their pain complaints at the time. Lastly, although sports stimulation is considered a separate component of treatment, it is important to mention that it is unclear whether the rehabilitants received advice for sports stimulation in the last week of treatment.

Occupational therapy

It is unclear whether the rehabilitants were part of an individual intake during the first week of the program. This is the case for all groups. In the VW and DM group, no standing or sitting activities were done in the third week. Further, the third week of occupational therapy was cancelled for the DM group. The program was thus postponed for one week. As a result, the facultative program started one week later. The first week of the DW group was also cancelled and thus postponed one week. In this case, however, the facultative program started in the fifth week as initially planned. Lastly, several treatment components were not executed in the DW group. These components include learning to get in and out of bed and filling in lists of standing and sitting activities. In this group, the focus was placed primarily on computer work as the rehabilitants indicated that this was what they needed.

Rehabilitation physician

In the third week, two rehabilitants of the DM groups had an extra consultation outside the planned hours. In the second week, two rehabilitants of the DW group had an extra consultation.

Vocational rehabilitation

None of the therapists filled in the evaluation forms as patient participation in this discipline was extremely low. Low participation was likely due to the fact that most patients receive worker's compensation for their disability.

CHAPTER 5

Pain expectations and experiences: replicating the match-mismatch model in daily life of chronic pain patients

Adapted from: C.A.J. Mes, R. Lousberg, C.G.M. Oudshoorn, G. Zilvold, P.A.E.G. Delespaul, D.C. Turk. Pain Expectations and Experiences: Replicating the Match-Mismatch Model of Pain in Daily Life (submitted).

5.1 Introduction

Several models have been proposed to explain features of chronic pain. Etiological and maintaining mechanisms are elusive and are not completely understood. There are several theories that focus on somatic, psychological, and social aspects of pain and the interrelationship among them[1,2]. There is increasing interest in the so-called Match-Mismatch (MM) model of pain, which is based on the MM model of fear developed by Rachman and Lopatka[3,4]. The MM model of pain offers a possible explanation for the maintenance of chronic pain[5,6]. The model states that the most important immediate consequence of a mismatch between expected and experienced intensity of a painful event is that the expectation for the next episode is adapted in the direction of the previous experience. When the event is more painful than anticipated (i.e., an underprediction), the expectation for the next event is elevated so that people anticipate greater pain. Conversely, when the preceding event is less painful than anticipated (i.e., an overprediction), the expectation for the next incidence is adjusted so that the anticipated pain severity is lowered. When the expectation for the anticipated pain severity is accurate (i.e., a match) no changes in future pain expectations would be expected to occur.

One of the central issues in the MM model of pain is the hypothesized negative effect of an underprediction of pain. Several studies have demonstrated that underpredictions are related to an increase of fear of pain, fear of movement. They lead to the appearance of escape and avoidance behavior[6,7,8,9,10,11]. Also, the experienced level of pain remains heightened following an underprediction and subjects remain cognitively 'disturbed'[6]. Thus, patients who repeatedly underpredict their pain may develop inaccurate and maladaptive thoughts that may in turn induce elevated emotional and physiological arousal that may exacerbate pain. Moreover, the concurrent escape and avoidance behavior may eventually not only cause a syndrome of disuse, but also prevent patients from testing their predictions and from adjusting them in case they are inaccurate[7].

Inconsistent evidence has been published that underpredicted painful stimuli are experienced as more aversive than correctly or overpredicted pain. Arntz and Hopmans have reported that underpredictions are followed by less pain[12], while studies of Von Baeyer, Carlson and Webb or Spafford, Von Baeyer and Hicks show that underpredictions result in more pain[13,14]. Also, inconsistent results regarding the structural tendency of making under versus overpredictions in

populations experiencing pain have been reported. Linton and Melin as well as Arntz, Van Eck and Heijmans report that pain sufferers tend to overpredict their pain[15,16], whereas studies from McCracken et al. and Arntz and Peters report the opposite[17,18]. Methodological differences between study designs might partly explain these inconsistencies. For example, Linton and Melin asked chronic low back pain (CLBP) patients to recall the pain they experienced several weeks ago[15], while McCracken et al. studied the predictions of CLBP patients with regard to impending physiotherapeutic exercises[17]. Moreover, the duration of the existing pain problem differs between studies. Arntz, Van Eck and Heijmans studied the predictions of dental patients[16], whose pain is of a more or less acute nature, whereas McCracken et al. and Arntz and Peters studied predictions of patients with long lasting low back pain problems[17,18]. Further, it seems that the level of anxiety modulates the tendency of making under or overpredictions[16,17]. Arntz and Peters have stated that it might be possible that there are two groups of pain patients: those who are afraid of pain and show avoidance might manifest the tendency to overestimate pain and are probably more anxious and depressed than those who tend to be tough and challenge themselves by engaging in (painful) activities, which is contra productive[18]. The last group will probably show a tendency to underpredict pain. Poulton et al. have suggested that it may be normative (or even adaptive) to overpredict pain initially and to modify these predictions as new information becomes available[19]. Due to lack of exposure, excessive or unrealistic overpredictions are never disconfirmated, and this perpetuates avoidance behavior and its negative consequences. So far, scientific literature does not give a conclusive clarification for the tendency of chronic pain patients to underpredict or overpredict pain.

It is important to note, however, that the majority of studies on the MM model of pain are based on results obtained in experimental or laboratory studies where the occurrence of mismatches is under experimental control. For example, in several studies participants typically are instructed to predict the intensity of pain of an electrical shock, while underpredictions are induced by temporally increasing the shock level[7,9,18]. Other studies concerned the prediction of pain in patients receiving dental treatment[16] or children undergoing ear piercing[13,14]. In an experimental situation, only a defined episode is covered whereas in daily life the chronic experiences are more part of the overall process of experience. Whether the MM model applies in daily life has not been studied and remains to be determined. In a natural environment, pain assessments are part of the ongoing stream of

experiences and a large range of factors that are part of normal ecological daily reality and challenges can modify the predictions. Therefore, the aim of this study is to examine the predictions of the MM model in the daily life situation of chronic pain patients. Based on the results of previous studies, the following hypotheses are tested in daily life: (1) Underpredictions of pain will be followed by increases in *predicted* pain; (2) Overpredictions of pain will be followed by decreases in *predicted* pain; (3) Predicted pain will not change after a correct match; (4) Underpredictions of pain will be followed by increases of *reported* pain; (5) Overpredictions will be followed by decreases in *reported* pain; and (6) Reported pain will not change after a match.

Before testing these hypotheses, it needs to be assessed whether the results of the study are not based on chance findings or statistical artefacts. An increase in subsequent pain predictions after an underprediction, or a decrease after an overprediction, can merely be the result of a regression to the mean effect. In a laboratory study on pain, Arntz et al. have demonstrated that their MM data reflected, besides a significant regression to the mean effect, also a 'real' psychological process[20]. The question is whether this is also the case for daily life data with regard to the MM model of pain.

5.2 Methods

5.2.1 Subjects

From June 2000 to June 2002, all chronic pain patients referred to an inpatient multidisciplinary, cognitive-behavioral rehabilitation program were recruited for participation in the study. The exclusion criteria were: age < 18 years, pain duration < six months, presence of serious psychopathology, lack of fluency in the Dutch language, participation in another study, geographical distance > 100 km. Recruitment was continued until 100 patients with informed consent were acquired. To reach this number, 179 patients had to be invited (a 55.9% response rate). On the basis of the selection criteria, 41 patients were excluded; 38 patients gave no reason for their non-participation. The mean age of the 100 included patients in the study (22 men and 78 women) was 41 years (SD = 10.4) and the mean duration of pain was 6.6 years (SD = 6.3). The local ethical committee approved the study. No significant differences were found in sex, age, pain duration, and pain location between the included and excluded patients (all *P*'s > .118).

5.2.2 Procedures

Data were collected using the Experience Sampling Method (ESM). ESM is a valid and reliable structured diary method. It allows collecting random snapshots of the participant's mental state (including assessments of pain intensity) within their natural environment[21,22]. It minimizes traditional bias of subjective reports by anticipation or retrospection and comes close to a direct in vivo observation of a patient[22,23,24]. The accuracy and applicability of ESM have been demonstrated in a number of studies involving patients with chronic pain[25,26,27,28,29,30].

All patients received a Seiko RC-4000 wristwatch and a set of ESM booklets, each containing the necessary Experience Sampling Forms (ESF's) for one day. The watch randomly signaled participants 10 times a day, between 7:30 a.m. and 10:30 p.m., for a period of 7 days. The time between two signals ranged from 15 minutes to three hours with an average interval of one and a half hour. The patients were instructed to reply to an ESF in the booklet as soon as possible after the auditory signal. The ESF's were 14-item questionnaires (Appendix 5A). They assess current pain and some contextual information, as well as expected activities and related pain levels.

By comparing the actual prompt moments with the log time-entries in the booklets it was possible to discard ESF's that were completed more than 10 minutes after the signal. These responses were considered invalid. Only patients who responded validly to more than 30% of the emitted beeps were included in the analysis. A Monte Carlo simulation by one of the co-authors (PhD) has demonstrated that the likelihood of fooling this protocol retrospectively and still be included as a valid subject in the sample is less than 1%.

5.2.3 Data reduction and analysis

A match is defined when the experienced pain level ($b[experienced]_x$ in Figure 5.1) is equal to the anticipated pain level one beep before ($b[predicted]_{x-1}$ in Figure 5.1). Mismatches can occur when the experienced pain is higher than anticipated in case of an underprediction ($b[experienced]_x > b[predicted]_{x-1}$) or when the experienced pain is lower than anticipated in case of an overprediction ($b[experienced]_x < b[predicted]_{x-1}$). Because the results of a prediction can only be assessed for two consecutive non-missing observations, the last (10th) signal of the day was always discarded leaving a maximum number of 9 observations per day for (mis)match analysis.

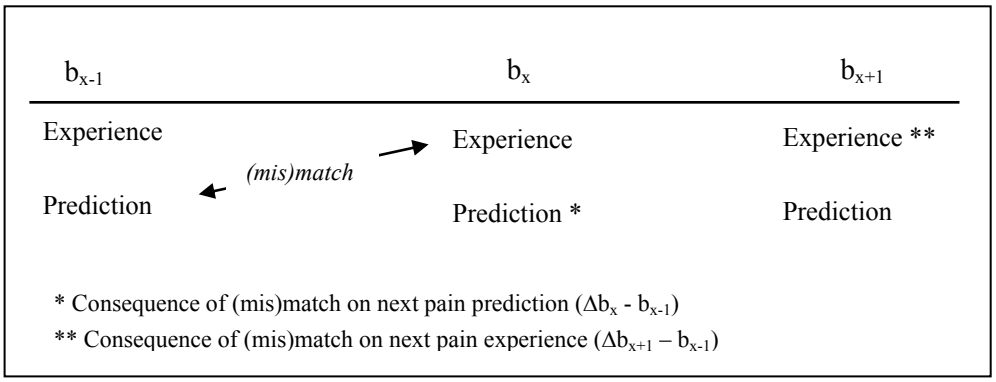


Figure 5.1: Defining a (mis)match and its consequences on subsequent pain experiences and predictions

To code moments as match or mismatch data had to be available about the current and previous beep. As can be viewed in Appendix 5A, predictions and experiences of pain are linked to the ongoing activity. Real (mis)matches should therefore be restricted to pain ratings during activities that were planned during the previous beep. Activities were translated into Metabolic Equivalent (MET) values and corrected for in the analyses[31,32,33].

To assess whether mismatches result in alterations in future predictions and experiences of pain, 3 consecutive non-missing observations were needed (resulting in a maximum of 8 reports for each day, 56 by subject). The consequences of every valid (mis)match on subsequent pain predictions and pain experiences was calculated by counting the number and estimated probabilities of increases, decreases, and no changes following each (mis)match. Further, the mean value of the effects of a (mis)match on pain prediction (Δb_x-b_{x-1} in Figure 5.1) and pain experience (Δb_{x+1}–b_{x-1} in Figure 5.1) was computed by averaging all non-missing scores, first within each person and next over all persons.

The data were analyzed in SPSS 11.5 and STATA 8.0. The data have a hierarchical structure: each patient participated for 70 observations. Following the procedure described by Arntz and Van den Hout, the consequences of (mis)matches on subsequent pain predictions and pain experiences were calculated by counting the number of increases, decreases, and no changes following each (mis)match[5]. The estimated probabilities of these numbers were obtained with multilevel logistic regression taking into account the hierarchical nature of the data.

To determine the mean effect of a (mis)match on subsequent pain experiences and pain predictions multi-level regression models were used[34]. Estimation and testing of the effects for the (non)linear mixed models is based on (restricted) Maximum Likelihood estimation. For these analyses, (mis)matches were recoded as –1 for the underpredictions, 0 for matches and 1 for the overpredictions.

To determine whether the results of the study were not only based on statistical artefacts, a regression model similar as proposed by Arntz et al. was analyzed[20]. The regression equation in their study was as follows:

$$P_{i+1} - P_i = \beta_0 + \beta_1 \times (P_i - M) + \beta_2 \times (P_i - E_i)$$

P = pain prediction; M = mean (individual) prediction level; E = pain experience; i = beep number, $P_i - M$ = regression to the mean effect; $P_i - E_i$ = mismatch effect

Since Arntz et al. did not apply multilevel analysis[20], the variables for the present study were slightly adapted: the dependent variable of the equation is the change in pain prediction $P_{i+1,j} - P_{ij}$ and the exploratory variables a regression to the mean effect ($P_{ij} - P_{.j}$) and the mismatch effect ($P_{ij} - E_{ij}$). The indices i stands for beep number and j for person. $P_{.j}$ is the mean pain prediction over all moments for person j.

5.3 Results

5.3.1 Number of signals and data-pairs

Of the 100 patients, 17 patients did not respond validly to more than 30% of the emitted beeps. These patients were excluded from the analyses, leaving a sample of N = 83. No significant differences were found with regard to sex, age, pain duration, and pain location, between patients who did and patients who did not have sufficiently valid answered signals.

A total number of 4579 valid signals were identified (78.8% of the maximum number of 83 x 7 x 10 = 5810 possible signals), 955 (16.4%) signals were missed and 276 (4.8%) were invalid (i.e., signals not answered within 10 minutes). No significant correlations were found between the number of missing beeps and the variables sex, age, pain duration and pain location (all *P*'s > .187).

For the (mis)match analyses, 3584 valid non-missing consecutive pairs of observations were available (68.5% of 83 x 7 x 9 = 5229). There were 1418

(39.6%) matches and 2166 mismatches; 1476 (41.2%) were underpredictions and 690 (19.2%) were overpredictions. Analysis of missing data revealed that missing data-pairs were randomly distributed with a tendency for more missing data-pairs on the first signal of the day (when patients were still asleep).

5.3.2 Testing the hypotheses

The correlation between experienced pain and actual activity level using MET ratings was significant (Pearson's r = 0.05; $P < 0.001$) but small. The regression models with activity as covariate yielded no significance (overall $R^2 = 0.018$). Therefore, models without correction for activity level are presented.

Figure 5.2 presents the number and estimated probabilities of increases, decreases, and no changes in pain prediction (5.2A) and pain experience (5.2B) following each (mis)match. This figure shows that overpredictions were mainly followed by decreases in predictions on the next prompt and underpredictions mainly by increases and no changes. Further, overpredictions tended to be followed by decreases of experienced pain on the next signal; underpredictions and 'no changes' were followed by no changes or increases in pain.

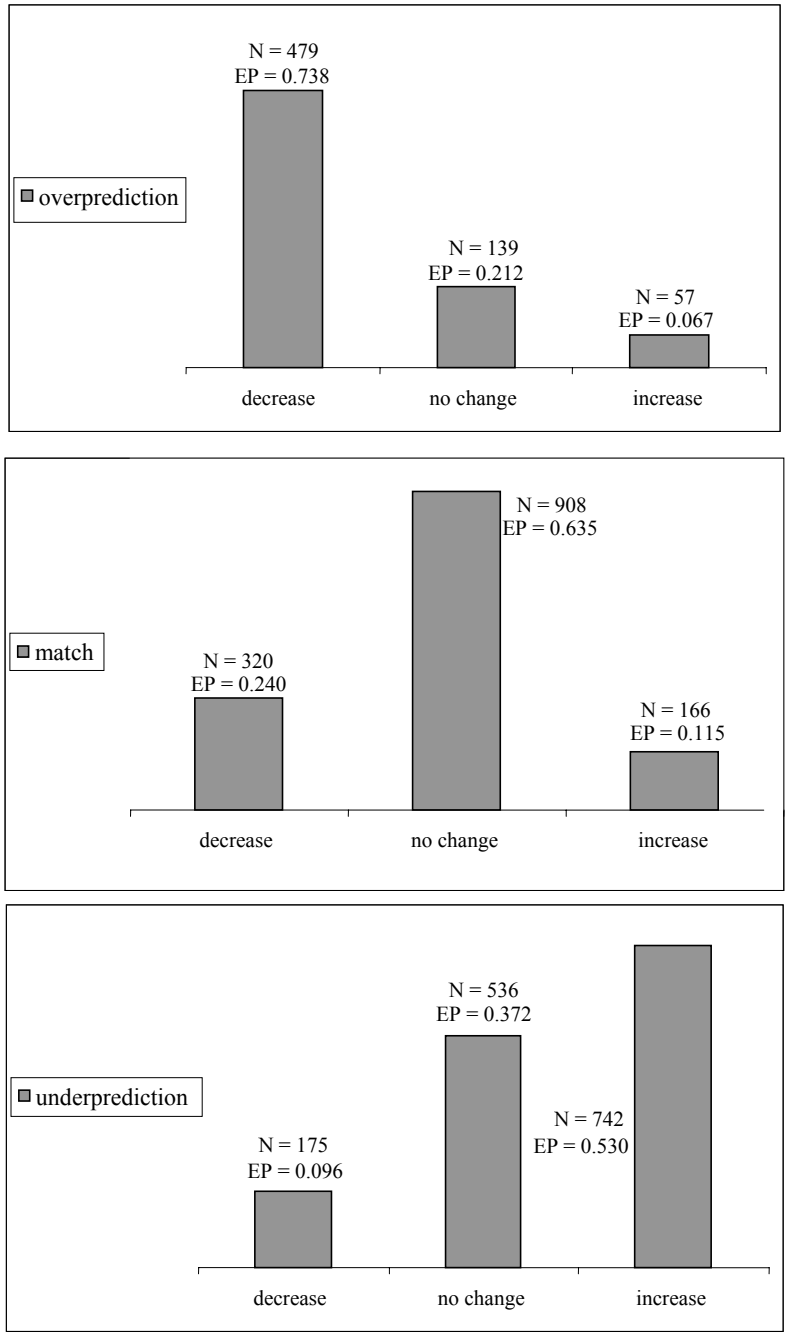


Figure 5.2A: Number of changes in pain prediction after a (mis)match (N=83)
(EP=estimated probability)

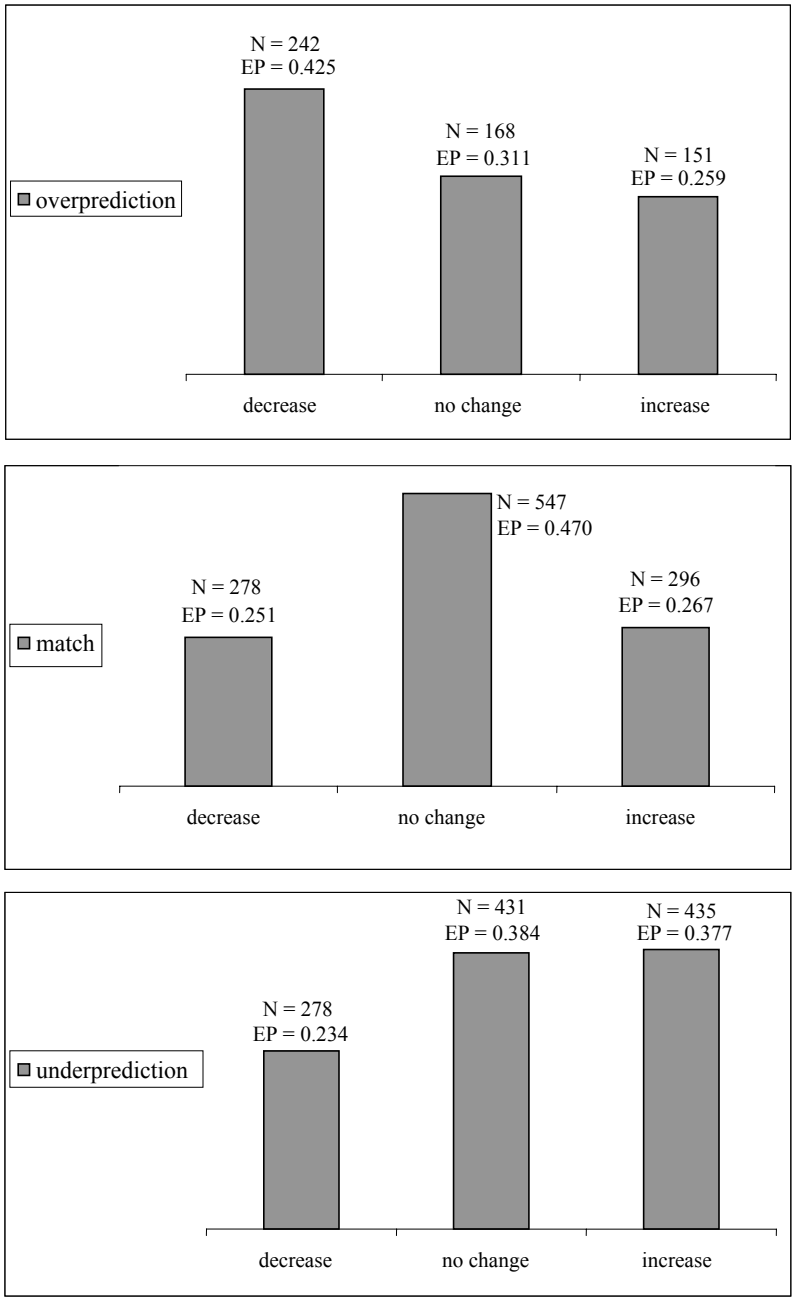


Figure 5.2B: Number of changes in pain experience after a (mis)match (N=83) (EP=estimated probability)

A significant day pattern was found for pain experience (β = 0.09; z = 3.83; $P < 0.001$). The optimal model also included a significant negative quadratic factor (β = -0.006; z = -2.69; $P = 0.007$): pain intensity increases during the day, an increase that was attenuated at the end of a day (Figure 5.3). Further, a significant negative day pattern was found for pain prediction (β = -0.02; z = -2.48; $P = 0.013$): during the day the prediction of pain decreased (Figure 5.3). However, since these factors only explained a very small part of the total moment-to-moment pain variance (overall $R^2 = 0.002$), these factors were not corrected for in subsequent analyses.

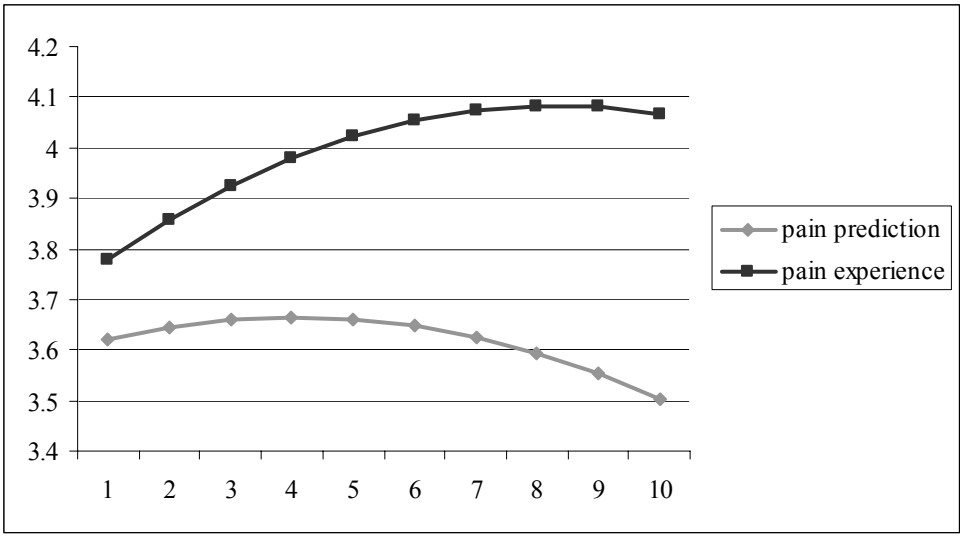


Figure 5.3: Mean effect of pain prediction and pain experience per beep number (N = 83)

Table 5.1 presents the overall changes in pain prediction and pain experience due to a regression to the mean effect as well as both a regression to the mean and mismatch effect. This table shows that the prediction of future pain levels was significantly altered after a (mis)match, even after controlling for the regression to the mean effect ($\chi^2_{(2)} = 2645.95$; $P < .0001$). Furthermore, when controlling for the regression to the mean effect the experience of pain was also significantly altered after a (mis)match ($\chi^2_{(2)} = 85.73$; $P < .0001$).

Table 5.1: Changes in pain prediction and pain experience (N = 83)

	Variables	β	z value	*P*	overall R^2	χ^2 (df)
Effect on pain prediction	Regrmean	-0.70	-44.06	< .001	0.36	1941.58 (1)*
	Regrmean	-0.46	-24.31	< .001	0.43	2645.95 (2)*
	Mismatch	-0.35	-21.32	< .001		
Effect on pain experience	Regrmean	-0.17	-9.18	< .001	0.03	84.27 (1)*
	Regrmean	-0.15	-5.86	< .001	0.03	85.73 (2)*
	Mismatch	-0.05	-1.99	0.047		

Regrmean = regression to the mean effect; mismatch = mismatch effect
* $P < .0001$

In Table 5.2, the mean effect of a (mis)match on future predicted and experienced pain levels is presented, corrected for the regression to the mean effect. This table shows that the alterations in consequent pain predictions and pain experiences are, except for the effect of an overprediction on consequent pain experience, all determined by a significant regression to the mean effect. Besides, Table 5.2 shows that pain prediction significantly increases after an underprediction and significantly decreases after an overprediction. Unexpectedly, after a match pain prediction also decreases significantly. Further, pain experience significantly increases after an underprediction, remains the same after a match, and shows a trend towards a decrease after an overprediction.

Table 5.2: Effect of underpredictions and overpredictions on pain prediction and pain experience (N = 83)

	(Mis)match	Variables	β	z value	*P*	df	χ^2 (df)
Effect on pain prediction	Underprediction	Regrmean	0.64	24.10	< .001	1452	939.20 (1)*
		Mismatch	0.30	7.93	< .001		
	Match	Regrmean	-0.56	-31.97	< .001	3520	2386.15 (2)*
		Mismatch	-0.09	-5.20	< .001		
	Overprediction	Regrmean	-0.58	-12.74	< .001	675	775.91 (1)*
		Mismatch	-0.53	-9.48	< .001		
Effect on pain experience	Underprediction	Regrmean	0.14	4.32	< .001	1143	47.50 (1)*
		Mismatch	0.15	3.23	0.001		
	Match	Regrmean	-0.14	-6.06	< .001	2823	95.95 (2)*
		Mismatch	0.03	1.37	0.170		
	Overprediction	Regrmean	-0.12	-1.94	0.052	561	20.66 (1)*
		Mismatch	-0.13	-1.66	0.096		

Regrmean = regression to the mean effect; mismatch = mismatch effect (underpredictions are coded as –1, matches as 0, and overpredictions as 1)
* $P < .0001$

5.4 Discussion

The primary objective of the present study was to determine the generalizability of the MM model in daily life of chronic pain patients. The results mainly confirm the hypotheses: mean effects of a (mis)match on pain prediction and experience were found, as well as on the number of increases and decreases as predicted by the model. Unexpectedly, a significant mean effect of a match on future pain prediction was found. At this moment, we do not have a clear explanation for this finding. One should keep in mind that the effects of the matches are meaningfully smaller than the effects of mismatches. Further studies are needed to find an explanation for the effect of matches.

A critical issue in this study concerns the question whether the results are based on a regression to the mean effect. In line with the findings of Arntz et al.[20], the present study found an independent significant mismatch effect after controlling for the regression to the mean effect. The latter was also significant. Therefore, it is concluded that the (mis)match effect in the present study is real and not based on some statistical artefect. In comparison with the study of Arntz et al.,

however, the regression to the mean effect in the present study was larger and the mismatch effect smaller. Differences in time interval between measuring moments, research setting (real life versus laboratory), subjects (pain patients versus healthy students), and scales of measurement may explain these differences in study results. Furthermore, in the study of Arntz et al. respondents were instructed to rate levels of pain experience and prediction, i.e. the respondents were more aware of the fact that the study was aimed at testing the match-mismatch model. In the present study, patients were not specifically told that the study was aimed at estimating levels of pain experience and pain predictions.

A second critical issue in this study was the potential confounding effect of (predicted) activity level. This issue is especially relevant in daily life studies where the activity of the event for which pain was predicted might be different from the activity of the event for which pain was rated. In fact, given the random time sampling procedure, it is quite likely that patients predicted pain levels for anticipated future activities that were not performed at the moment of pain rating. Although the correlation between the experienced pain and actual activity level was significant, its predictive contribution to the regression equations was very small. Furthermore, the coefficients of the models did only very slightly change for models that did not control for the activity level. Therefore, it was concluded that for chronic pain patients the mismatch patterns are independent of activity level.

A remarkable finding in this study was the larger number of underpredictions versus overpredictions. This finding is consistent with the results of studies reported by McCracken et al.[17] and Arntz and Peters[18], but is in contrast to other studies where pain patients overpredicted pain[15,16]. Although no definite conclusions can be drawn, strengths of the present study favoring the underprediction mechanism are that the present study was based on a relatively large sample, was executed in the natural environment of patients, and used real life daily data. Additional studies are needed to clarify the observations.

Underpredictions seem to play a pronounced role in the maintenance of chronic pain. As was shown in Table 5.2, the absolute mean effect on pain experience was practically the same for underpredictions and overpredictions (0.15 versus 0.13). Given this and the fact that the number of underpredictions was twice the number of overpredictions (1467 versus 690), the net effect is that pain experience increases during a day. Figure 5.3 shows that pain prediction lies structurally beneath the pain experience level. However, the prediction of pain (in

contrast to the experience) decreases over the day. Consequently, mismatches may be induced by the lower pain prediction in the direction of an underprediction. More research is needed to gain insight into the contextual, emotional and personal factors influencing the patterns in pain prediction and pain experience of chronic pain patients.

An important limitation of using the ESM is that it relies on self-reports. No external check on the validity of the data is available. On the other hand, since ESM collects data in the natural living environment, its results have better external validity compared to research in laboratories. ESM offers the advantage of ecological validity because the data accounts for the patient's own contextual situation, mental state, and behavior[22]. Another limitation of the ESM is directly related to the use of ESM and the prompting 10 times a day. Although ESM-based observations are unaffected by anticipation (because of the random time scheduling of signals), it is possible that the frequency of recording might influence the nature of patients' responses. That is, the repeated questioning might draw the patients' attention to aspects of their lives that they may typically not attend to normally. Higher focus on pain may lead to an increased sensitization or vigilance[35]. Thus, the additional attention might subtly alter the phenomena of interest. Because in this ESM study people have to carry booklets and stop what they are doing 10 times a day to complete the logs, the method is obtrusive. In fact, it might explain part of the missing data. Although there is no reason to believe that the missing data would have altered the results (i.e., it rather can be assumed since the missing data were random) it is possible that the available data presented a skewed if not biased view of the relationship among (mis)matches, predictions, and the experience of pain. Therefore, future research is needed to determine the relationship between these variables in chronic pain populations using ESM methodology.

As a final point of criticism, in paper diaries patients can see previous ratings. This might influence subsequent ratings[29,36]. Using palmtop computers, previous responses can be hided from the subject and this may be preferred in future (mis)match studies. In fact, Jamison et al. have shown that data collection with palmtop computers is more reliable than paper diaries and patients using these palmtop computers had much higher rates of compliance than patients using paper diaries[37]. The discussion is not closed yet and further studies comparing paper and computerized diaries are needed. For the present study, however, because of the

number and contents of questions in the ESF's, as well as the available study budget, the use of paper and pencil was chosen.

In sum, the results of this study largely confirm the role of the MM model in daily life of chronic pain sufferers. In addition, the results provide an indication of the importance of the frequency of underpredictions in daily life of chronic pain patients. Future studies have to be planned to reveal the causes of underpredictions as well as the role of underpredictions in pain maintaining mechanisms in chronic patients.

Acknowledgement

With special thanks to Dr. Philippe Delespaul for his contribution to the analysis and interpretation of data, as well as to Mandy Bergman and Idwine van der Blij for the collection and analysis of data during their internship.

References

1. Feuerstein, M., Papciak, A.S., & Hoon, P.E. (1987). Biobehavioral mechanisms of chronic low back pain. *Clin Psychol Rev*, *7*, 243-273.
2. Turk, D.C., & Flor, H. (1999). Chronic pain: A biobehavioral perspective. In R. Gatchel & D.C. Turk (Eds.). *Psychosocial factors in pain: Critical perspectives*. New York: Guilford Press, p.18-34.
3. Rachman, S., & Lopatka, C. (1986). Match and mismatch in the prediction of fear-I. *Behav Res Ther, 24*, 387-893.
4. Rachman, S., & Lopatka, C. (1986). Match and mismatch of fear in Gray's theory-II. *Behav Res Ther, 24*, 395-401.
5. Arntz, A., & Hout, M.A. van den (1988). Generalizability of the match/mismatch model of fear. *Behav Res Ther,* 26, 207-223.
6. Arntz, A., & Lousberg, R. (1990). The effects of underestimated pain and their relationship to habituation. *Behav Res Ther, 28*, 15-28.
7. Arntz, A., Eck, M. van, Jong, P. de, & Hout, M.A. van den (1990). The relationship between underpredicted pain and escape. *Behav Res Ther, 28*, 87-90.
8. Arntz, A., Eck, M. van, & Jong, P. de (1991). Avoidance of pain of unpredictable intensity. *Behav Res Ther, 29*, 197-201.
9. Arntz, A., Hout, M.A. van den, Berg, G. van den, & Meijboom, A. (1991). The effects of incorrect pain expectations on acquired fear and pain responses. *Behav Res Ther, 29*, 547-560.
10. Rachman, S., & Arntz, A. (1991). The overprediction and underprediction of pain. *Clin Psychol Rev, 11*, 339-355.
11. Crombez, G., Baeyens, F., Vervaet, L., &Vlaeyen, J. (1997). Vermijdingsgedrag bij patiënten met chronische lage rugpijn. *Gedrag Gezond, 25*, 126-135.
12. Arntz, A., & Hopmans, M. (1998). Underpredicted pain disrupts more than correctly predicted pain, but does not hurt more. *Behav Res Ther*, *36*, 1121-1129.
13. Baeyer, C.L. von, Carlson, G., & Webb, L. (1997). Underprediction of pain in children undergoing ear piercing. *Behav Res Ther*, *35*, 399-404.
14. Spafford, P.A., Baeyer, C.L. von, & Hicks, C.L. (2002). Expected and reported pain in children undergoing ear piercing: a randomized trial of preparation by parents. *Behav Res Ther*, *40*, 253-266.
15. Linton, S.J., & Melin, C. (1982). The accuracy of remembering chronic pain. *Pain, 13*, 281-285.
16. Arntz, A., Eck, M. van, & Heijmans, M. (1990). Predictions of dental pain: The fear of any expected evil is worse than the evil itself. *Behav Res Ther, 28*, 29-42.
17. McCracken, L.M., Gross, R.T., Sorg, P.J., & Edmands, T.A. (1993). Prediction of pain in patients with chronic low back pain: effects of inaccurate prediction and pain-related anxiety. *Behav Res Ther, 31*, 647-652.
18. Arntz, A., & Peters, M. (1995). Chronic low back pain and inaccurate predictions of pain: is being too tough a risk factor for the development and maintenance of chronic pain? *Behav Res Ther, 33*, 49-53.
19. Poulton, R., Trevena, J. Reeder, A.I., & Richards, R. (2002). Physical health correlates of overprediction of physical discomfort during exercise. *Behav Res Ther*, *40*, 401-414.

20. Arntz, A., Hout, M.A. van den, Lousberg, R., & Schouten, S. (1990). Case history and shorter communication: Is the match/mismatch model based on a statistical artefact? *Behav Res Ther*, *28*, 249-253.
21. Czikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience sampling method. *J Nerv Ment Dis, 175*, 526-536.
22. Delespaul, P.A.E.G. (1995). *Assessing schizophrenia in daily life: The Experience Sampling Method*. Maastricht: University Press.
23. Vries, M.W. de (1992). *The experience of psychopathology*. Cambridge: Cambridge University Press.
24. Vries, M.W. de, & Delespaul, P.A.E.G. (1993). Zelfobservatie en intensieve time-sampling in de psychiatrie. In C.A.L. Hoogduin, P. Schnabel, W. Vandereycken, K. van der Velden, & F.C. Verhulst (Eds.). *Jaarboek psychiatrie en psychotherapie*. Meppel: Van Loghum Slaterus, p.94-108.
25. Lousberg. R., Schmidt, A.J., Groenman, N.H., Vendrig, L., & Dijkman-Caes, C.I. (1997). Validating the MPI-DLV using experience sampling data. *J Behav Med, 20*, 195-206.
26. Honkoop, P.C., Sorbi, M.J., Godaert, G.L., & Spierings, E.L. (1999). High-density assessment of the HIS classification criteria for migraine without aura: a prospective study. *Cephalalgiam, 19*, 201-206.
27. Peters, M.L., Sorbi, M.J., Kruise, D.A., Kerssens, J.J., Verhaak, P.F.M., & Bensing, J.M. (2000). Electronic diary assessment of pain, disability and psychological adaptation in patients differing in duration of pain. *Pain, 84*, 181-192.
28. Stone, A.A., Shiffman, S., Schwartz, J.E., Broderick, J.E., & Hufford, M.R. (2003). Patient compliance with paper and electronic diaries. *Control Clin Trials, 24*, 182-199.
29. Stone, A.A., Broderick, J.E., Schwartz, J.E., Shiffman, S., Litcher-Kelly, L., & Calvanese, P. (2003). Intensive momentary reporting of pain with an electronic diary: reactivity, compliance, and patient satisfaction. *Pain*, *104*, 343-51.
30. Stone, A.A., Broderick, J.E., Shiffman, S., & Schwartz, J.E. (2004). Understanding recall of weekly pain from a momentary assessment perspective: absolute agreement, between- and within-person consistency, and judged change in weekly pain. *Pain*, *107*, 61-69.
31. Willmore, J.H., & Costill, D.L. (1994). *Physiology of sport and exercise*. Champaign (IL), USA: Human Kinetics.
32. Ainsworth, B.E., Haskell, W.L., Whitt, M.C., Irwin, M.L., Swartz, A.M., Strath, S.J., O'Brien, W.L., Bassett D.R., Schmitz, K.H., Emplaincourt, P.O., Jacobs, D.R., & Leon, A.S. (2000). Compendium of physical activities: an update of activity codes and MET intensities. *Med Sci Sports Exerc*, S498-S516.
33. Ainsworth, B.E. (2002). *The compendium of physical activities tracking guide*. Prevention Research Center, Norman J. Arnold School of Public Health, University of South Carolina. Retrieved [27-09-2004] from the world wide web http://prevention.sph.sc.edu/tools/docs/documents_compendium.pdf.
34. Snijders, T.A.B., Bosker, R.J. (1999). *Multilevel analysis: an introduction to basic and advanced multilevel modeling*. London (etc.): Sage.
35. Arntz, A., Dreessen, L., & Merckelbach, H. (1991). Attention, not anxiety, influences pain. *Behav Res Ther, 29*, 41-50.

36. Stone, A.A., Broderick, J.E., Porter, L.S., & Kaell, A.T. (1997) The experience of rheumatoid arthritis pain and fatigue: examining momentary reports and correlates over one week. *Arthritis Care Res, 10*, 185-192.
37. Jamison, R.N., Raymond, S.A., Levine, J.G., Slawsby, E.A., Nedeljkovic, S.S., & Katz. N.P. (2001) Electronic diaries for monitoring chronic pain: 1-year validation study. *Pain*, *91*, 277-285.

Appendix 5A: Contents of the experience sampling form (ESF)

1. Severity of the pain:

0 1 2 3 4 5 6 7
no pain very much pain

2. Interference of pain

0 1 2 3 4 5 6 7
no very much
interference interference

3. Pain control:

0 1 2 3 4 5 6 7
no control very much control

4. Where are you now? ……..…

5. With whom are you now? ……..……

6. What are you doing at this moment?
…….

7. How intensive do you think this activity is?

0 1 2 3 4 5 6 7
absolutely not very
intensive intensive

8. How strained are you during this activity?

0 1 2 3 4 5 6 7
absolutely not very
strained strained

9. What are you going to do immediately?
……..……...….

10. How skilled are you in performing this activity?

0 1 2 3 4 5 6 7
absolutely not very skilled
skilled

11. How much pain do you expect to feel by doing this activity?

0 1 2 3 4 5 6 7
no pain very much pain

12. How well do you think you'll be able to cope with
this pain?

0 1 2 3 4 5 6 7
absolutely not very well

13. This beep disturbed me

0 1 2 3 4 5 6 7
not very much

14. What is the time now? …. hour ... minutes

CHAPTER 6

The influence of cognitive-behavior therapy on the match-mismatch mechanism in chronic pain patients: an experience sampling study

Adapted from: C.A.J. Mes, R. Lousberg, G. Zilvold, P.A.E.G. Delespaul, D.C. Turk. The influence of cognitive behavioral treatment on the match-mismatch mechanism in chronic pain patients: an experience sampling study (submitted).

6.1 Introduction

Several models have been proposed to explain the development and maintenance of chronic pain and related disability. The Match-Mismatch (MM) model of pain[1,2], based on the MM model of fear developed by Rachman and Lopatka[3,4], states that the most important immediate consequence of a mismatch between expected and experienced intensity of a painful event is that the expectation for the next experience is adjusted in the direction of the most immediate prior experience. When the event is more painful than anticipated (i.e., an underprediction), the expectation for the next event is predicted to be greater. Conversely, when the preceding event is less painful than anticipated (i.e., an overprediction), the expectation is adjusted so that the anticipated pain severity is lowered. When the expectation for the anticipated pain severity is accurate (i.e., a match) no changes in future pain expectations should occur.

Mes et al. demonstrated that the mechanisms of the MM model studied mainly in laboratory situations can be applied to daily life situations of chronic pain patients (Chapter 5)[5]. The results of this study indicated that chronic pain patients more often underpredict than overpredict their pain and that these underpredictions seem to play a crucial role in the maintenance of chronic pain. These results were in line with several other studies, demonstrating that underpredictions are related to an increase of fear of pain, fear of movement, as well as to escape and avoidance behavior[6,7,8].

There are indications that the experienced level of pain remains heightened for a longer period of time after an underprediction. In addition, underpredictions result in long-term cognitive effects such as increased pain expectations as well as a long-term uncertainty about these expectations[2]. For these reasons, changing pain expectations of patients who are inclined to underpredict their pain may be clinically relevant since escape and avoidance behavior, resulting from repeated underpredictions of pain, may eventually instigate a disuse syndrome. Further, escape and avoidance behavior may prevent patients from testing their predictions and from adjusting them in case the predictions prove to be inaccurate[9]. Therefore, specific attention should be given to the (mis)match mechanism during chronic pain treatment.

In general, there is consensus about the positive effects of cognitive-behavior therapy (CBT) for example on pain severity, mood, cognitive coping, pain

behavior, activity levels, and social role functioning[10,11]. CBT attempts to address the psychosocial and behavioral as well as the physical contributors to chronic pain[12]. However, it is unclear what exact influence CBT has on the MM mechanisms of chronic pain patients, since no attention is specifically paid to this mechanism. The present study is part of an ongoing program of research about the MM mechanism of pain and builds on a previous study (Chapter 5) to expand the understanding about this mechanism. Specifically, changes with regard to (mis)matches as a result of CBT for chronic pain were investigated. It was hypothesized that the emphasis of CBT on helping patients to understand the role of their beliefs and expectations might, although not directly targeted, reduce the occurrence of mismatches. If this hypothesis is correct, then the number and size of mismatches, as well as the effect of a mismatch on the next pain experiences and pain predictions should be reduced following treatment. The following hypotheses were tested in this study:

(1) Following CBT, the number of mismatches will decrease in favor of the number of matches;

(2) After CBT, the size of mismatches will decrease compared to the size of the mismatches that occur prior to treatment; and

(3) After CBT, the effect of mismatches on consequent pain experiences and pain expectations will decrease compared to the effect of mismatches before treatment.

6.2 Methods

6.2.1 Participants

From June 2000 to June 2002, chronic pain patients referred to an inpatient multidisciplinary, cognitive-behavioral rehabilitation program were recruited for participation in the study. The exclusion criteria were: age < 18 years, pain duration < six months, presence of serious psychopathology, lack of fluency in the Dutch language, geographical distance > 100 km and no participation in randomized clinical trial that was performed at the same time. Recruitment was continued until 100 patients with valid data were acquired. To reach this number, 179 patients had to be invited (a 55.9% response rate). On the basis of the selection criteria, 41 patients were excluded; 38 patients gave no reason for their non-participation. Unfortunately, data from 19 of the 100 included patients had to be excluded from analysis, since these patients were unavailable at follow-up. No differences were found between these 19 ‘unavailable’ patients and the remaining

81 patients with respect to age, sex, pain duration, and pain level at baseline. The sample consisted of 20 men (mean age 42.9 (SD = 9.1)) and 61 women (mean age 39.5 (SD = 9.9) with divers pain diagnoses. The mean duration of pain was 6.4 years (SD = 6.0). All participants gave their written informed consent. This study was approved by the local ethical committee.

Fifty-five of the 81 patients participated in the ESM measurement at t_{post}. The other 26 patients terminated prematurely because they did not participate in all measurement days at t_{pre} (15.4%), or found the measurements too strenuous (23.1%); were "too busy" (15.4%); became pregnant (7.7%), had no transportation (3.8%); did not want to be confronted with pain all the time (7.7%); dropped out from treatment (7.7%); or gave no reason (19.2%). Of the remaining 55 respondents, 5 respondents had excessive missing data at t_{pre} and were excluded from the analysis. Further, of the remaining 50 respondents, 7 had excessive data at t_{post} and were also excluded from the analysis. No differences were found between completers and drop-outs with regard to sex, age, pain level at t_{pre}, duration of pain, and subgroup classification as measured at t_{pre} with the Multidimensional Pain Inventory - Dutch Language Version (MPI-DLV)[13]. For this study only results of patients who performed both t_{pre} and t_{post} were included (N = 43).

6.2.2 Data collection

Data were collected using the Experience Sampling Method (ESM). ESM is a valid and reliable structured diary method. It allows collecting random snapshots of the participant's mental state (including assessments of pain intensity) within their natural environment[14,15]. It minimizes traditional bias of subjective reports by anticipation or retrospection and comes close to a direct in vivo observation of a patient[15,16,17]. The accuracy and applicability of ESM have been demonstrated in a number of studies involving patients with chronic pain[18,19,20].

The ESM took place approximately two weeks prior to the initiation of CBT treatment (t_{pre}) as well as four months after treatment termination (t_{post}). For both measurements, all patients received a Seiko RC-4000 wristwatch and a set of ESM booklets, each containing the necessary Experience Sampling Forms (ESFs) for one day. The ESFs consisted of a 14-item questionnaire (Appendix 6A). They assess current pain and some contextual information, as well as expected activities and related pain levels. The watch randomly prompted participants 10 times a day, between 7:30 a.m. and 10:30 p.m., for a period of 7 days. The minimum time between two signals was 15 minutes; the maximum time was three hours with an

average interval of one and a half hour. As soon as possible after the auditory signal, the patient was instructed to reply to an ESF in the booklet. By comparing the actual prompt moments with the log time-entries in the booklets it was possible to discard ESFs that were completed more than 10 minutes after the signal. These responses were considered invalid. Only patients who responded validly to more than 30% of the emitted prompts were included in the analysis. A Monte Carlo simulation by one of the co-authors (PhD) has demonstrated that the likelihood of fooling this protocol retrospectively and still be included as a valid participant in the sample is less than 1%.

Treatment effects of the CBT program were based on responses to the MPI-DLV[13] as well as by corresponding questions measured with the ESFs. The MPI-DLV is a comprehensive self-report measure that assesses pain, the impact of pain, responses by significant others, and general activities. The instrument has been shown to be reliable and valid for use with diverse samples of chronic pain[21]. In the present study, only three of the MPI-DLV scales were used - Pain Severity, Perceived Control, and Interference.

6.2.3 Data reduction and analysis

A match is defined when the experienced pain level ($b[experienced]_x$ in Figure 6.1) is equal to the anticipated pain level one beep before ($b[predicted]_{x-1}$ in Figure 6.1). Mismatches can occur when the experienced pain is higher than anticipated in case of an underprediction ($b[experienced]_x > b[predicted]_{x-1}$) or when the experienced pain is lower than anticipated in case of an overprediction ($b[experienced]_x < b[predicted]_{x-1}$). Because the results of a prediction can only be assessed for two consecutive non-missing observations, the last (10th) signal of the day was always discarded leaving a maximum number of 9 observations per day for (mis)match analysis.

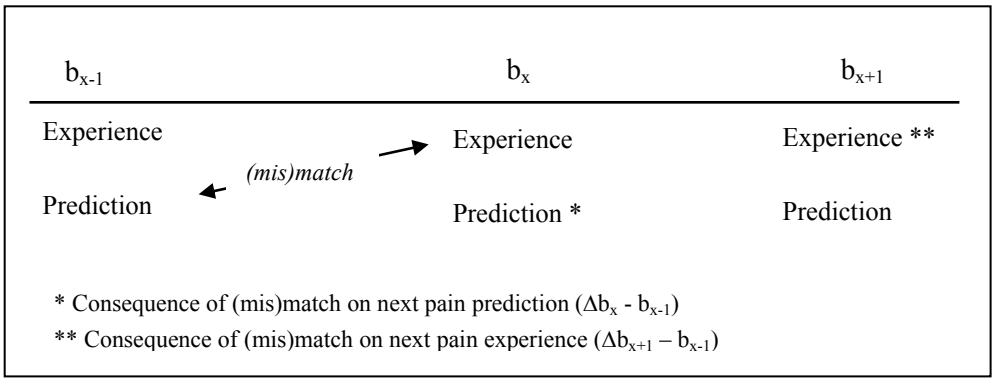


Figure 6.1: Defining a (mis)match and its consequences on pain prediction and pain experience

As can be observed in Appendix 6A, predictions and experiences of pain are linked to the ongoing activity. Real (mis)matches should therefore be restricted to pain ratings during activities that were planned during the previous signal. Activities were translated into Metabolic Equivalent (MET) values and corrected for in the analyses[22,23].

Moments were coded as match or mismatch. To assess whether mismatches result in alterations in future predictions and experiences of pain, 3 consecutive non-missing observations were needed (resulting in a maximum of 7x8 valid reports for each participant). For every valid (mis)match, the consequent effects on pain prediction ($\Delta b[predicted]_x$-$b[predicted]_{x-1}$) and pain experience ($\Delta b[experienced]_{x+1}$–$b[experienced]_{x-1}$) were calculated (Figure 6.1). The mean values of these effects were computed by averaging all non-missing scores, first within each person and next over all persons.

The analyses were performed in SPSS 11.5 as well as STATA 8.0. The data have a hierarchical structure: each patient contributed observations over two periods of 7 days, 10 times each day. The days and assessment periods were not treated as separate levels, but as attributes of the prompt. Thus, the analyses were performed only on two levels: participant and signal. To determine the mean effect of a (mis)match on subsequent pain experiences and pain predictions multi-level regression models were used[24]. Estimation and testing of the effects for the (non)linear mixed models is based on (restricted) Maximum Likelihood estimation. For these analyses, the (mis)matches were recoded in two orthogonal dummy variables (resulting in separate estimations for underpredictions, matches and

overpredictions). When, for example, determining the difference between pre-treatment and post-treatment in effect of an underprediction on pain experience, the analysis is corrected for the effect of overpredictions and matches as well. Further, all analyses were corrected for present level of pain experience as well as for a regression to the mean effect. This is a statistical artifact and no real psychological process. That is, after a low score the next rating tends to increase and after a high score ratings decrease. Finally, since it may be expected that treatment effects interact with mismatch effects, also interaction effects between the moment of measurement and the effect of mismatches were calculated.

6.3 Results

Before testing the specific hypotheses, two basic analyses were performed: 1) determination of the effect of treatment based on the pain severity, interference and control scales of the MPI-DLV, and 2) determination of the number of valid signals.

6.3.1 The effect of the CBT

The effect of the CBT treatment was determined on three key variables of the MPI-DLV -- Pain Severity, Pain Control and Pain Interference. Since Lousberg et al. demonstrated significant correlations between these three MPI-DLV scales and corresponding ESM items[25], it was expected to find comparable treatment effects for the ESM measures in the present study. In order to determine the effect of treatment, the mean values per subject over all signals on t_{pre} and t_{post} were calculated and compared with a t-test. Table 6.1 shows the treatment results on corresponding MPI-DLV and ESM measures. As can be seen, the MPI-DLV results show a significant decrease in pain severity and pain interference. Further, a trend towards an increase of perceived control was demonstrated. With regard to the ESM results, for all three variables a significant improvement was found. The analyses were repeated using multi-level regression techniques. Similar results were found confirming the initial analyses. Based on these results, it was concluded that the treatment has a positive effect.

Table 6.1: Differences between variables on t_{pre} and t_{post} (N = 43)

	MPI-DLV Measurement			ESM Measurement		
	M (SD) t_{pre}	M (SD) t_{post}	Δ t_{pre}-t_{post} (t value, (*P* value)	M (SD) t_{pre}	M (SD) t_{post}	Δ t_{pre}-t_{post} (t value, (*P* value)
Pain Severity	4.10 (0.90)	3.53 (1.17)	3.72 (.001)	4.08 (0.85)	3.63 (1.19)	2.79 (.008)
Pain Control	3.51 (1.14)	3.97 (1.42)	1.87 (.068)	3.57 (1.23)	4.12 (1.26)	-5.04 (<.001)
Pain Interference	4.42 (0.83)	4.23 (0.73)	-2.72 (.009)	3.49 (0.87)	3.12 (1.03)	2.20 (.033)

6.3.2 Number of Signals

At t_{pre}, a total number of 2449 valid signals for (mis)match analysis were identified (81.4% of the maximum number of 3010 possible signals), 459 (15.2%) signals were missed and 102 (3.4%) signals were invalid (i.e., signals not answered within 10 minutes). At t_{post}, a total number of 2342 valid signals for (mis)match analysis was defined (77.8% of the maximum number of 3010 possible signals), 557 (18.5%) signals were missed and 111 (3.7%) signals were invalid. The relative number of missed and invalid signals together significantly increased between t_{pre} and t_{post} with 3.6% ($t_{(42)} = 2.42$; $P = .020$).

6.3.3 Number and size of data-pairs / (mis)matches

At t_{pre}, 1925 valid (i.e., non-missing) consecutive data-pairs of observations (71.1% of 2709, i.c. 43 patients for 7 days, 9 pairs per day) were available for (mis)match analysis. There were 767 (39.9%) matches and 1158 mismatches; 807 (41.9%) were underpredictions and 351 (18.2%) were overpredictions. At t_{post}, 1838 valid consecutive data-pairs (67.8% of the maximum number of 2709 possible data-pairs) were available for (mis)match analysis. There were 917 (49.9%) were matches and 921 mismatches: 644 (35.0%) were underpredictions and 277 (15.1%) were overpredictions. The number of occasions when a specific size of a (mis)match occurs before as well as after treatment is presented as a percentage of the total number of valid (mis)matches in Figure 6.2. The negative values of the mismatches represent the underpredictions, the zero value indicates the number of matches and the positive values represent the overpredictions.

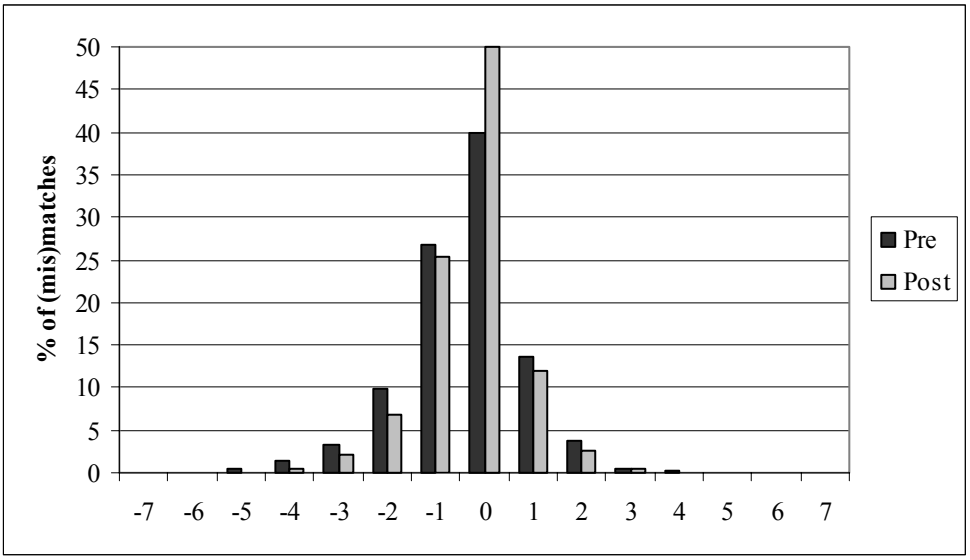


Figure 6.2: The occurrence of (mis)matches (N = 43)

Results of a chi-square test revealed that, although the proportion of underpredictions and overpredictions remained relatively the same, there was a significant change in the proportion of matches versus mismatches in favor of the number of matches ($\chi^2 = 38.40$; $P < 0.001$).

Analysis of missing data indicated that for t_{pre} as well as t_{post}, missing data-pairs were randomly distributed with a tendency for more missing data-pairs on the first signal of the day (patients still asleep). At t_{post}, the percentage of missing data-pairs (32.2%) was not significantly larger than the percentage of missing data-pairs at t_{pre} (28.9%).

Although the correlation between the experienced pain and actual activity level (as measured using the MET scores) was found to be significant (Pearson's r = 0.05; $P < 0.001$), its predictive contribution to the regression equations was small. The following regression models were corrected for activity level, but the contribution of the actual activity level was not significant in all analyzed models. Furthermore, the coefficients of the models did only very slightly change after removing the activity level. Therefore, only results of the models without the correction for activity level are presented.

Table 6.2 presents the mean sizes of underpredictions and overpredictions at t_{pre} and t_{post}. Results of multilevel regression analysis revealed that the size of the underpredictions significantly decreased after treatment; there was only a trend towards a significant decrease of the size of the overpredictions. Further, these results indicated that the effect of treatment was twice as large for the underpredictions as for the overpredictions.

Table 6. 2: Mean size of mismatches (N = 43)

Mismatch	**Mean size at t_{pre} (SD)**	**Mean size at t_{post} (SD)**	**Δ t_{pre} - t_{post}**
Underprediction	1.55 (0.89)	1.37 (0.68)	β = 0.18 (z = 4.35; *P* < .001)
Overprediction	1.31 (0.62)	1.24 (0.50)	β = -0.08 (z = -1.66; *P* = .096)

6.3.4 Effect of a mismatch

In Table 6.3, the results of the regression analyses of the mismatch related differences between t_{pre} and t_{post} are presented, corrected for the regression to the mean effect as well as the experienced level of pain. The mismatch effect is significant overall: both overpredictions and underpredictions (only for pain prediction) resulted in changes in pain experience as well as later pain prediction.

Table 6.3: Regression analysis of changes in effect of (mis)matches, corrected for regression to the mean effect and level of experienced pain (N = 43)

	Effect on pain prediction				**Effect on pain experience**			
	β	z	*P*	$\chi2$ (df) (*P*)	β	z	*P*	$\chi2$ (df) (*P*)
Treatment effect	-0.12	-3.43	0.001	2125.73 (7) (<.0001)	-0.18	-4.79	<.001	1102.22 (7) (<.0001)
Mismatch effect from underprediction	0.31	4.20	<.001		0.13	1.67	0.095	
Interaction effect underprediction x treatment	0.03	0.75	0.454		0.11	2.23	0.026	
Mismatch effect from overprediction	-0.26	-2.73	0.006		-0.25	-2.57	0.010	
Interaction effect overprediction x treatment	-0.10	-1.64	0.102		-0.01	-0.20	0.845	

After treatment, mismatch related changes were significantly reduced on both pain prediction and pain experience (Table 6.3). The interaction effect was significant

for underpredictions, thereby reducing the treatment effect in pain experience (Figure 6.3).

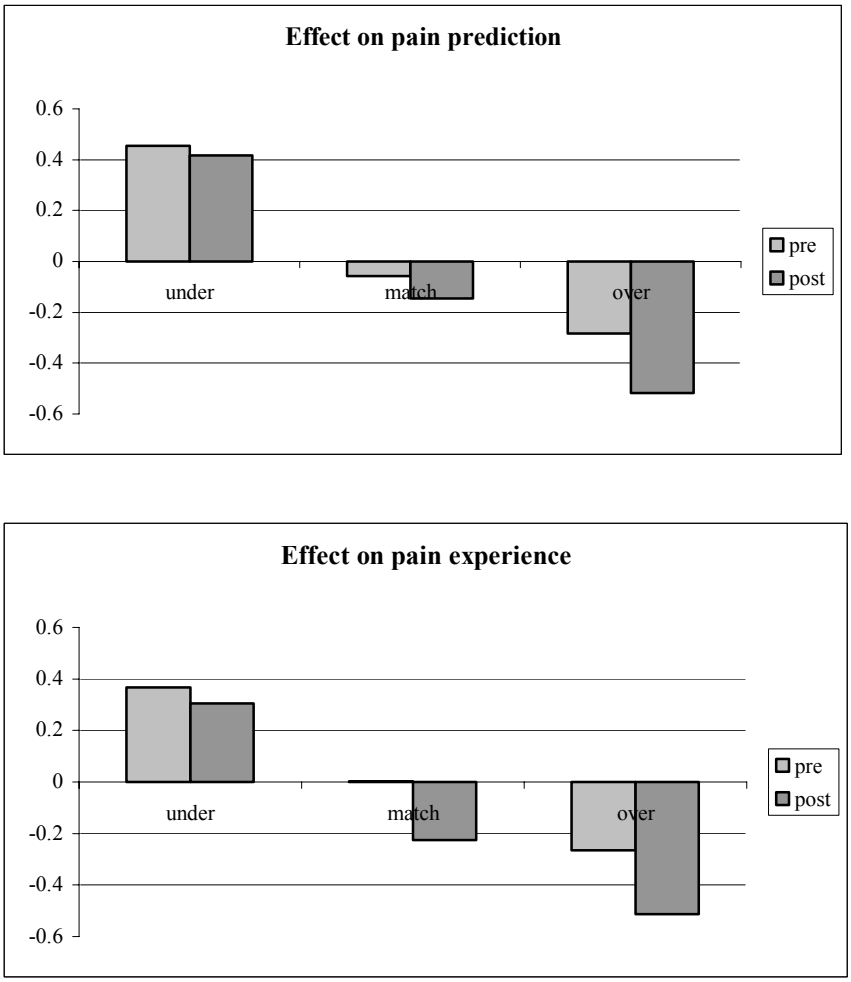


Figure 6.3: The effect of (mis)matches on subsequent pain predictions and pain experiences before and after treatment for an average level of pain experience (score 4) and corrected for regression to the mean effect (for the situation were the pain level equals patients' mean score on pain prediction). For the appropriate significant levels is refered to Table 6.3.

6.4 Discussion

Between t_{post} and t_{pre}, the number of underpredictions had significantly decreased and the number of matches had significantly increased. Also, the size of underpredictions proved to be significantly smaller at t_{post}. For the size of overpredictions, only a trend towards a decrease was found after treatment. These results confirm the first and partially the second hypothesis. Further, the effect of mismatches on consequent pain predictions and pain experiences significantly decreased after treatment. These results confirm the third hypothesis. Although there was no explicit attention for the MM mechanism during treatment, the data show a genuine change in pain self-monitoring and evaluation. Thus, the multidisciplinary CBT program appears to have had a positive, indirect influence on the number, size and effect of mismatches in chronic pain patients, and, for that matter, also on the mechanisms of pain assessment and chronification. One should bear in mind that the data were collected using the ESM. Participants did not actually have feedback on their previous ratings when rating actual pain levels. They were, therefore, unaware of the fact that their current rating was a match or mismatch. Moreover, the interval between ratings was between 15 minutes and 3 hours. With such intervals, most patients probably did not know the actual situation they would be engaged in during the next rating. We believe that finding such an impact and alterations in this process of moment-to-moment pain monitoring after an a-specific treatment, is a powerful demonstration of the relevance of the match-mismatch theory for daily life experience of chronic pain patients.

The results of this study indicate that chronic pain patients are more inclined to underpredict than overpredict their pain. These results confirm previous findings[8,26]. This may explain the fact that the decrease of the total number of mismatches is almost completely accounted for by a decrease of underpredictions, as well as by the fact that the treatment had a larger effect on the mean size of the underpredictions. Underpredictions of pain seem to have a more detrimental effect than correctly predicted or overpredicted pain[6,27,28]. Therefore, CBT might be expected to have a larger effect on the occurrence of underpredictions as compared to matches and overpredictions. The results of the present study demonstrated that indeed this was the case.

The observation that chronic pain patients more often underpredict than overpredict their pain suggests that underpredictions and their consequences seem to play a crucial role in the maintenance of chronic pain. Therefore, treatment

should focus more explicitly on the MM mechanism in chronic pain patients, in particular on the underpredictions. Additional research is needed to explain the etiology of (mis)matches. Factors such as social context and (pain) cognitions may be responsible for the occurrence of (mis)matches and should be studied in more detail.

Several limitations of this study should be acknowledged. First, the effectiveness of CBT with regard to the match-mismatch mechanism in chronic pain patients was demonstrated in relation with three key variables as measured with the MPI-DLV. However, these positive results have to be interpreted with caution since no control group was included for comparison. This makes it difficult to attribute the findings to the treatment program per se. On the other hand, extensive literature is available demonstrating that CBT is an effective treatment and similar findings are demonstrated[10,29]. Therefore, it may be assumed that the findings in the present study are likely due to the CBT program.

Second, data on only 55 chronic pain patients were available for analysis, whereas 81 patients were measured at t_{pre}. It might be argued that a selection bias may have distorted the results. However, the general characteristics of the patients who terminated treatment prematurely with regard to sex, age, pain level at $t_{pre,}$ and duration of pain did not differ from the participating patients. Consequently, it is plausible that the results of this study are generalizable to the total sample of patients.

A third limitation concerns the use of ESM and the prompting 10 times a day. It is possible that the frequency of recording might draw the patients' attention to aspects of their lives that they may typically not attend, or to their pain which may lead to an increased sensitization or vigilance[30]. Further, the method — carrying around booklets and stop what you are doing 10 times a day — may be more obtrusive than anticipated. In fact, it might explain part of the amount of missing data. Although there is no reason to believe that the missing data would have altered the results (since the missing data were random), it is possible that the available data presented a skewed if not biased view of the relationship among (mis)matches, predictions, and the experience of pain. Future research is needed to investigate the relationship between these variables in chronic pain populations.

As a final point of criticism, it might be argued that paper diaries allow patients to see previous ratings that might influence subsequent ratings[31,32]. Using palmtop computers with programmed log out of previous responses may be

preferred in future studies. In fact, Jamison et al. have shown that data collection with palmtop computers is more reliable than paper diaries and patients using these palmtop computers had much higher rates of adherence than patients using paper diaries[33].

In sum, CBT had a positive influence on the number, size and effect of (mis)matches. In addition, the present study gives an indication of the importance of the frequency of underpredictions in daily life of chronic pain patients. Future investigations have to be performed to reveal the causes of underpredictions as well as the role of underpredictions in pain maintaining mechanisms.

Acknowledgement

With special thanks to Dr. Philippe Delespaul for his contribution to the analysis and interpretation of data, as well as to Mandy Bergman and Idwine van der Blij for the collection and analysis of data during their internship.

References

1. Arntz, A., & Van den Hout, A. (1988). Generalizability of the match/mismatch model of fear. *Behav Res Ther, 26*, 207-223.
2. Arntz, A., & Lousberg, R. (1990). The effects of underestimated pain and their relationship to habituation. *Behav Res Ther, 28*, 15-28.
3. Rachman, S., & Lopatka, C. (1986). Match and mismatch in the prediction of fear-I. *Behav Res Ther, 24*, 387-893.
4. Rachman, S., & Lopatka, C. (1986). Match and mismatch of fear in Gray's theory-II. *Behav Res Ther, 24*, 395-401.
5. Mes, C.A.J., Lousberg, R., Zilvold, G., Delespaul, P.A.E.G., & Turk, D.C. (submitted). Pain expectations and experiences: replicating the match-mismatch model of pain in daily life.
6. Arntz, A., Van den Hout, A., Van den Berg, G., & Meijboom, A. (1991). The effects of incorrect pain expectations on acquired fear and pain responses. *Behav Res Ther, 29*, 547-560.
7. Crombez, G., Baeyens, F., Vervaet, L., & Vlaeyen, J. (1997). Vermijdingsgedrag bij patiënten met chronische lage rugpijn. *Gedrag Gezond 25*, 126-135.
8. McCracken, L.M., Gross, R.T., Sorg, P.J., & Edmands, T.A. (1993). Prediction of pain in patients with chronic low back pain: effects of inaccurate prediction and pain-related anxiety. *Behav Res Ther, 31*, 647-652.
9. Arntz, A., Van Eck, M., De Jong, P., & Van den Hout, M.A. (1990). The relationship between underpredicted pain and escape. *Behav Res Ther, 28*, 87-90.
10. Morley, S., Eccleston, C., & Williams, A (1999). Systematic review and meta-analysis of randomized controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache. *Pain, 80*, 1-13.
11. Turk, D.C., & Okifuji, A. (2002). Psychological factors in chronic pain: Evolution and revolution. *Consult Clin Psychol, 70*, 278-290.
12. Turk, D.C., Meichenbaum, D., & Genest, M. (1983). *Pain and behavioral medicine: A cognitive-behavioral perspective*. New York: Guilford Press.
13. Lousberg, R., Breukelen, G.J.P., van, Groenman, N., Schmidt, A.J.M., Arntz, A., & Winter, F.A.M. (1999). Psychometric properties of the Multidimensional Pain Inventory, Dutch language version (MPI-DLV). *Behav Res Ther, 37*, 167-182.
14. Czikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience sampling method. *J Nerv Ment Dis, 175*, 526-536.
15. Delespaul, P.A.E.G. (1995). *Assessing schizophrenia in daily life: the Experience Sampling Method*. Maastricht: University Press.
16. Vries, M.W. de (1992). *The experience of psychopathology*. Cambridge: Cambridge University Press.
17. Vries, M.W. de, Delespaul, P.A.E.G. (1993). Zelfobservatie en intensieve time-sampling in de psychiatrie. In C.A.L. Hoogduin, P. Schnabel, W. Vandereycken, K. van der Velden, & F. Verhulst (Eds.). *Jaarboek psychiatrie en psychotherapie*. Meppel: Van Loghum Slaterus, p. 94 - 108.

18. Peters, M.L., Sorbi, M.J., Kruise, D.A., Kerssens, J.J., Verhaak, P.F.M., & Bensing, J.M. (2000). Electronic diary assessment of pain, disability and psychological adaptation in patients differing in duration of pain. *Pain, 84*, 181-192.
19. Stone, A.A., Broderick, J.E., Schwartz, J.E., Shiffman, S., Litcher-Kelly, L., & Calvanese, P. (2003). Intensive momentary reporting of pain with an electronic diary: reactivity, compliance, and patient satisfaction. *Pain, 104*, 343-351.
20. Stone, A.A., Broderick, J.E., Shiffman, S., & Schwartz, J.E. (2004). Understanding recall of weekly pain from a momentary assessment perspective: absolute agreement, between- and within-person consistency, and judged change in weekly pain. *Pain*, *107*, 61-69.
21. Turk, D.C., & Rudy, T.E. (1990). The robustness of an empirically derived taxonomy of chronic pain patients. *Pain, 43*, 27-35.
22. Ainsworth, B.E., Haskell, W.L., Whitt, M.C., Irwin, M.L., Swartz, A.M., Strath, S.J., O'Brien, W.L., Bassett D.R., Schmitz, K.H., Emplaincourt, P.O., Jacobs, D.R., & Leon, A.S. (2000). Compendium of physical activities: an update of activity codes and MET intensities. *Med Sci Sports Exerc*, S498-S516.
23. Ainsworth, B.E. (2002). *The compendium of physical activities tracking guide*. Prevention Research Center, Norman J. Arnold School of Public Health, University of South Carolina. Retrieved [27-09-2004] from the world wide web http://prevention.sph.sc.edu/tools/docs/documents_compendium.pdf.
24. Snijders, T.A.B., & Bosker, R.J. (1999). Multilevel analysis: an introduction to basic and advanced multilevel modeling. London (etc.): Sage.
25. Lousberg, R., Schmidt, A.J., Groenman, N.H., Vendrig, L., & Dijkman-Caes, C.I. (1997). Validating the MPI-DLV using experience sampling data. *J Behav Med, 20*, 195-206.
26. Arntz, A., & Peters, M. (1995). Chronic low back pain and inaccurate predictions of pain: is being too tough a risk factor for the development and maintenance of chronic pain? *Behav Res Ther*, 33, 49-53.
27. Arntz, A., Van Eck, M., & De Jong, P. (1991). Avoidance of pain of unpredictable intensity. *Behav Res Ther, 29*, 197-201.
28. Rachman, S., & Arntz, A. (1991). The overprediction and underprediction of pain. *Clin Psychol Rev, 11*, 339-355.
29. Flor, H., Fydrich, T., & Turk, D.C. (1992). Efficacy of multidisciplinary pain treatment centers: a meta-analytic review. *Pain, 49*, 221-230.
30. Arntz, A., Dreessen, L., & Merckelbach, H. (1991). Attention, not anxiety, influences pain. *Behav Res Ther, 29*, 41-50.
31. Stone, A.A., Broderick, J.E., Porter, L.S., & Kaell, A.T. (1997). The experience of rheumatoid arthritis pain and fatigue: examining momentary reports and correlates over one week. *Arthritis Care Res, 10*, 185-192.
32. Stone, A.A., Shiffman, S., Schwartz, J.E., Broderick, J.E., & Hufford, M.R. (2003). Patient compliance with paper and electronic diaries. *Control Clin Trials, 24*, 182-199.
33. Jamison, R.N., Raymond, S.A., Levine, J.G., Slawsby, E.A., Nedeljkovic, S.S., & Katz. N.P. (2001). Electronic diaries for monitoring chronic pain: 1-year validation study. *Pain*, *91*, 277-285.

Appendix 6A: Contents of the experience sampling form (ESF)

1. Severity of the pain:

0 1 2 3 4 5 6 7
no pain very much pain

2. Interference of pain

0 1 2 3 4 5 6 7
no interference very much interference

3. Pain control:

0 1 2 3 4 5 6 7
no control very much control

4. Where are you now? ……..…

5. With whom are you now? ……..……

6. What are you doing at this moment?
…….

7. How intensive do you think this activity is?

0 1 2 3 4 5 6 7
absolutely not intensive very intensive

8. How strained are you during this activity?

0 1 2 3 4 5 6 7
absolutely not strained very strained

9. What are you going to do immediately?
……..……...….

10. How skilled are you in performing this activity?

0 1 2 3 4 5 6 7
absolutely not skilled very skilled

11. How much pain do you expect to feel by doing this activity?

0 1 2 3 4 5 6 7
no pain very much pain

12. How well do you think you'll be able to cope with
this pain?

0 1 2 3 4 5 6 7
absolutely not very well

13. This beep disturbed me

0 1 2 3 4 5 6 7
not very much

14. What is the time now? …. hour ... minutes

CHAPTER 7

The match-mismatch model in daily practice: why do chronic pain patients tend to underpredict their pain?

Adapted from: C.A.J. Mes, R. Lousberg, G. Zilvold, D.C. Turk. The match-mismatch model in daily practice: why do chronic pain patients tend to underpredict their pain? (submitted).

7.1 Introduction

The match-mismatch (MM) model of pain addresses the processing and anticipation of pain that is overpredicted, rightly predicted, or underpredicted[1,2,3]. The MM model of pain states that the most important immediate consequence of a mismatch between expected and experienced intensity of a painful event is that the expectation for the next experience is adjusted in the direction of the most immediate prior experience. When the event is perceived as being more painful than anticipated (i.e., an underprediction), the expectation for the next event is predicted to be greater. Conversely, when the preceding event is less painful than anticipated (i.e., an overprediction), the expectation is adjusted so that the anticipated pain severity is lowered. When the expectation for the anticipated pain severity is accurate (i.e., a match) no changes in future pain expectations should occur.

In a previous study (Chapter 5), we demonstrated that the mechanisms of the MM model, studied mainly in laboratory situations, can be applied to daily life situations of people experiencing chronic pain[4]. The results of this study indicated that people with chronic pain underpredict their pain more than twice as much than they overpredict (respectively 41.2% underpredictions, 39.6% matches and 19.2% overpredictions). Given the negative effects of underpredictions on fear of pain, fear of movement, escape and avoidance behavior[3,5,6], it is of both clinical and theoretical importance to determine why chronic pain patients tend to underpredict the severity of pain they believe that they experience when they engage in an activity. Therefore, the primary aim of the present study was to test a model that explains the discrepancy between the numbers of underpredictions versus overpredictions observed in the study described in Chapter 5. The present study builds on this study to expand the understanding about the role the MM model plays in the maintenance of chronic pain.

The MM mechanism is a cognitive process whereby an estimate is based on individuals' prior personal experiences, psychological variables, and situational factors. In the present study, the influence of a selected number of potentially important psychological variables (available from the study in Chapter 5) that are assumed to interfere with the match-mismatch process was investigated.

7. 2 Methods

7.2.1 Hypothesized path model

In order to test a multidimensional, cognitive-behavioral pain model, a path model was developed where the inclusion of several dependent relationships is allowed (Figure 7.1). To correct for the total number of valid mismatches per patient, the discrepancy between the number of underpredictions minus the number of overpredictions was chosen as the dependent variable.

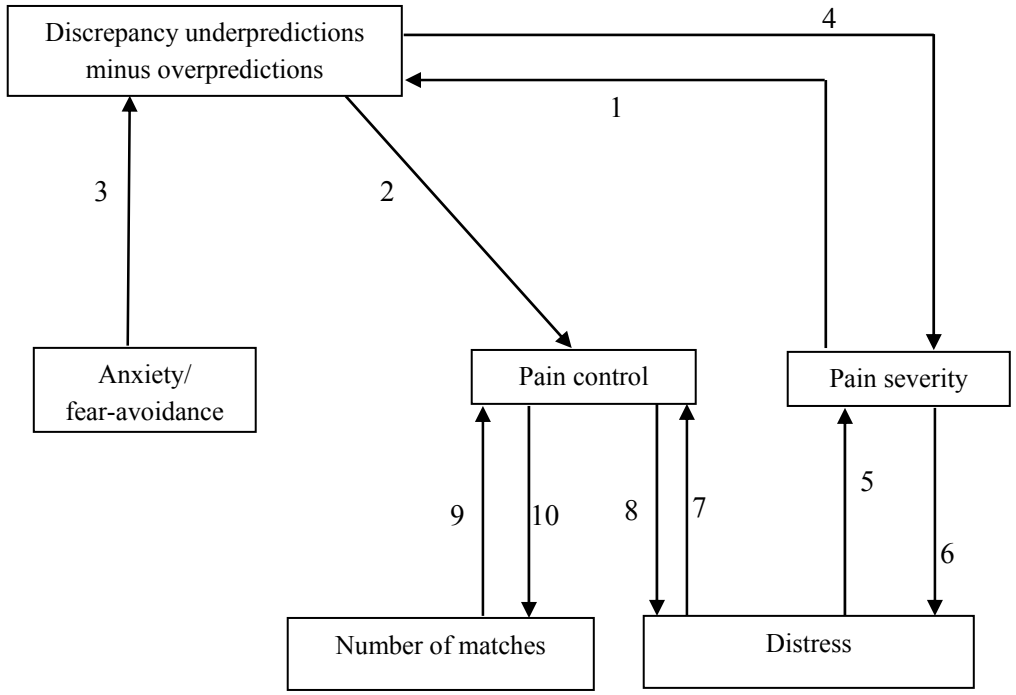


Figure 7.1:Hypothesized relationships between variables

The main question to be answered was to what extent the a priori path model in Figure 7.1 would fit in explaining the discrepancy between the number of underpredictions minus the number of overpredictions. Subsequent analyses concerned the question of which of the following variables contributed significantly to the model:

1. *Pain severity → discrepancy:* Increases in subsequent pain predictions after an underprediction, or decreases after an overprediction, are expected to be a result of

a statistical regression to the mean effect. In fact, previous studies have demonstrated that (mis)match data indeed contain a significant regression to the mean effect besides a 'real' independent mismatch effect[4,7]. Based on this regression to the mean effect, there is a greater chance of an underprediction when pain severity is high. Thus, it may be expected that the more pain a person feels, the larger the discrepancy between the number of underpredictions and overpredictions.

2. *Discrepancy → pain control*: A study by Arntz and Lousberg demonstrated that uncertainty is related to underpredictions[2]. Therefore, it may be assumed that pain control becomes smaller when more underpredictions are made. Thus, we hypothesized a negative relationship between the discrepancy between underpredictions and overpredictions and pain control.

3. *Fear-avoidance → discrepancy*: Several studies have demonstrated that fear plays a role in the prediction of pain. For example, McCracken et al. demonstrated that patients high in pain-anxiety tended to overpredict pain whereas low pain-anxiety patients were generally tended to make underpredictions[8]. Further, Arntz et al. found that anxious dental patients had a tendency to expect greater pain severity than fearless patients[9]. In addition, Arntz and Peters suggested that there are two groups of pain patients[10]. The first group consists of patients who are fearful of pain and show avoidance behavior. They might manifest the tendency to overestimate pain and are relatively anxious and depressed. The second group of patients tends to be tough and confront themselves with (painful) activities that are counterproductive. These patients will likely show a tendency to underpredict their pain. Based on these assumptions, it is hypothesized that the more fear-avoidant a patient is, the larger the number of overpredictions there will be and therefore the smaller the discrepancy between the number of underpredictions and overpredictions.

4. *Discrepancy → Pain severity*: The issue whether or not an underprediction of pain leads to an increase of pain experience is still unresolved. A study by Arntz and Hopmans showed that underpredictions do not "hurt" more, but their impact is higher[11]. On the other hand, other studies demonstrate that after an underprediction, increased levels of pain are experienced[2,12,13,14]. Because of the higher impact underpredictions may have, we hypothesized that the larger and more positive the discrepancy between the number of underpredictions and overpredictions is, the more pain will be experienced.

5. *Distress → pain severity*: According to cognitive-behavioral theories of chronic pain, the influence of distress as a magnifier of pain severity is generally accepted[15,16,17]. In addition, studies on the psychometric properties of the Multidimensional Pain Inventory (MPI) reveal a positive relationship between distress and pain severity[18,19,20].
6. *Pain severity → distress:* Results from studies by Banks and Kerns[21] and McCracken et al.[22] demonstrated that emotional distress can also be a result of chronic pain, as opposed to the fifth hypothesized path. The more pain one experiences, the more a person will feel distressed.
7. *Distress → pain control*: Similar to the fifth path, a (negative) causal relationship is assumed between distress and pain control. The more distress a person experiences, the less control he will feel over his pain.
8. *Pain control → distress*: Similar to the fifth path, a (negative) causal relationship is assumed between pain control and distress. The more pain control a person experiences, the less distress he will feel.
9. *Number of matches → pain control:* A large number of matches may indicate that a person with chronic pain is able to make a adequate prediction of his pain level. It may therefore be assumed that the larger the number of correct matches, the more control one has over the pain.
10. *Pain control → number of matches:* According to Turner and Romano, the "rationale for applying cognitive-behavioral treatment strategies to chronic pain is that learning new cognitive and behavioral responses to pain can provide an individual with a sense of control over pain and among others decrease judgments related to pain" (p. 1711)[23]. On the basis of this rationale, it may be expected that the more pain control a patient experiences, the more certain he is in his predictions. Pain control is not just being able to accurately predict the level of pain, but also the belief that when pain increases, this will be manageable and not lead to serious consequences. Therefore, it is assumed that more matches will be made when an individual experiences more control over his pain.

No other relationships were hypothesized in view of the relatively small number of available participants.

7.2.2 Participants

From June 2000 to June 2002, all chronic pain patients referred to an inpatient multidisciplinary, cognitive-behavioral rehabilitation program were recruited for participation in the study. The exclusion criteria included: age < 18 years, pain duration < six months, presence of serious psychopathology, lack of fluency in the Dutch language, participation in another study, geographical distance > 100 km from the rehabilitation center. Recruitment was continued until 100 patients with informed consent were acquired. A total sample of 179 patients was invited to participate to obtain this number (a 55.9% response rate). On the basis of the selection criteria, 41 patients were excluded; 38 patients gave no reason for their non-participation. The mean age of the 100 included patients in the study (22 men and 78 women) was 41 years (SD = 10.4) and the mean duration of pain was 6.6 years (SD = 6.3). The local ethical committee approved the study. No significant differences were found in sex, age, pain duration, and pain location between the included and excluded patients (all *P*'s > .118).

7.2.3 Instruments and procedures

Data were collected approximately two weeks before the start of the rehabilitation program.

7.2.3.1 Experience sampling

First data collection method applied concerned the Experience Sampling Method (ESM), which is a valid and reliable structured diary method. It allows collecting random snapshots of the participant's mental state (including assessments of pain intensity) within their natural environment[24,25,26]. The accuracy and applicability of ESM have been demonstrated in a number of studies involving patients with chronic pain[27,28,29].

All patients received a Seiko RC-4000 wristwatch and a set of ESM booklets, each containing the necessary Experience Sampling Forms (ESFs). The watch randomly signaled participants 10 times a day, between 7:30 a.m. and 10:30 p.m., for a period of 7 days. The minimum time between two signals was 15 minutes; the maximum time was three hours with an average interval of one and one-half hours. As soon as possible after the auditory signal, the patient was instructed to reply to an ESF in the booklet. The ESFs include 14-item

questionnaires (Appendix 7A). Questions address current pain and some contextual information, as well as expected activities and related pain levels.

By comparing the actual prompt moments with the log time-entries in the booklets it was possible to discard ESFs that were completed more than 10 minutes after the signal. These responses were considered invalid. Only patients who responded validly to more than 30% of the emitted prompts (a tone) were included in the analysis. Delespaul (personal communication) has demonstrated that the likelihood of fooling this protocol retrospectively and still be included as a valid subject in the sample is less than 1%.

7.2.3.2 Multidimensional Pain Inventory

The Multidimensional Pain Inventory (MPI), Dutch language version, was included to assess psychological variables related to pain[20]. The MPI has a good internal consistency and satisfactory test-retest reliability[20]. Further, the MPI has been shown to be reliable and valid for use with diverse samples of chronic pain[30]. For this study, the MPI subscales pain severity, negative distress, as well as the 22th item of part 1 of the MPI (measuring pain control) were used to test the hypotheses of interest.

7.2.3.3 Tampa Scale of Kinesiophobia

The Dutch version of the Tampa Scale for Kinesiophobia (TSK) was used to measure fear of physical activity and fear of re-injury during physical activity[31]. The TSK consists of 17 items that have to be scored on 4-points scales. The total score varies between 17 and 68. A score of more than 37 means that a person has high fear (avoider), a score of 37 or below means that a person has low fear (confronter)[32]. The TSK has good evidence to support its internal consistency, construct validity, and criterion validity[32]. Moreover, the total score does not correlate with age or duration of complaints.

7.2.4 Data reduction and analyses

A match is established when the experienced pain level ($b[experienced]_x$ in Figure 7.2) is equal to the anticipated pain level one prompt before ($b[predicted]_{x-1}$in Figure 7.2). Mismatches can occur when the experienced pain is higher than anticipated in case of an underprediction ($b[experienced]_x > b[predicted]_{x-1}$) or when the experienced pain is lower than anticipated in case of an overprediction ($b[experienced]_x < b[predicted]_{x-1}$). Because the results of a prediction can only be

assessed for two consecutive non-missing observations, the last (10th) signal of the day was always discarded leaving a maximum number of 9 observations per day for (mis)match analysis. Moments were coded as match or mismatch. For every participant, the total number of (mis)matches was counted. As can be viewed in the Appendix, predictions and experiences of pain are linked to ongoing activity. Real (mis)matches should therefore be restricted to pain ratings during activities that were planned during the previous prompt. Activities were translated into Metabolic Equivalent (MET) values[33,34] and corrected for in the analyses, since the activity for which pain is predicted might be different from the activity for which pain is rated. In fact, given the random time sampling procedure, it is quite likely that different events might be rated.

In a previous study (Chapter 5), we demonstrated that the correlation between experienced pain and actual activity level using MET ratings was significant, but small. The contribution of activity as a covariate in the investigated regression models of that study was not significant and coefficients of these models did only very slightly change when not controlling for activity level. These observations were the justification, in the present study, for not correcting (mis)matches for the level of activity.

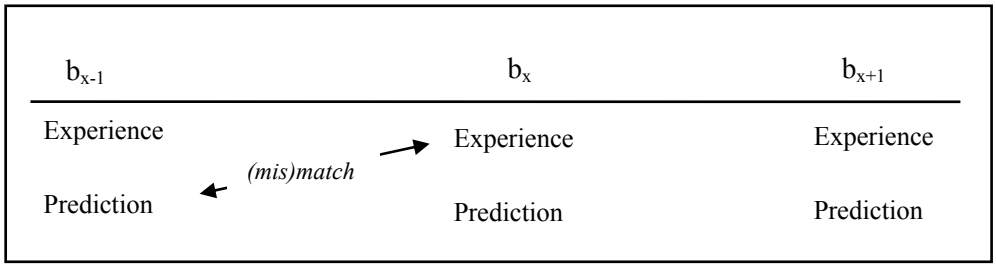


Figure 7.2:Defining a (mis)match

Data were analyzed by means of path analysis, using MPlus 3.0 software[35]. Predictors for the discrepancy of underpredictions minus the number of overpredictions included the variables MPI pain severity, MPI pain control (part 1, item 22), MPI affective distress, the TSK score and the number of matches. Scores on all these variables were standardized into z-scores. The strenght of the association between endogenous and exogenous variables is represented by β coefficients and in the path diagram by straight directed arrows. No arrows are drawn between variables that are hypothesized to be unrelated. The β coefficients

are partial regression coefficients quantifying the strength of the association between two variables if all other variables are held constant in the model[36]. Model-fitting was hypothesis-driven. Model fit was evaluated using three parameters: (1) the χ^2 value; large χ^2 values relative to the degrees of freedom indicate an inadequate fit of the model to the data[37]; (2) the root mean square error of approximation (RMSEA) is provided as it provides sample-size adjusted estimates, indicating good model fit when < 0.050[38]; and (3), the comparative fit index (CFI)[39]; CFI may vary from 0 to 1. A CFI close to 1 indicates a very good fit. Since the relationships between all hypothesized paths were explicitly either positive or negative, it was decided to perform one-tailed tests for the path coefficients. With an α of .05, t values of these coefficients are significant at 1.66.

7.3 Results

7.3.1 Number of signals and data-pairs

Of the 100 participants, 17 did not respond validly to more than 30% of the emitted prompts. These patients were excluded from the analyses, leaving a sample of 83. No significant differences were found with regard to sex, age, pain duration, and pain location, between patients who did and patients who did not have sufficiently valid responses. The mean age of the 83 included patients (16 men and 67 women) was 41.1 years (SD = 10.6) and the mean duration of pain was 7.0 years (SD = 6.9).

A total number of 4579 valid responses were identified (78.8% of the maximum number of 5810 possible responses), 955 (16.4%) prompts were missed, and 276 (4.8%) were invalid (i.e., prompts not answered within 10 minutes). No significant correlations were found between the number of missing prompts and the variables sex, age, pain duration, and pain location. For the (mis)match analyses, 3584 valid non-missing consecutive pairs of observations were available (68.5% of 5229, i.e. 83 patients for 7 days, 9 pairs per day). There were 1418 (39.6%) matches, 1476 (41.2%) underpredictions, and 690 (19.2%) overpredictions. Analysis of missing data revealed that missing data-pairs were randomly distributed with a tendency for more missing data-pairs on the first prompt of the day (participants still asleep).

7.3.2 Explaining the discrepancy between underpredictions and overpredictions

In Table 7.1, the correlations between the discrepancy between underpredictions minus overpredictions and the independent variables (z-scores) are presented.

Table 7.1: Correlations (Pearson's r; P value) between the z-scores of the discrepancy between the number of underpredictions minus overpredictions and the independent variables (N = 83)

	Number of matches	**Pain severity**	**Pain control**	**Fear-avoidance**	**Distress**
Discrepancy underprections/ overpredictions	-.209 (.058)	.141 (.203)	-.200 (.070)	-.147 (.186)	-.023 (.836)
Number of matches		-.043 (.701)	.101 (.365)	-.053 (.634)	.023 (.840)
Pain severity			-.182 (.099)	-.106 (.342)	.161 (.145)
Pain control				-.013 (.911)	-.619 (.000)
Fear-avoidance					-.037 (.740)

The results of the path analysis showed that all relationships were in the hypothesized direction with the exception of the paths 'pain severity → discrepancy', 'number of matches → pain control', 'distress → pain control', and 'pain severity → distress'. The hypothesized model (Figure 7.3) provided a reasonable initial fit (χ^2 = 4.031, df = 5, *P* = 0.545). The RMSEA was 0.000 (90% CI 0.000 – 0.137; probability RMSEA < .05 = 0.652), and the CFI 1.000. None of the variables significantly predicted the discrepancy between the number of underpredictions and overpredictions. The model showed a trend of a role of fear-avoidance in the prediction of the discrepancy between the number of underpredictions and overpredictions.

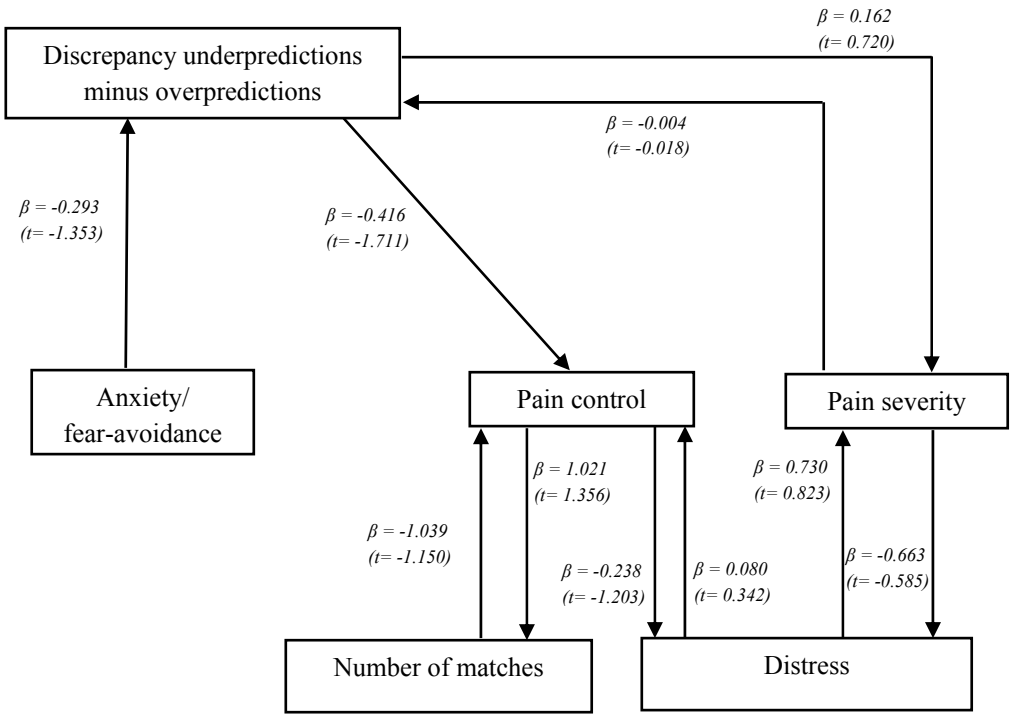


Figure 7.3: A priori path model (weighted standardized path coefficients with t values)

The modification indices suggested that two paths should be added to the path model, namely 'fear-avoidance → pain control' and 'fear-avoidance → distress'. A second model with these modifications was tested. This post hoc model resulted in a better fit (χ^2 =0.369, df = 3, *P* = 0.947) with an RMSEA of 0.000 (90% CI 0.000 – 0.024; probability RMSEA < .05 = 0.960), and a CFI of 1.000. Further modification indices of the post hoc model showed that no significant improvement could be obtained by adding or removing more paths. The final path model is represented in Figure 7.4.

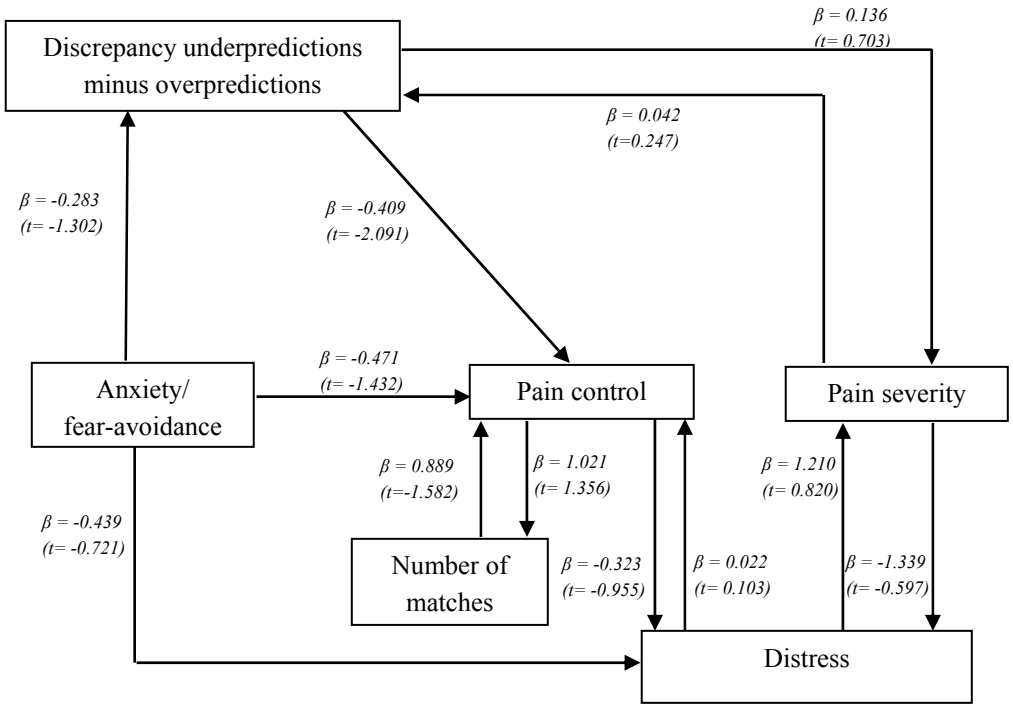


Figure 7.4:Post hoc path model (weighted standardized path coefficients with t values)

7.4 Discussion

In this study, we evaluated a model explaning the relatively large number of underpredictions of pain in chronic pain patients as observed from ESM data in a daily life situation. The small χ^2 value, the RMSEA < 0.001 and the CFI of 1, all indicated that there was a good fit of the final path model. Almost all relationships were in the hypothesized direction. The ones that were not, were non-significant. For the path 'matches → pain control', an unhypothesized trend towards a negative relationship was found. This trend could not be explained on the base of the available data and needs further exploration.

The two paths that were added to the a priori model need a post hoc clarification. First, since fear-avoidance is, over the long-term, related to increases of disability and pain, as described by the fear-avoidance model[31], it may be assumed that fear-avoidant patients will experience less control over their pain.

This negative relationship was already demonstrated by Philips[40]. She showed that avoidance behavior was negatively associated with self-efficacy beliefs - the more avoidance, the less control over pain patients reported. Adding the relationship between fear-avoidance and pain control in the final model can therefore be theoretically grounded. Second, adding the relationship 'fear avoidance → distress' also has a theoretical ground on the base of the fear-avoidance model[31], which describes the relationship between fear, avoidance behavior, and consequent physical deconditioning, disabilities, and depressive symptoms.

The results of the present study provide a heuristic model, in which there was a trend for fear-avoidance in the prediction of the discrepancy between the number of underpredictions and overpredictions. That is, the more fear-avoidant a person with pain is, the smaller this discrepancy. The small number of participants is probably the cause of this relationship not to be significant. Nevertheless, this information increases our knowledge about the underlying cognitive-behavioral mechanisms of chronic pain and may help to develop more effective treatments for patients with chronic pain[41]. Several studies have demonstrated that graded exposure in vivo is successful in decreasing levels of pain-related fear[42,43]. Graded exposure in vivo may also be applied to make pain patients more aware of the role of their estimations of pain and to learn to make more realistic judgments about their pain levels and related behaviors. Thereby, a different approach is needed for fear-avoidant patients. The possibility and clinical consequences of applying graded exposure in vivo to reduce fear-avoidance in combination with the improvement of judgments about pain needs to be further explored.

There are several limitations of our study. First, in the case of path models based on non-experimental data, one must be careful with causal interpretations. Strictly spoken, path analysis can only determine whether data are consistent with the model being tested[44]. With respect to the present study this means that, although a trend of a predictor for the discrepancy between underpredictions and overpredictions was found, this does not necessarily imply a causal nature of fear-avoidance as a cause of the mismatch discrepancy. However, causality does not have to be ruled out, since the presence of a correlation is an essential condition for causality and a trend towards a significant correlation between the investigated variables indeed was found.

A second critical point relates to the fact that the instruments applied in the previous study determined the choice of variables to investigate in the present study. It is likely that there are more variables that influence the discrepancy

between the number of mismatches as well as the other variables included in the model tested. The results of the post hoc model demonstrated a significant residual variance for the variables pain severity, and pain control. In other words, for these variables there is still a significant amount of unexplained variance. Thus, although the present study provides a better understanding of the factors related to the discrepancy between underpredictions and overpredictions, and thus probably of a mechanism for the maintenance of chronic pain, additional research is needed to gain insights into the contextual, emotional, and personal factors influencing the patterns in pain prediction and pain experience of chronic pain patients.

A third point of concern relates to power. Because of the relatively small number of participants (N = 83) in relation to the number of variables in the path model, we can not rule out the possibility that some relationships would have reached significancy in case of a larger sample size.

A final critical issue in this study concerns the validity of the measure of underpredictions and overpredictions and the potential confounding effect of (predicted) activity level. This issue is especially relevant in daily life studies where the activity of the event for which pain was predicted might be different from the activity of the event for which pain was rated. For reasons mentioned in section 2.4, we decided not to correct for activity level. This decision, however, may be questioned, as it probably may not be possible to determine (mis)matches independently of activity level. A more extensive investigation of the role of activity level in pain predictions and pain expectations is needed to find a solution for this methololegical problem.

The results of our study, that chronic pain patients are more inclined to underpredict their pain, are consistent with results reported elsewhere in literature. McCracken et al., for example, investigated predictions and experiences of pain in a population of 43 chronic low back pain (CLBP) patients exposed to six trials of a passive straight leg raising test during a physical examination[8]. The predictions of these patients, who were aware of the fact that they participated in a study of prediction of pain, concerned in 38.4% of the cases an underprediction, in 42.0% a correct match and in 19.5% of the cases an overprediction. Arntz and Peters also demonstrated that CLBP patients are more inclined to underpredict their pain[10]. They investigated 20 CLBP patients and 20 healthy controls, who underwent six trials of laboratory induced pressure pain. The CLBP patients showed a tendency to underpredict pain whereas the controls made more accurate predictions. The

participants were aware of their participation in a study regarding subjective and physical pain responses.

On the other hand, the results of our study are not consistent with the results reported by Linton and Melin[45] and Arntz et al.[9]. Arntz et al. investigated 40 patients receiving extensive dental treatment[9]. These patients showed a tendency to overpredict pain. In the study of Linton and Melin, twelve chronic pain patients predicted pain before entering a treatment program[45]. At dismissal 3-11 weeks later, they were asked to remember how much pain they had at baseline. The patients remembered having significantly more pain than the pain that was rated at baseline. It may be concluded that they generally overpredicted their pain. However, this result might highlight a form of (retrospective) memory bias, which is less likely to occur using the ESM and the prospective mismatches in the present study.

With respect to the issue whether chronic pain patients systematically underpredict or overpredict their pain, it is not possible to draw a definite conclusion. Conflicting results are most likely due to methodological differences between studies: differences in number of participants and (mismatch) trials, the moment of measurement, and the interval between prediction and actual experience. Further, in some of the studies participants have to estimate an experimentally-induced pain stimulus whereas in others participants had to rate 'real life' pain stimuli. A valid and definite comparison of our results with those studies mentioned above is also difficult, since our study is the only one in which ESM was used in order to obtain (mis)match data. Additionally, the participants in the present study were people with chronic pain referred to a rehabilitation program. Since this is a selected group of individuals, it may be possible that more mismatches were made in this group compared to persons with chronic pain who are not seeking treatment. The extent to which the selection of participants affected (biased) the results of the present study needs to be investigated.

In sum, the present study, based on ESM data, investigated the issue why chronic pain patients are more inclined to underpredict than to overpredict their pain. The path model revealed a role of fear-avoidance in the prediction of this discrepancy. Since the present data represent cross-sectional data, it is inappropriate to make a causal interpretation of the modeled relationships. Future studies are needed to explain the causal mechanisms between the investigated variables.

References

1. Arntz, A., & Hout, M.A. van den (1988). Generalizability of the match/mismatch model of fear. *Behav Res Ther, 26*, 207-223.
2. Arntz, A., & Lousberg, R. (1990). The effects of underestimated pain and their relationship to habituation. *Behav Res Ther, 28*, 15-28.
3. Rachman, S., & Arntz, A. (1991). The overprediction and underprediction of pain. *Clin Psychol Rev, 11*, 339-355.
4. Mes, C.A.J., Lousberg, R., Oudshoorn, C.G.M., Zilvold, G., Delespaul, P.A.E.G., & Turk. D.C. (submitted). Pain expectations and experiences: replicating the match-mismatch model in daily life of chronic pain patients.
5. Arntz, A., Hout, M.A. van den, Berg, G. van den, & Meijboom, A. (1991). The effects of incorrect pain expectations on acquired fear and pain responses. *Behav Res Ther, 29*, 547-560.
6. Crombez, G., Baeyens, F., Vervaet, L., & Vlaeyen, J. (1997). Vermijdingsgedrag bij patiënten met chronische lage rugpijn. *Gedrag Gezond, 25*, 126-135.
7. Arntz, A., Hout, M.A. van den, Lousberg, R., & Schouten, S. (1990). Case history and shorter communication: Is the match/mismatch model based on a statistical artefact? *Behav Res Ther, 28*, 249-253.
8. McCracken, L.M., Gross, R.T., Sorg, P.J., & Edmands, T.A. (1993). Prediction of pain in patients with chronic low back pain: effects of inaccurate prediction and pain-related anxiety. *Behav Res Ther, 31*, 647-652.
9. Arntz, A., Eck, M. van, & Heijmans, M. (1990). Predictions of dental pain: The fear of any expected evil is worse than the evil itself. *Behav Res Ther, 28*, 29-42.
10. Arntz, A., & Peters, M. (1995). Chronic low back pain and inaccurate predictions of pain: is being too tough a risk factor for the development and maintenance of chronic pain? *Behav Res Ther, 33*, 49-53.
11. Arntz A., & Hopmans, M. (1998). Underpredicted pain disrupts more than correctly predicted pain, but does not hurt more. *Behav Res Ther, 36*, 1121-1129.
12. Crombez G, Baeuens F, Eelen P. Sensory and temporal information about impending pain: the influence of predictability on pain. *Behav Res Ther, 32*, 611-622.
13. Baeyer, C.L. von, Carlson, G., & Webb, L. (1997). Underprediction of pain in children undergoing ear piercing. *Behav Res Ther, 35*, 399-404.
14. Spafford, P.A., Baeyer, C.L. von, & Hicks, C.L. (2002) Expected and reported pain in children undergoing ear piercing: a randomized trial of preparation by parents. *Behav Res Ther,* 2002, 40: 253-266.
15. Feuerstein, M., & Beattie, P. (1995). Biobehavioral factors affecting pain and disability in low back pain: mechanisms and assessment. *Phys Ther, 75*, 267-280.
16. Bradley, L.A. (1996). Cognitive-behavioral therapy for chronic pain. In: R.J. Gatchel, & D.C. Turk (Eds.). *Psychological approaches to pain management: a practitioner's handbook*. New York: Guilford Press.
17. Turk, D.C. (1999). The role of psychological factors in chronic pain. *Acta Anaesthesiol Scand, 43*, 885-888.
18. Kerns, R.D., Turk, D.C., & Rudy, T.E. (1985). The West Haven-Yale Multidimensional Pain Inventory (WHYMPI). *Pain, 23*, 345-356.

19. Flor, H., Rudy, T.E., Birbaumer, N., & Schugens, M.M. (1990). Zur Anwendbarkeit des West Haven-Yale Multidimensional Pain Inventory im deutschen Sprachraum. *Der Schmerz, 4*, 82-87.
20. Lousberg, R., Breukelen, G.J.P. van, Groenman, N.H., Schmidt, A.J.M., Arntz, A., & Winter, F.A.M. (1999). Psychometric properties of the Multidimensional Pain Inventory, Dutch language version (MPI-DLV). *Behav Res Ther, 37*, 167-182.
21. Banks, S.M., & Kerns, R.D. (1996). Explaining high rates of depression in chronic pain: a diathesis-stress framework. *Psychol Bull*, *119*, 95-100.
22. McCracken, L.M., Faber, S.D., & Janeck, A.S. (1998). Pain-related anxiety predicts non-specific physical complaints in persons with chronic pain. *Behav Res Ther, 36*, 621-630.
23. Turner, J.A., Romano, J.M. (1990). Cognitive-behavioral therapy. In: J.J. Bonica, C.R. Chapman, W.E. Fordyce, & J.D. Loeser (Eds.). *The management of pain in clinical practice*. 2nd edition. Philadelphia: Lea Febiger, p. 1711-1721.
24. Czikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience sampling method. *J Nerv Ment Dis, 175*, 526-536.
25. Delespaul, P.A.E.G. (1995). *Assessing schizophrenia in daily life: The Experience Sampling Method*. Maastricht: University Press.
26. Vries, M.W. de, & Delespaul, P.A.E.G. (1993). Zelfobservatie en intensieve time-sampling in de psychiatrie. In: C.A.L. Hoogduin, P. Schnabel, W. Vandereycken, K. van der Velden, & F.C. Verhulst (Eds*.). Jaarboek psychiatrie en psychotherapie*. Meppel: Van Loghum Slaterus, p. 94-108.
27. Lousberg, R., Schmidt, A.J., Groenman, N.H., Vendrig, L., & Dijkman-Caes, C.I. (1997). Validating the MPI-DLV using experience sampling data. *J Behav Med, 7*, 20: 195-206.
28. Peters, M.L., Sorbi, M.J., Kruise, D.A., Kerssens, J.J., Verhaak, P.F.M., & Bensing, J.M. (2000). Electronic diary assessment of pain, disability and psychological adaptation in patients differing in duration of pain. *Pain, 84*, 181-192.
29. Stone, A.A., Broderick, J.E., Shiffman, S., & Schwartz, J.E. (2004). Understanding recall of weekly pain from a momentary assessment perspective: absolute agreement, between- and within-person consistency, and judged change in weekly pain. *Pain, 107*, 61-69.
30. Turk, D.C., & Rudy, T.E. (1990). Neglected factors in chronic pain treatment outcome studies: referral patterns, failure to enter treatment, and attrition. *Pain, 43*, 7-25.
31. Vlaeyen, J.W.S., Kole-Snijders, A.M.J., Boeren, R.G.B., & Eek, H. van (1995). Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain, 62*, 363-372.
32. Goubert, L., Crombez, G., Vlaeyen, J.W.S., Damme, S. van, Broeck, A. van den, & Houdenhove, B. van (2000). De Tampa Schaal voor Kinesiofobie. Psychometrische karakteristieken en normering. *Gedrag Gezond, 28*, 54-62
33. Willmore, J.H., & Costill, D.L. (1994). *Physiology of sport and exercise*. USA (Champaign, IL.): Human Kinetics.
34. Ainsworth, B.E., Haskell, W.L., Whitt, M.C., Irwin, M.L., Swartz, A.M., Strath, S.J., O'Brien, W.L., Bassett, D.R., Schmitz, K.H., Emplaincourt, P.O., Jacobs, D.R., & Leon, A.S. (2000). Compendium of physical activities: an update of activity codes and MET intensities. *Med Sci Sports Exerc*, S498-S516.

35. Muthen, L.K., & Muthen, B.O. (2004). *Mplus User's Guide*. Los Angeles: Muthen and Muthen.
36. Sobel, M.E. (1996). An introduction to causal inference. *Sociol Methods Res, 24*, 353-379.
37. Scott Long, J. (1983). *Covariance structure models: an introduction to LISREL*. Newbury Park: Sage Publication.
38. Arminger, G., Clogg, C.C., & Sobel, M.E. (1995). *Handbook of statistical modeling in the social and behavioral science*s. New York: Plenum.
39. Bentler, P.M. (1990). Comparative fit indexes in structural model. *Psychol Bull, 107*, 238-246.
40. Philips, H.C. (1987). The effects of behavioural treatment on chronic pain. *Behav Res Ther, 25*, 365-377.
41. Vlaeyen, J.W.S., & Morley, S. (2005). Cognitive-behavioral treatments for chronic pain: What works for whom? *Clin J Pain, 21*, 1-8.
42. Jong, J.R. de, Vlaeyen, J.W., Onghena, P., Cuypers, C., Hollander, M. den, & Ruijgrok, J. (2005). Reduction of pain-related fear in complex regional pain syndrome type I: the application of graded exposure in vivo. *Pain, 116*, 2640-275.
43. Vlaeyen, J.W., Jong, J.R. de, Onghena, P., Kerckhoffs-Hanssen, M., & Kole-Snijders, A.M. (2002). Can pain-related fear be reduced? The application of cognitive-behavioural exposure in vivo. *Pain, 7*, 144-153.
44. Streiner, D.L. (2005). Finding our way: an introduction to path analysis. *Can J Psychiatry, 50*, 115-122.
45. Linton, S.J., & Melin, C. (1982). The accuracy of remembering chronic pain. *Pain, 13*, 281-285.

Appendix 7A: Contents of the experience sampling form (ESF)

1. Severity of the pain:

0 1 2 3 4 5 6 7
no pain very much pain

2. Interference of pain

0 1 2 3 4 5 6 7
no interference very much interference

3. Pain control:

0 1 2 3 4 5 6 7
no control very much control

4. Where are you now? ……..…

5. With whom are you now? ……..……

6. What are you doing at this moment?
…….

7. How intensive do you think this activity is?

0 1 2 3 4 5 6 7
absolutely not intensive very intensive

8. How strained are you during this activity?

0 1 2 3 4 5 6 7
absolutely not strained very strained

9. What are you going to do immediately?
……..……...….

10. How skilled are you in performing this activity?

0 1 2 3 4 5 6 7
absolutely not skilled very skilled

11. How much pain do you expect to feel by doing this activity?

0 1 2 3 4 5 6 7
no pain very much pain

12. How well do you think you'll be able to cope with
this pain?

0 1 2 3 4 5 6 7
absolutely not very well

13. This beep disturbed me

0 1 2 3 4 5 6 7
not very much

14. What is the time now? …. hour ... minutes

CHAPTER 8
General discussion

8.1 Introduction

The central idea of this dissertation is that almost all pain treatment programs are suboptimal. Although many patients do improve, in every treatment program, there are patients who show no improvement. Some even deteriorate following treatment. This dissertation has explored the reasons for treatment variability in a cognitive behavior therapy program (CBT-R) offered at the Roessingh Center for Rehabilitation (RCR) in the Netherlands. Treatment variability was demonstrated using a randomized control trial (RCT). Additionally, an attempt was made to explain treatment variability by means of three additional studies. The first endeavored to determine whether treatment variability could by explained by the existence of homogeneous subgroups of chronic pain patients in the total heterogeneous group of chronic pain patients. The second sought to explain treatment variability by looking at failures in the treatment process, namely the lack of congruence between treatment protocol and the program that was actually implemented. The third study investigated whether treatment variability could be explained by incomplete or incorrect underlying theoretical mechanisms. In this last study, the generalizability of the match-mismatch (MM) mechanism of pain was determined, as well as the influence of the CBT-R program on the consequences of this mechanism in daily life. This effort to establish the causes of variability in treatment results creates new opportunities and ideas for improving the quality of the CBT-R program.

8.2 Program evaluation of the CBT-R program

Both the process and the effect of CBT-R program were evaluated. The effect study investigated both general and subgroup specific effects. In the following sections, the general effect study, the subgroup-specific effect study and the process study will be discussed and appraised on their value for clinical practice.

8.2.1 General outcome of the CBT-R program

The general outcome of the CBT-R program was determined by means of a randomized controlled trial (RCT). The results of this RCT, described in Chapter 2, demonstrated that the CBT-R program had a significant influence on both the degree to which pain interferes with daily life of chronic pain patients and the

experienced severity of pain. As such, we can conclude from the RCT that the general rehabilitation aim of the CBT-R program was attained. After all, the general rehabilitation aim is to "learn to more adequately cope with pain, so that the influence of pain is diminished and one becomes better able to function on a physical and mental level"[1,2]. However, the results of the RCT were quite meager and less significant than the results found in an earlier study conducted by Winter[1] Additionally, the RCT results were less significant than results found in numerous other scientific publications that claim that multidisciplinary rehabilitation programs that are based on cognitive-behavioral principles are an effective means of treating chronic pain[3,4]. Although the current treatment components of the CBT-R program appeared comparable to the treatment components used in Winter's study, some differences between the current RCT and Winter's study exist. First and foremost, methodological differences can be found. Additionally, differences in patients' level of pain severity and differences in treatment vision, approach and logistics could all have been responsible for the divergent findings of the Winter's study and the RCT. Unfortunately, a clear and unambiguous explanation for the difference in treatment effect could not be found.

With the above in mind, an important question remains: Why did the CBT-R program fail to generate the hypothesized effect? The RCT demonstrated that a large variability exists with respect to CBT-R treatment results. For all outcome parameters, some patients improved, some deteriorated and some showed no change at all. It is logical that when the mean of all these patients (improved, deteriorated and no change) is used, no large overall effect can be generated. However disappointing the effect size, it is more important to understand why some patients deteriorated following the CBT-R program. Several explanations for this have been proposed in the previous chapters and will be summarized here below.

One explanation is that the program was insufficiently adapted to the specific needs and characteristics of the patients. The results of the study described in Chapter 3 demonstrated that, based on the MPI-DLV classification system, differential treatment responses between subgroups of patients exist. However, the differences between the MPI-DLV subgroups were small and mostly non-significant. We can thus conclude that, based on the meager differences between MPI-DLV subgroups, it is unlikely that the lack of general treatment effect can be satisfactorily explained by differential subgroup effects.

In Chapter 4, another plausible explanation for the lack of large general, as well as the subgroup-specific, treatment effect was presented. This chapter explored whether factors in the treatment process could have been responsible for the lack of overall treatment effect. The RCT was executed in a period when several therapists were ill or on holiday and that, in this time period, a significant shortage of staff was present. Additionally, some therapy sessions were canceled and some patients were unable to attend all sessions due to double bookings with other sessions. As a result of the above, some patients were unable to fully complete the total CBT-R program. This inability to complete the program may explain the lack of desired treatment effect found in the RCT.

In addition to poor adaptation to subgroup characteristics and process-related factors, it is possible that an explanation for the RCT's failure to demonstrate a large treatment effect can be found in the methodology of the RCT. Several methodological problems have been identified. The first problem is linked to the amount of measurements done. During the RCT, patients were required, on six occasions, to complete a battery of questionnaires. It is thus possible that patients got tired of answering the same kinds of questions over and over again. As a result, patient compliance in completing questionnaires may have significantly decreased throughout the course of the program. This lack of compliance may have influenced the way patients rated their functioning and explain the noteworthy amount of missing data.

A second problem concerns measurement timing. It is possible that the occasions upon which measurements were conducted were poorly suited for determining a treatment effect. This can be illustrated by the following: Some measurements were done directly after the program. Given that the CBT-R aims to attain behavioral change and that behavioral change is a long- term process that cannot be completed and fully attained within eight weeks (the duration of the CBT-R program), it is unlikely that a measurement done directly after treatment will show any large effect. Knowing this, the question then becomes, when exactly should treatment effect be measured? Some therapists at the RCR claim that the maximum effect is reached three months after treatment, while others claim that this maximum effect cannot be reached prior to one year after treatment. This disparity in opinion poses a significant question for the present study: Is the decrease in effect, measured at three months follow-up (as demonstrated in Chapter 3) caused by the fact that the program itself is poorly suited to needs of the patients or is the decrease in effect the result of poorly timed measurements? Should the

final treatment effect be measured at a different moment in time? And, if indeed, treatment effect should be measured at a different moment in time, what is the most appropriate time for measuring effect? Is it three months, six months, a year, two years or longer? Unfortunately, a long period of follow-up is, in most cases, less than feasible due to time and financial restraints. Furthermore, a long period of follow-up is hard to measure using an RCT. An RCT that investigates the effect of a specific treatment requires controlled settings. The only difference between the intervention group and the control group ought to be the treatment itself. It is only when all other factors are comparable that an effect can be attributed to the treatment. Unfortunately, after the eight weeks of the CBT-R program, patients can no longer be compared in controlled settings. It is impossible to know exactly what patients do following treatment. They may enroll in other treatment programs or experience life events or other situations that may impact their functioning and pain. In a long period of follow-up, the chance that such intervening situations or events occur becomes greater. It then becomes more plausible that changes in complaints and behavior can be attributed to these situations and events instead of to the treatment.

The third methodological problem concerns the choice of instruments used to determine treatment effect. Given that the treatment protocol of the CBT-R program was vague and that the aims of treatment were insufficiently described and operationalized (see Chapter 4), it was difficult to determine which instruments should be used to measure the effect of the CBT-R program. Which instruments were most appropriate was determined by looking at the kinds of changes expected to occur as a result of participation in the CBT-R program. It is possible that the instruments employed were less appropriate than expected. Although the instruments used appeared to be appropriate measures of the general rehabilitation aim of the program, namely to "learn to more adequately cope with pain, so that the influence of pain is diminished (MPI-DLV Interference) and one becomes better able to function on a physical and mental level (SCL-90, RAND-36)", the instruments may not have been specific or sensitive enough to measure the detailed rehabilitation goals of the CBT-R program or the sub-goals of the various treatment disciplines. Another argument supporting the contention that the instruments used were inappropriate measures of treatment effect is related to the multidimensional nature of chronic pain. It has already been established that an international consensus about the multidimensional nature of chronic pain exists[5,6]. If we accept this, it is logical that, not only clinical practice (diagnostics, treatment and

prevention), but also outcome measurement should be of a multidimensional nature. Under auspices of the Initiative on Methods, Measurement and Pain Assessment in Clinical Trials (IMMPACT), multidimensional core outcome domains and outcome measures that should be considered in clinical trials of chronic pain treatment were identified and recommended[7,8]. These domains include: pain, physical functioning, emotional functioning, participant ratings of global improvement, symptoms and adverse events and participant disposition. The measures applied in the RCT do partially coincide with the measures recommended by IMMPACT (MPI-DLV, RAND-36, MPQ-DLV), but do not cover all the recommended outcome domains. Indeed, the description of the treatment effect of the CBT-R program would have been more complete if all outcome domains were covered. However, it is doubtful whether a more complete set of instruments would have led to different conclusions on the effect of the CBT-R program.

A final methodological problem and explanation for the lack of large treatment effects in the CBT-R program relates to the RCT's power. The power of the RCT is likely to be one of the main reasons for the RCT's inability to demonstrate large statistically significant treatment effects. Although the RCT was powered at n=60 on the MPI-DLV parameter Interference for both the intervention group and the control group, whether this sample size was sufficient enough for determining an effect on the other parameters is questionable. Unfortunately, the number of subjects included in a study is usually less than ideal due to time restraints and financial limitations. Additionally, it is usually the research organization that determines most aspects of a study's design and, consequently, the number of subjects to be included in a research population. Moreover, a large research population is also not always ideal. Including a large number of participants can create a situation in which very small effects become statistically significant. Even then, whether small effects are also clinically meaningful is something that should be given consideration. Statistical significance is not the only way to judge the importance of a program's results[9]. In Chapter 3, a ½ SD from the mean group score at baseline was used as a criterion for clinical relevance[10]. The results of this analysis indicated that the percentage of patients who improved on a specific outcome parameter was, in most cases, lower than the percentage of patients who worsened or showed no change at all. Evidently, the clinical relevance of the CBT-R results was small. This means that even if a larger research population had been included and statistically significant results had been obtained on all outcome parameters, the CBT-R program still would have had a

minimal effect on the overall functioning of chronic pain patients. The clinical relevance of the results would have remained limited.

8.2.2 Subgroups of chronic pain patients

The Multidimensional Pain Inventory (MPI) is advantageous in that it offers a multidimensional measure of patient psychosocial and behavioral functioning. Unfortunately, the MPI does not cover all dimensions of functioning. Turk and Rudy have suggested that pain should be assessed on the three axes of the Multiaxial Assessment of Pain (MAP) system[11]. The MPI only covers axes two and three of this system, namely a psycho-social axis and an axis concerning the quantification of pain behavior. It can be contended that a useful and sufficient classification system for chronic pain patients must not only include psycho-social pain behavior axes, but also an axis that covers somatic and medical aspects of pain. Nonetheless, the MPI is accepted worldwide as a useful, valid and reliable tool for the classification of pain patient subgroups[12,13,14,15,16].

Not only do we need to consider the multidimensional nature of chronic pain in the treatment of chronic pain, we also need to satisfy the individual needs and characteristics of patients within each subgroup. A classification system can never fully incorporate and include all features of chronic pain patients[17] but patient features can definitely influence treatment effect. Turk, Zaki and Rudy have developed a treatment program specifically tailored to the clinical needs of patients with temporomandibular disorder (TMD) classified by the MPI as dysfunctional[18]. A proportion of patients failed to obtain clinically significant improvements. It is possible that this failure to improve in a tailored program is an indication that the knowledge about the underlying mechanisms of chronic pain is still insufficient. Another study by Broderick, Junghaenel and Turk has demonstrated that MPI classifications may not be stable[19]. The acquisition of additional information to support adaptation style has thus been suggested. Additional knowledge about the underlying mechanisms of every theoretical-empirical subgroup of pain patients is needed. Additionally, it is necessary that we generate further knowledge about the ways in which certain patients react in to specific situations and how to tailor treatment programs to the needs of a patient. This kind of knowledge may not only help to improve programs for patients who failed to improve after treatment, but it may also enlarge the efficacy of programs for the patients that already show improvements in existing programs. Theoretically speaking, without substantial knowledge on the etiological and maintaining mechanisms of the MPI subgroups,

one should be very cautious in tailoring treatment to the characteristics of these subgroups.

Currently, the RCR pain division is tailoring the semi-inpatient CBT-R program. Firstly, during the intake procedure patients are required to show that they are prepared for self-management and prepared to work towards behavioral change. Secondly, in the first few weeks of treatment, the appropriate treatment components for the individual patient are determined, personal treatment goals are formulated and factors needing specific attention are noted. Patients then follow the treatment modules (either on a group or an individual basis) that are most likely to meet the patients' needs. In this way, a step towards a more tailored and personalized program is taken. Whether tailoring and personalizing the CBT-R program actually decreases treatment variability in the CBT-R program needs to be scientifically investigated.

8.2.3 Process variables as determinants for treatment outcome

In Chapter 4, the results of a CBT-R program process evaluation were described. Despite the fact that most of the implemented content and goals of the program corresponded with the protocol, some differences between planned program and the actual program implemented were evident. The differences related mostly to drop out and double booking of therapy sessions. Additionally, there are some cases in which the program was adapted slightly to meet the specific needs and requests of the rehabilitants. It may be that process factors had an impact on whether or not patients were able to gain significant or clinically relevant therapy effects. The mentioned differences between the planned program and the actual program may thus be partly responsible for variability in general and subgroup specific treatment results.

A critical investigation of the CBT-R treatment protocol revealed that neither the selection criteria nor the indication criteria of the program were fully operationalized or described. When clear and well defined criteria are unavailable or cannot be measured, it is likely that patients who do not fulfill the criteria are, on occasion, included. As a result, we may assume that during the RCT, some patients in the CBT-R were inappropriate candidates and did not fulfill the selection criteria. We could thus expect, a priori, that for these patients, either no effect or a very small treatment effect would be generated. These patients may very well be the patients that failed to improve following treatment and therefore the patients responsible for variability in treatment results. A clearer, more objective and better

operationalized description of the selection and indication criteria is required so that we can determine whether or not patients participated in the CBT-R program who should not have been admitted. Since the RCT, the RCR pain division has improved the intake procedure for the CBT-R program. Stricter selection criteria have been established in conjunction with the shift towards a more tailored program. Recent analyses indicate that patients included after the implementation of stricter selection criteria experience more clinically relevant improvements with respect to interference of pain and mood (Schreurs, personal communication 2006). Further investigation is needed to properly demonstrate these observed positive consequences of the adapted intake procedure.

In addition to vague selection and indication criteria, the CBT-R program protocol contains poorly described and inadequately operationalized program goals and discipline-related sub-goals. Although the overall theoretical basis of the CBT-R program is described in terms of the principles of cognitive-behavioral approach of pain and load versus load capacity, the theoretical bases for the separate program goals and activities of the treatment disciplines are limited. A clear, objective and operationalized description and theoretical background of these goals and activities is necessary if we are to understand how certain activities result in the attainment of certain goals. The inclusion of this kind of description in the protocol is of further relevance as it can indicate a) the kinds of instruments needed to measure treatment outcome; and b) the causes of the variability in treatment results, namely program failures or theoretical failures. It is important to note that the RCR has made changes since the results of the RCT and process evaluation were made available. The description of the goals, activities and theoretical background of the CBT-R program has been improved in such a way that these are more objective and better operationalized. The RCR is also working on establishing which instruments are most effective for measuring treatment outcome.

Until recently, evaluations in rehabilitation research almost always focused on outcomes of rehabilitation programs. Process evaluations were more common in social science studies. More recently, process evaluations have been included in rehabilitation research[20]. In 2004, a national project called Transform was started by the Center for Rehabilitation-UMCG (CR-UMCG) in the Netherlands. This project aimed to uncover the black box of rehabilitation treatment in four Dutch rehabilitation programs by means of theory-driven evaluation. The content and theoretical bases of the four rehabilitation programs were investigated in an effort to uncover how programs can best be improved and evaluated[21]. This is a positive

shift towards the supplementation of outcome-based evaluations that, often times, fail to determine exactly which factors are responsible for the success or failure of treatment. In short, the inclusion of factors that influence the treatment process is imperative. As a result, the CR-UMCG's recognition that the investigation of treatment theory in rehabilitation research and practice is important is a very positive development. Indeed, an extensive evaluation of treatment process is a timely and costly endeavor but without it, we will never be able to fully understand and sufficiently explain variability in treatment or a lack of effectiveness in treatment programs.

8.3 Investigating the match-mismatch model of pain using experience sampling

The previous paragraphs have shown how treatment variability in the CBT-R program can be explained by looking at differences between subgroups of chronic pain patients, as well as by looking at factors in the treatment process. It is also possible that treatment variability is the result of inaccurate or incomplete theoretical knowledge. It could thus be the case that, due to insufficient knowledge about the origin and maintenance of chronic pain, patients receive non-optimal treatment. So far, many theoretical mechanisms have been proposed and have been shown to be applicable in explaining the maintenance of chronic pain. In this dissertation, the match-mismatch (MM) model was selected for investigation as it may offer a feasible explanation for the development and maintenance of chronic pain. The MM model of pain was investigated in three unifying studies. The first study investigated whether this model can be applied in the daily life situations of chronic pain patients. In the second study, the influence of the CBT-R program on the consequences of this mechanism in daily life was measured. The third study attempted to establish the influence of several psychological variables on the MM mechanism. In the following paragraphs, these three studies, their methods and their results are discussed and appraised with respect to their value in clinical practice.

8.3.1 Predictions and experiences of (chronic) pain

The chapters concerning the MM model of pain have shown that this model describes a psychological mechanism that may be (partly) responsible for the development and maintenance of pain. Chapter 5 demonstrated that the

relationships between pain predictions and pain experiences, as stated by the MM model and mainly investigated in laboratory situations, actually appeared in daily life situations. Pain predictions were mainly adapted in the direction of the last experience.

Crombez, Eelen and Baeyens have criticized the MM model[22]. They have stated that the MM model ignores processes formulated by other learning theories to explain the habituation phenomenon. Additionally, they point out that the MM model is unclear with respect to whether or not an underprediction of pain leads to an increase in *experienced* pain. Studies by Arntz et al. address the habituation phenomenon[23,24] but the discussion with regard to the consequences of underpredictions on subsequent pain experiences remains. While Crombez, Baeyens and Eelen[25] found that underpredictions led to increases in pain, Arntz and Hopmans[26] demonstrated that underpredictions were associated with less pain. In the present study, underpredictions led to an increase in experienced pain. More studies are needed to confirm this finding. However, given that the present study was executed using a relatively large sample of chronic pain patients in their natural environment (which generally means greater ecological validity), we can assume that the observed relationship between underpredictions and experienced pain is realistic.

The results of the study described in Chapter 5 revealed that chronic pain patients tend to underpredict their pain. Additionally, the mean effect of underpredictions on subsequent pain experiences was larger in comparison with the effect of overpredictions. The role of the discrepancy between pain predictions and pain experiences in the development of chronic pain has already been demonstrated in studies by Philips and Vlaeyen et al.[27,28]. They developed the fear-avoidance model of pain which describes how a catastrophizing pain coping style can transform acute pain into chronic pain (see Chapter 1). The MM model is closely related to or implicitly part of both the fear-avoidance model and the avoidance-endurance model of pain[29], as these models describe how pain patients' behavior is influenced by perceptions, beliefs and memories about pain. Expectations with respect to pain are formed under certain conditions and on the basis of previous experiences. Avoidance or endurance may be a result of these expectations. The fear-avoidance model and avoidance-endurance model of pain thus describe the transition from acute to chronic pain. As the results of the present study were demonstrated in chronic pain patients, it could be argued that the MM model is applicable in all stages of pain, including the maintenance of chronic pain.

No definite conclusions can be drawn about the causal mechanisms underlying the discrepancy between pain predictions and pain experiences or the reasons why pain prediction lays structurally beneath the experience of pain (Figure 5.3). In Chapter 7, an attempt was made to explain why pain patients tend to underpredict pain. The results of this study indicated that fear-avoidance is a direct and independent factor in predicting the discrepancy between the number of underpredictions and overpredictions. In essence, the more fear-avoidant a patient is, the more likely he or she will be to make overpredictions. This results in a smaller discrepancy between the number of overpredictions and underpredictions. Since this study is cross-sectional, causal interpretations of the modeled relationships cannot be made. As a result, more experimental, and in particular prospective, studies are needed to elucidate the causal mechanisms of underpredictions. Currently, new studies are being designed and executed. These studies are aimed at investigating whether the MM mechanism is applicable in multiple chronic pain populations, such as fibromyalgia and chronic whiplash syndrome. These studies also aim to investigate, in greater depth, a) the relationship between pain predictions and pain experiences; and b) the influence of pain cognitions, emotions, physical limitations and activity level on the discrepancy between underpredictions and overpredictions. A better understanding of the causal mechanisms underlying the discrepancy between underpredictions and overpredictions will be helpful in developing, and incorporating into the CBT-R program, clear guidelines for dealing with (mis)matches.

It is imperative that we pay attention to the MM mechanism of pain in chronic pain treatment in view of the relationship between underpredictions and fear of pain and movement, as well as avoidance behavior and increases of pain[24,30,31]. Unfortunately, the current CBT-R pays no specific attention to this mechanism. The study described in Chapter 6 demonstrated that the CBT-R program had a positive influence on the number, size and effect of (mis)matches. It is important to note that a decrease in mismatches is not explicitly included in the aims of the CBT-R program. However, we can hypothesize that if the MM model is explicitly incorporated in the CBT-R program, the influence demonstrated in Chapter 6 may very well increase. It is imperative that pain patients gain more insight into the relationship between their judgments of pain, their activity level, their behaviors and their cognitions. The present CBT-R program already focuses on teaching patients to make more realistic judgments about personal boundaries with regard to physical and mental functioning. However, in view of the

importance of mismatches, more attention should be paid to making realistic judgments about pain itself and pain-provoking situations. A suitable method for this may be offered by graded in vivo exposure towards possible pain provoking situations. This means that patients would be presented with manageable pain-evoking situations over a period of time and that eventually the intensity of emotional reactions to pain would decrease.

Although the occurrence of mismatches in the MM model has been demonstrated to be a pain-maintaining mechanism, the MM model does not cover all maintaining factors of chronic pain. It is likely that several different mechanisms function simultaneously to impact pain in different subgroups of patients. The unique and specific contribution of each of these mechanisms in each subgroup needs to be clarified. Additionally, interactions between these mechanisms have to be ascertained. Clarifying which mechanisms play a role is not a matter of determining which mechanism is better. Rather, clarification of pain maintaining mechanisms is a matter of determining which mechanism is most applicable for a given patient subgroup. In the case of the MM model, it would be interesting to investigate the differential responses of pain patient subgroups to establish whether or not the MM mechanism differs among patients. We could explore issues such as whether the tendency to underpredict pain is equal in all pain diagnoses and whether the consequences of mismatches are equal for all pain patients. In a post hoc analysis of the ESM data of the present MM study, a comparison was made of the number and mean size of (mis)matches between the dysfunctional, interpersonally distressed and average types of patients as determined with the MPI-DLV classification system. The analysis revealed a trend towards a differential response between MPI-DLV clusters with regard to the number of matches in favor of the dysfunctionals ($P = .079$). A more extensive investigation that includes a sufficient number of patients per MPI-DLV subgroup is necessary if we want to draw definite conclusions about differential patient response concerning the MM mechanism. Further investigation is also necessary to determine if chronic pain treatment should be adapted based on differential MM mechanisms between patients and, if so, how.

As the present study was aimed at determining the nature of the relationship between mismatches and various psychosocial variables, it may be interesting to investigate the effect of mismatches on psycho-physiological and central parameters. Currently, at Roessingh Research & Development (RRD),

several studies investigating the influence of pain on muscle activity as measured by means of electromyography (EMG) are being conducted[32,33]. With regard to central measures, RRD recently initiated studies on the relationship between pain-related psychosocial aspects and Event Related Potential's (ERP's)[34]. Similar studies can be designed to determine the relationship between mismatches and muscle activity or ERP's. These studies can investigate how physical substrates play a role in the maintenance of pain or in the discrepancy between underpredictions and overpredictions. Using the results of the studies currently being conducted at RRD, a measurement instrument could be developed in which ESM and EMG or EEG (ERP) measurements can take place at the same time. If we were to succeed in establishing relationships between the mentioned variables, it would be possible to develop a feedback system that warns patients at the moment that a mismatch is made. Such an instrument may also be able to warn for the consequences of mismatches so that patients could be conditioned to create realistic matches between pain predictions and pain experience.

8.3.2 Experience sampling in chronic pain research

For the studies conducted on the MM mechanism of pain, the experience sampling method (ESM) was applied using paper diaries. The accuracy and applicability of the ESM has been demonstrated in a number of studies involving patients with chronic pain. Until recently, studies on the MM mechanism were mainly executed in laboratory settings. In a laboratory setting, all sorts of influences are filtered out and patients may react differently than they would in their own private setting. For example, in a laboratory setting, the display of social desirability is much higher than in private settings. Real life pain experience and pain behavior is often influenced by several factors in a patient's natural environment. For example, the presence or attitude of a patient's partner may influence pain experience and pain interference[35,36]. It is thus important to take all these factors into account when measuring the extent the MM mechanism and other mechanisms determine pain and pain behavior. Evidently, there is a need for more ecologically valid studies on the MM mechanism and other pain mechanisms in general. The ESM can be applied in daily life situations and has already shown its value in offering more realistic and ecological valid data with regard to the MM mechanism in chronic pain patients.

In Chapters 5-7, several limitations of the ESM using paper diaries were mentioned. One limitation is that no external check on the validity of the data is

available. It is impossible to know whether or not patients referred to previous ratings, by looking at previous diary entries, to rate subsequent pain experiences. This limitation can be resolved by using a palmtop computer (PTC). For the present study, however, paper diaries were used for two reasons. Firstly, the costs related to purchasing and distributing PTCs was beyond this study's budget. Secondly, extensive open questions on contextual situations cannot be answered effectively on a PTC. A further limitation of the ESM is directly related to the very use of ESM and the prompting that occurs. Although ESM-based observations seem to be unaffected by anticipation because of the random time scheduling of signals[37,38], it is possible that the frequency of recording influenced the nature of patients' responses. Patients were required to carry the paper diaries with them and to stop what they were doing ten times a day. This may have been experienced as intrusive. Although a study of Cruise et al. has demonstrated that there are no reactive effects of intensive pain diary methods on levels of pain[39], the amount of signals and items to be answered in the present study needs to be considered. Evidently, we endeavored to obtain as much information as possible on pain levels during the course of a given day. Unfortunately, there is no instrument available that measures pain continuously and objectively. As a result, the number of measuring days, the number of prompts and the number of diary questions had to be considered. Although some of the patients experienced the ESM study as intrusive, most patients were indeed highly willing to participate. Patients were free to discontinue their participation at any given moment without explanation. Participation also took place separately from participation in the CBT-R program. For these reasons, we can assume that patients that participated did not feel overly obligated to participate and complete the diaries. We can thus assume that the design of the present study was not excessively intrusive and that the responses given were a fair reflection of what we attempted to measure.

One final issue with regard to the ESM methodology concerns the diary questions on pain predictions and pain experiences. In the present study, these questions were related to ongoing activity. Since a random time sampling procedure was applied and the duration between two prompts took at least 15 minutes, the activity for which pain was predicted did not necessarily have to correspond with the activity for which pain was experienced. If only pain ratings were included for a comparable activity, there would have been a very large number of missing values and consequently the power of the study would have decreased. Instead, the analyses were corrected for activity level operationalized

into Metabolic Equivalent (MET) values. This resulted in comparable findings even when there was no correction for activity. The question then is: To what extent did the correction for activity level solve the methodological problem? Given that the ESM study was already highly demanding of participants, increasing the number of prompts per day was not an option. A plausible and likely more feasible solution would have been to split the pain ratings from activity level. The consequences of this kind of action need to be further investigated.

In the current CBT-R program, patients are required to make a list of their daily activities for two days prior to starting treatment. In addition, patients with chronic fatigue syndrome that are admitted to the RCR pain division are required to carry an activity monitor for two weeks prior to start of their treatment. At this time, a study that investigates the relationship between the registration of activities by means of activity monitoring, by means of ESM and by means of subjective recording with questionnaires, is being performed. One aim of this study is to determine whether ESM can replace the other two methods of activity registration. This is preferable as the ESM is easily applicable in daily life and factors influencing activity patterns may be easily measured at the same time.

Section 8.2.1 mentions that, based on studies relating ESM results to physical substrates like EMG registrations, it may become possible to develop a feedback system that warns patients when they are inclined to make a mismatch. This kind of system may also be able to provide consequences for mismatches on a physical level. Because of its applicability in daily life, ESM can eventually offer possibilities for tele-treatment, such as monitoring and treatment independent of time and place (for instance in the home situation) utilizing a mobile service infrastructure. Tele-treatment is expected to be effective and efficient as patients can practice and train new behaviors more intensively in their own environment and professionals can treat several patients at the same time[40]. More studies are needed of find out how direct feedback on matches and mismatches can be provided through tele-treatment.

8.4 Conclusion and recommendations

In conclusion, the studies conducted in this dissertation have presented the reader with numerous recommendations and guidelines for the optimalization of the CBT-R program through the reduction of treatment variability. Unfortunately, the problem of variability in treatment results of the CBT-R program remains partly unresolved. Given this, how can we improve the effect of not only the CBT-R program at the RCR but also other pain management programs? What kind of studies or what kind of knowledge is needed to reduce the variability in results? A few recommendations can be offered. In the absence of significant financial and time restraints, the following can be done: First and foremost, the CBT-R program in its totality and the separate disciplines in the CBT-R program can be extensively described. We need to describe, chart and operationalize: a) the treatment itself: it's components or activities, intensity, frequency and duration; b) the aims and goals of the program; and c) the implementation-environment. This means that all these domains and their underlying theoretical assumptions should be clear, objective and operationalized. Secondly, we can then determine which instruments are most appropriate for measuring the outcome of the CBT-R program. The instruments selected must correspond with the aims and theoretical assumptions of the program. Thirdly, an extensive evaluation of the treatment process should be conducted. The incongruencies between the planned treatment program and actual treatment process need to be understood. Ineffectiveness of a treatment program can be caused by either program failures or theoretical failures. It is imperative that we investigate program failure first. When we are able to rule out incongruencies between the planned treatment program and the actual treatment situation (a program failure) as an explanation for treatment ineffectiveness, we can then investigate the theoretical background of the program. Lastly, treatment effect can be established by means of the selected instruments. Since approximately a year, RCR patients' data on applied measurement instruments are stored in a database. In this way, treatment effect can be monitored more constantly and treatment contents can be adapted if necessary.

To date, the knowledge on how chronic pain persists, on why some patients continue to experience chronic pain long after the physical cause for pain disappears, and on how chronic pain patients can best be treated remains insufficient. As a result, it is imperative that we not only alternately evaluate

treatment process and effect but also simultaneously execute experiments that endeavor to expand our knowledge on the development and maintenance of chronic pain. Further, it is important that studies be conducted on how, using complementary theoretical concepts, the CBT-R program can best be adapted. Currently, we can see a trend in pain research towards studies that attempt to determine which treatment factors are effective for specific patient sub-groups. Not only factors predicting the success of treatment are investigated, but also characteristics of patients that may determine the degree to which a given patient will profit from the available treatment. At the same time, studies are being conducted on the development of objective indicators that can be used to determine which patients should or should not be admitted to a certain treatment program. Using such studies, we should then be able to distinguish between those pain patients who will develop chronic pain and those who will not. Once we are able to recognize these characteristics, we can sharpen the procedures by which patients are admitted to chronic pain treatment programs and thus offer patients that are less suitable for the CBT-R program an alternative. By optimizing the selection procedure and by customizing the treatment program to the specific characteristics, needs and profiles of patients or patient subgroups, we can decrease treatment variability and, in the end, improve the overall effectiveness of chronic pain treatment.

References

1. Winter, F.A.M. (1992). *Pijnrevalidatie voor chronische pijnpatiënten: De pijn de baas*. Enschede: Intermap
2. Divisie pijnrevalidatie Het Roessingh (1999). *Beslissingsondersteunend en verrichtingenprotocol van het Pijn de Baas programma*. Enschede: Het Roessingh.
3. Morley, S., Eccleston, C., & Williams, A. (1999). Systematic review and meta-analysis of randomized controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache. *Pain, 80*, 1-13.
4. Flor H, Fydrich T, Turk DC (1992). Efficacy of multidisciplinary pain treatment centers: a meta-analytic review. *Pain, 49*, 221-230.
5. Flor, H., Birbaumer, N., & Turk, D.C. (1990). The psychobiology of chronic pain. *Advances in Behaviour Research and Therapy, 12*, 47, 84.
6. Feuerstein, M., & Beattie, P. (1995). Biobehavioral facotrs affecting pain and disability in low back pain: mechanisma and assessment. *Phys Ther*, 75, 267-280.
7. Turk, D.C., Dworkin, R.H., Allen, R.R., Bellamy, N., Brandenburg, N., Carr, D.B., Cleeland, C., Dionne, R., Farrar, J.T., Galer, B.S., Hewitt, D.J., Jadad, A.R., Katz, N.P., Kramer, L.D., Manning, D.C., McGormick, C.G., McDermott, M.P., McGrath, P., Quessy, S., Rappaport, B.A., Robinson, J.P., Royal, M.A., Simon, L., Stauffer, J.W., Stein, W., Tollett, J., & Witter, J. (2003). Core outcome domains for chronic pain trials: IMMPACT recommendations. *Pain, 106*, 337-345.
8. Dworkin, R.H., Turk, D.C., Farrar, J.T., Haythornthwaite, J.A., Jensen, M.P., Katz, N.P., Kerns, R.D., Stucki, G., Allen, R.R., et al. (2005). Core outcome measures for chronic pain trials: IMMPACT recommendations. *Pain*, *113*, 9-19.
9. Dyer, I. (1998). The significance of statistical significance. *Accid Emerg Nurs, 6*, 92-98.
10. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.
11. Turk, D.C., & Rudy, T.E. (1986). Assessment of chronic pain states: MAP-ping out the terrain. *J Pain Symptom Manage*, *1*, 235-237.
12. Dahlström, L., Widmark, G., & Carlsson, S.G. (1997). Cognitive-behavioral profiles among different categories of orofacial pain patients: diagnostic and treatment implications. *Eur J Oral Sci, 105*, 377-383.
13. Turk, D.C., Okifuji, A., Sinclair, J.D., & Starz, T.W. (1998). Differential responses by psychosocial subgroups of Fibromyalgia syndrome patients to an interdisciplinary treatment. *Arthritis Care Res*, *11,* 397-404.
14. Bergström, G., Bodin, L., Jensen, I.B., Linton, S.J., & Nygren, A.L. (2001). Long-term, non-specific spinal pain: reliable and valid subgroups of patients. *Behav Res Ther, 39,* 75-87.
15. Hankin, H.A., Spencer, T., Kegerreis, S., Worrell, T., & Rice, J.M. (2001). Analysis of pain behavior profiles and functional disability in outpatient physical therapy clinics. *J Orthop Sports Phys Ther, 31,* 90-95.
16. Vollenbroek-Hutten, M.M., Hermens, H.J., Wever, D., Gorter, M., Rinket, J., & IJzerman, M.J. (2004). Differences in outcome of a multidisciplinary treatment between subgroups of chronic

low back pain patient defined using two multiaxial assessment instruments: the multidimensional pain inventory and lumbar dynamometry. *Clin Rehabil, 18*, 566-579.
17. Turk, D.C. (2005). The potential treatment matching for subgroups of patients with chronic pain: lumping versus splitting. *Clin J Pain, 21*, 44-55.
18. Turk, D.C., Zaki, H.S., Rudy, T.E. (1993). Effects of intraoral applicance and biofeedback/stress management alone and in combination in treating pain and depression in TMD patients. *J Prosthet Dent*, *70*, 158-164.
19. Broderick, J.E., Junghaenel, D.U., & Turk, D.C. (2004). Stability of patient adaptatiohn classifications on the multidimensional pain inventory. *Pain, 109*, 94-102.
20. Mes, C.A.J., Winter, F.A.M., & Wijck R. Van (1999). Programma-evaluatie in de pijnrevalidatie: analyse van het Pijn de Baas programma. *Gedrag Gezond, 27*, 136-147.
21. Lettinga, A., & Poels, B. (2005). Theoriegestuurde evaluatie van behandelprogramma's: toetsen om te verbeteren. *Revalidatiemagazine*, *4*, 18-19.
22. Crombez, G., Eelen, P., & Baeyens, F. (1990). Het 'match/mismatch'-model van vrees en pijn: kritische bedenkingen. *Gedragstherapie, 3*, 195-201.
23. Arntz, A. (1996). Why do people tend to overpredict pain? On the asymmetries between underpredictions and overpredictions of pain. *Behav Res Ther*, *34*, 545-554.
24. Arntz, A., Van den Hout, A., Van den Berg, G., & Meijboom, A. (1991). The effects of incorrect pain expectations on acquired fear and pain responses. *Behav Res Ther, 29*, 547-560.
25. Crombez, G., Baeyens, F., & Eelen, P. (1994). Sensory and temporal information about impending pain: the influence of predictability on pain. *Behav res Ther, 32*, 611-622.
26. Arntz, A., & Hopmans, M. (1998). Underpredicted pain disrupts more than correctly predicted pain, but does not hurt more. *Behav Res Ther*, *36*, 1121-1129.
27. Philips, H.C. (1987). Avoidance behaviour and its role in sustaining chronic pain. *Behav Res Ther, 25*, 273-279.
28. Vlaeyen, J.W.S., Kole-Snijders, A.M.J., Boeren, R.G.B., & Eek, H. Van (1995). Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain*, *62*, 363-372.
29. Hasenbring, M., Hallner, D., & Klasen, B. (2001). Psychologische Mechanismen im Prozess der Schmerzchronifizierung. *Schmerz*, *15*, 442-447.
30. Crombez, G., Baeyens, F., Vervaet, L., & Vlaeyen, J. (1997). Vermijdingsgedrag bij patiënten met chronische lage rugpijn. *Gedrag Gezond, 25*, 126-135.
31. McCracken, L.M., Gross, R.T., Sorg, P.J., & Edmands, T.A. (1993). Prediction of pain in patients with chronic low back pain: effects of inaccurate prediction and pain-related anxiety. *Behav Res Ther, 31*, 647-652.
32. Nederhand, M.J., Hermens, H.J., IJzerman, M.J., Turk, D.C., & Zilvold, G. (2003). Cervical muscle dysfunction in chronic whiplash-associated disorder grade 2: the relevance of the trauma. *Spine, 27*, 1056-1061.
33. Kallenberg, L.A., & Hermens, H.J., (2006). Motor unit action potential rate and motor unit action potential shape properties in subjects with work-related chronic pain. *Eur J Appl Physiol*, *96*, 203-208.
34. Lousberg, R., Vuurman, E., Lamers, T., Breukelen, G. Van, Jongen, E., Rijnen, H., Maessen, C. & Hermens, H. (2005). *Clin J Pain, 21*, 262-271.

35. Flor, H., Kerns, R.D., & Turk, D.C. (1987). The role of spouse reinforcement, perceived pain, and activity levels of chronic pain patients. *J Psychosom Res, 31*, 251-259.
36. Lousberg, R., Schmidt, A.J., & Groenman, N.H. (1992). The relationship between spouse solicitousness and pain behavior: searching for more experimental evidence. *Pain*, *51*, 75-79.
37. Delespaul, P.A.E.G. (1995). *Assessing schizophrenia in daily life: The Experience Sampling Method*. Maastricht: University Press.
38. Vries, M.W. de, & Delespaul, P.A.E.G. (1993). Zelfobservatie en intensieve time-sampling in de psychiatrie. In C.A.L. Hoogduin, P. Schnabel, W. Vandereycken, K. van der Velden, & F.C. Verhulst (Eds.). *Jaarboek psychiatrie en psychotherapie*. Meppel: Van Loghum Slaterus, p.94-108.
39. Cruise, C.E., Broderick, J., Porter, L., Kaell, A., & Stone, A.A. (1996). Reactive effect of diary self-assessment in chronic pain patients. *Pain, 67,* 253-258.
40. Vollenbroek Hutten, M.M., Hermens H.J., & Huis In't Veld, R.M. (2004). Context aware tele-monitoring and tele-treatment services*. Stud Health Technol Inform, 114*, 230-234.

SUMMARY

The outcome of almost all chronic pain treatment programs is (partly) non-optimal. In most treatment programs, the greater majority of patients do show improvement after treatment. Despite these successes, there are almost always patients who show no improvement or who even deteriorate in functioning after completion of the treatment program. Failing significant improvements in treatment programs, this kind of variability in treatment results will remain. In this dissertation, the treatment variability of the cognitive-behavioral therapy (CBT-R) program at the Roessingh Center for Rehabilitation (RCR) is demonstrated through a randomized controlled trial (RCT) (Chapter 2). Additionally, an attempt is made to explain treatment variability of this program by: 1) determining whether treatment variability is the result of the existence of subgroups of chronic pain patients; 2) relating treatment variability to factors in the treatment process; and 3) investigating whether treatment variability is attributable to insufficient or incorrect underlying theoretical mechanisms. For the latter, the generalizability and value of the match-mismatch (MM) mechanism in explaining the maintenance of chronic pain is explored. Additionally, the influence of the CBT-R program on pain expectations in daily life is examined. The search for explanations for the existing variability in treatment results of the CBT-R program is imperative if we want to be able to identify what needs to be changed in order to improve the overall treatment effect of the CBT-R program.

In Chapter 3, an attempt to explain the variability of the treatment results of the CBT-R program is made by investigating the differential treatment responses of subgroups of pain patients as determined with the Multidimensional Pain Inventory-Dutch Language Version (MPI-DLV). Based on the results of the analyses, we can conclude that indeed a differential treatment response between clusters of patients exists. It appears that the differential cluster response is dependent upon not only the mode of treatment patients receive but also the specific measurement scale utilized. While Interpersonally Distressed patients seem to profit more from outpatient treatment, Dysfunctional patients seem to benefit more from the semi-inpatient treatment mode. These differences emerge particularly on the MPI-DLV response scales. Firstly, Interpersonally Distressed patients in the outpatient treatment program appear to improve more on MPI-DLV Punishing and Solicitous Responses, while Interpersonally Distressed patients in

the semi-inpatient treatment program appear to improve more on MPI-DLV Distracting Responses. Secondly, Dysfunctional patients in the semi-inpatient mode benefit most on MPI-DLV Solicitous Responses and, to a lesser extent, on MPI-DLV Punishing Responses. Thirdly, Interpersonally Distressed patients improve most on MPI-DLV Life Control and Affective Distress. Fourthly, Average patients show a positive change with regard to the SCL-90 variable Psychoneuroticism and the RAND-36 variables Social Functioning and Mental Health. This effect is greater for patients in the semi-inpatient treatment program than for patients in the outpatient treatment program.

Although some differential treatment effects are obvious immediately following treatment, we can expect that the differences in treatment effect between the MPI-DLV clusters will become more visible over time. Measuring differences in treatment effect immediately following the program may underestimate the actual differences as the CBT-R program aims to establish behavioral change and behavioral chance is a long-term process that may not be recognizable or sufficiently established after only eight weeks of treatment. However, the subgroup specific study demonstrates that the treatment results are less significant and, in some cases, even worse when measured at a three month follow-up. However, given that the CBT-R program is not adapted to the specific characteristics of the different MPI-DLV clusters, we can hardly expect to find a large differential treatment effect. The diminishment of the differential cluster effects at follow-up, and in particular the decrease in effects on the MPI-DLV response scales three months after treatment, may be the result of difficulties experienced by the patient with respect to retaining newly learned behavior in the home environment where, among other factors, the influence of significant others is present.

The fact that differential cluster effects are found supports the idea that customizing the CBT-R program to differential cluster effects may improve the overall treatment effectiveness of the program. Consequently, taking a patient's cluster type into account when determining which treatment program is most appropriate is important.

Chapter 4 investigates the degree to which treatment variability can be explained by factors in the treatment process. The congruence between the planned treatment, as described in the treatment protocol, and the actual treatment implemented in daily practice is investigated. Unfortunately, the existing protocol contains insufficient underlying theory for treatment strength, treatment aims and activities.

Additionally, the protocol fails to fully operationalize selection criteria, indication criteria and the aims of treatment. Further, no description or guidelines for measuring treatment outcome are provided. As a result of these shortcomings in the protocol, determining the congruence between the planned and the actual treatment implemented has proven to be impossible. More importantly, a poorly described and operationalized protocol will almost inevitably generate substantial variability with respect to treatment effect. Nonetheless, the greater majority of the program content and goals do correspond with the existing protocol. The differences between the planned program and the program that was actually implemented relate mostly to a) the cancellation of certain therapy sessions due to the absence of therapists; and b) double-booking of some therapy sessions so that two components of the program were offered at the same time. Additionally, some program adaptations occur in accordance with the specific needs and requests of the rehabilitants. It is important to note that unexpected differences in treatment effect between semi-inpatient groups have been found. The only differences that ought to exist between these groups is the location of treatment in the rehabilitation center and the involved therapists, both of which are assumed not be of influence on treatment outcome.

The differences between the planned treatment program and the actual treatment program may be partly responsible for the failure of some patients to gain a significant or clinically relevant therapy effect and thus may explain some of the variability in the overall and subgroup specific treatment results. However, before any conclusions can be drawn with regard to possible program or theory failures, it is imperative that the treatment process is better understood. The treatment process must be extensively described and operationalized. We must also take a closer look at the congruence between the planned program and the actual treatment situation. Only then can we link factors in the treatment process to treatment effect and variability. This is of significant importance in a multidisciplinary treatment environment as the involvement of several disciplines in the treatment process makes it difficult to contribute the overall treatment outcome to specific discipline-related activities.

It is also possible that treatment variability is partly or wholly attributable to insufficient or incorrect underlying theoretical mechanisms. To investigate this possibility, the match-mismatch (MM) model of pain was utilized.. The MM model contends that individuals make predictions of future pain by evaluating previous

painful events. When an event is perceived as being more painful than anticipated (an underprediction), the expectation for the next event is that the pain will be greater. Conversely, when the preceding event is less painful than anticipated (an overprediction), the expectation is adjusted in such a way that the anticipated pain severity is less than the previous expectation. When the expectation for the anticipated pain severity is accurate (a match), no changes in future pain expectations occur.

The MM model has been studied extensively but mostly in laboratory settings. Chapter 5 explores whether the MM model can be generalized to the daily life situations of chronic pain patients. For this study, the Experience Sampling Method (ESM) was applied. One hundred chronic pain patients carried a paging device that randomly prompted them 10 times a day over the course of seven days. Upon being prompted, patients answered a number of questions about their pain experiences, their pain expectations and some contextual aspects. The results of this study indicate that underpredictions were followed by significant increases in predicted pain as well as experienced pain. Overpredictions were followed by significant decreases in predicted pain as well as a trend towards a decrease in pain experience. Matches were mainly followed by no changes in pain experience. The results of the ESM study on the MM model shows that chronic pain patients are more likely to underpredict than to overpredict their pain. Consequently, the overall increase of pain experience due to underpredictions during the day is larger than the overall decrease of pain due to overpredictions. We can assume that the mismatch effects in this study reflect a 'real' psychological mechanism in the daily life of chronic pain patients. This study also shows how pain complaints can be maintained by means of mismatches between pain prediction and pain experience.

In Chapter 6, the influence of the CBT-R program on the MM mechanism is measured on the basis of the same data set. Only data from patients who had a valid measurement prior to and following completion of the CBT-R program were included (n=43). The results of this study demonstrate that after completion of the CBT-R program, patients showed a significant decrease in the number, size and effect of mismatches. The effect of overpredictions appears to be more resistant to the influence of treatment. Based on these results, we can conclude that the CBT-R program has a positive effect on changing expectations of pain, despite the fact that this effect is not an explicit aim or goal of the program.

Chapter 7 presents an additional study on the MM model. Here, the tendency of chronic pain patients to make more underpredictions than

overpredictions is explored. It is important to understand why this tendency exists as underpredictions may have negative effects, such as an increased fear of pain, greater fear of movement, more avoidance behavior and, most importantly, increased pain. As a result, explanations for the discrepancy between the number of underpredictions versus the number of overpredictions reported by chronic pain patients were sought. It was hypothesized that the variables pain severity, fear-avoidance, pain control, and emotional distress can explain this discrepancy. Using the ESM data from 83 chronic pain patients, together with the corresponding MPI-DLV data and data from the Tampa Scale for Kinesiophobia for these patients, analyses were conducted. The results of a path analysis show that there is a trend for fear-avoidance to be a direct and independent factor predicting the discrepancy between the number of underpredictions and overpredictions. Given that the data are cross-sectional, causal interpretations of the modeled relationships cannot be made. Future experimental and, in particular, prospective studies are necessary to elucidate the causal mechanisms between the investigated variables.

Despite the fact that several indications were found that can guide a process of improvement of the CBT-R program, the issue of variability in treatment results of the CBT-R program remains unresolved. To decrease treatment variability in not only the CBT-R program at the RCR but also other pain management programs, it is imperative that we, first and foremost, extensively investigate, describe and thus understand: a) the treatment itself: it's components, duration, frequency, and intensity; b) the aims of treatment; and c) the implementation-environment. This means that all these domains and their underlying theoretical assumptions should be clear, objective and operationalized. We must, secondly, determine which instruments are most appropriate for measuring the outcome of the CBT-R program. Thirdly, we must extensively evaluate the treatment process. Ineffectiveness of a treatment program can be caused by either program failures or theoretical failures. It is imperative that we investigate program failure first. When we are able to rule out incongruencies between the planned treatment program and the actual treatment situation (a program failure) as an explanation for treatment ineffectiveness, we can then investigate the theoretical background of the program. Once we are that far that program or theory failures can be ruled out, treatment effect can be established by means of the selected instruments.

In addition to alternately evaluating the treatment process and treatment efficacy and thereby unraveling the active, effective components of the CBT-R

program, studies on the underlying theoretical mechanisms that may explain the development and maintenance of chronic pain in different patients or patient subgroups are recommended. The knowledge obtained with these studies can then be applied to, firstly, improve the intake and selection procedures for pain treatment and, secondly, customize and tailor the treatment program to the specific characteristics and profiles of pain patients. Ultimately, the optimized selection procedures and tailored treatments may lead to an improvement in the overall effect of chronic pain treatment.

SAMENVATTING

Het resultaat van de meeste behandelprogramma's voor chronische pijn is (deels) niet optimaal: behalve dat het merendeel van de patiënten verbeteren na afloop van behandeling, zijn er ook patiënten die geen verandering tonen of zelfs achteruit gaan na behandeling. Wanneer bestaande behandelprogramma's niet verbeterd worden, zal deze variabiliteit in behandelresultaat blijven bestaan. In dit proefschrift wordt de variabiliteit van het cognitief-gedragsmatige pijnrevalidatieprogramma (CBT-R) van Revalidatiecentrum Het Roessingh (RCR) aangetoond aan de hand van de resultaten van een randomized controlled trial (RCT) (hoofdstuk 2). Daarnaast wordt op drie manieren geprobeerd om een verklaring te vinden voor deze variabiliteit: 1) door te bepalen of de variabiliteit verklaard kan worden op basis van de aanwezigheid van subgroepen van patiënten; 2) door de variabiliteit te relateren aan factoren in het behandelproces; en 3) door te onderzoeken of de variabiliteit toegeschreven kan worden aan onvolledige of incorrecte onderliggende theoretische mechanismen. Voor dit laatste wordt de generaliseerbaarheid en waarde bepaald van het match-mismatch (MM) mechanisme van pijn als verklaringsmodel voor het voortduren van chronische pijn. Tevens wordt de invloed van het CBT-R programma op pijnschattingen in het dagelijks leven onderzocht. Het zoeken naar verklaringen voor de variabiliteit van het behandelresultaat van het CBT-R programma is noodzakelijk als we aangrijpingspunten voor aanpassing van dit programma willen achterhalen en het algemene behandeleffect van dit programma willen vergroten.

In hoofdstuk 3 wordt geprobeerd om de variabiliteit van het behandelresultaat van het CBT-R programma te verklaren op basis van verschillen in behandelrespons tussen subgroepen van pijnpatiënten zoals bepaald met de Multidimensionele Pijn Vragenlijst (MPI-DLV). Op basis van de resultaten van deze analyses kan geconcludeerd worden dat er inderdaad een differentiële behandelrespons bestaat tussen clusters van patiënten. Deze differentiële respons blijkt niet alleen afhankelijk te zijn van de behandelvorm, maar ook van de specifieke meetschaal. Terwijl 'Interpersonally Distressed' patiënten meer lijken te profiteren van de poliklinische behandeling, verbeteren de 'Dysfunctional' patiënten meer in de semiklinische behandeling. Deze verschillen komen met name tot uiting op de MPI-DLV respons schalen: poliklinische 'Interpersonally Distressed' patiënten verbeteren meer ten aanzien van straffende en bezorgde responsen, terwijl

semiklinische ‘Interpersonally Distressed’ patiënten meer verbeteren ten aanzien van afleidende responsen. Semiklinische ‘Dysfunctional’ patiënten profiteren meer op bezorgde responsen en voor een kleiner deel op straffende responsen. ‘Interpersonally Distressed’ patiënten verbeteren het meest op MPI-DLV levenscontrole en negatieve stemming. ‘Average’ patiënten, met name van de semiklinische behandeling, behalen over het algemeen positieve veranderingen ten aanzien van SCL-90 psychoneuroticisme, RAND-36 sociaal functioneren en mentale gezondheid.

Hoewel er duidelijk differentieel behandelresultaat bestaat direct na afloop van behandeling, kan verwacht worden dat de verschillen in behandeleffect tussen de MPI-DLV clusters op de langere termijn meer zichtbaar zullen worden. Door de verschillen in behandeleffect direct na afloop van het programma te bepalen, kan het zijn dat aanwezige verschillen onderschat worden, aangezien het CBT-R programma gericht is op gedragsverandering en gedragsverandering een langdurig proces is dat niet herkend of onvoldoende vastgesteld kan worden na slechts acht weken van behandeling. De subgroepspecifiek studie laat echter zien dat de behandelresultaten bij drie maanden follow-up minder significant en in sommige gevallen zelfs slechter zijn. Aangezien het CBT-R programma niet aangepast is aan de specifieke kenmerken van de verschillende MPI-DLV clusters, kan ook bijna niet verwacht worden dat een groot differentieel behandeleffect gevonden zal worden. Het feit dat de clustereffecten bij follow-up zijn afgenomen, vooral de effecten op de MPI-DLV respons schalen, kan veroorzaakt zijn door het feit dat het moeilijk is om nieuw aangeleerd gedrag vast te houden in de thuisomgeving, waar met name de invloed van significante anderen aanwezig is.

Het feit dat differentiële clustereffecten gevonden zijn, ondersteunt het idee dat het aanpassen van het CBT-R programma aan deze differentiële clustereffecten kan leiden tot een verbetering van de algehele effectiviteit van dit programma. Om deze reden is het van belang dat rekening wordt gehouden met het clustertype van een patiënt bij het bepalen van het meest geschikte behandelprogramma.

In hoofdstuk 4 wordt onderzocht in welke mate de variabiliteit van het CBT-R programma verklaard kan worden door factoren in het behandelproces. De congruentie wordt bepaald tussen het geplande CBT-R programma, zoals omschreven in het behandelprotocol, en het programma zoals daadwerkelijk geïmplementeerd in de dagelijkse praktijk. Helaas is het protocol onvolledig voor wat betreft onderliggende theorie voor behandelsterkte, behandeldoelen en

activiteiten. Daarnaast zijn de selectie- en indicatiecriteria, alsook de doelstellingen van behandeling nauwelijks geoperationaliseerd en wordt niet beschreven hoe het behandelresultaat gemeten kan worden. Als gevolg van deze tekortkomingen is het bij voorbaat onmogelijk om de congruentie tussen planning en praktijk adequaat te bepalen en, veel belangrijker, is het haast onontkoombaar dat een substantiële variabiliteit in behandelresultaat optreedt.

Niettemin correspondeert de actuele inhoud en doelstelling van het CBT-R programma grotendeels met het protocol. Een paar verschillen zijn gevonden die veelal betrekking hebben op uitval of dubbelplanning van behandelsessies. Daarnaast wordt het programma soms enigszins aangepast aan de specifieke behoeften en wensen van de revalidanten. Verder zijn verschillen in effect gevonden tussen de semiklinische behandelgroepen. Tussen deze groepen mag normaliter geen verschil bestaan, behalve voor wat betreft de behandellocatie en de betrokken therapeuten, maar verondersteld wordt dat deze verschillen niet van invloed mogen zijn op het behandelresultaat.

De verschillen tussen het geplande en het daadwerkelijk uitgevoerde CBT-R programma kunnen (deels) verantwoordelijk zijn voor het feit dat sommige patiënten er niet in slagen om een significant of klinisch relevant behandeleffect te bereiken, en daarmee een verklaring bieden voor de variabiliteit van de totale en subgroepspecifieke resultaten van het CBT-R programma. Voordat conclusies getrokken kunnen worden met betrekking tot eventuele programma- of theoriefouten van het CBT-R programma, is het nodig dat het behandelproces van dit programma verder wordt uitgewerkt, onderbouwd en geoperationaliseerd en dat de congruentie tussen planning en praktijk uitgebreider wordt bepaald. Pas dan kunnen factoren in het behandelproces gerelateerd worden aan het behandeleffect en de variabiliteit in resultaat. Dit is in het bijzonder van belang in een multidimensionele behandelomgeving, aangezien de betrokkenheid van meerdere disciplines in het behandelproces het moeilijker maakt om het behandelresultaat toe te schrijven aan disciplinegerelateerde activiteiten.

Het is mogelijk dat de variabiliteit in behandelresultaat geheel of gedeeltelijk toe te schrijven is aan onvolledige en incorrecte onderliggende theoretisch mechanismen. Om deze mogelijkheid te onderzoeken, werd het match-mismatch (MM) model van pijn onderzocht. Het MM model stelt dat mensen voorspellingen doen over pijn door eerdere pijnlijke gebeurtenissen te evalueren. Wanneer een gebeurtenis ervaren wordt als pijnlijker dan geanticipeerd (een onderpredictie), neemt de

verwachting van pijn voor de volgende gebeurtenis toe. Omgekeerd, wanneer een gebeurtenis als minder pijnlijk wordt ervaren dan verwacht (een overpredictie), zal de verwachting van pijn voor de volgende gebeurtenis afnemen. Wanneer de pijnverwachting voor een gebeurtenis accuraat is (een match), stelt het model dat er geen verandering in pijnverwachting plaatsvindt.

Het MM model is uitgebreid onderzocht, voornamelijk in laboratoriumsituaties. In hoofdstuk 5 wordt onderzocht of het MM model ook gegeneraliseerd kan worden naar het dagelijkse leven van chronische pijnpatiënten. Hiervoor wordt gebruik gemaakt van de Experience Sampling Methode (ESM). Honderd patiënten droegen een horloge dat gedurende 7 dagen op 10 willekeurige momenten per dag een alarmsignaal gaf. Na dit signaal dienden zij een aantal vragen met betrekking tot pijnervaring, pijnverwachting en situationele aspecten te beantwoorden. De resultaten van deze studie wijzen erop uit dat onderpredicties van pijn gevolgd worden door significante toenames van niet alleen pijnverwachting, maar ook van pijnervaring. Overpredicties worden gevolgd door een significante afname van de pijnverwachting, alsook door een trend richting afname van pijnervaring. Matches worden voornamelijk gevolgd door geen veranderingen in pijnervaring. De resultaten van de ESM studie naar het MM model laten daarnaast zien dat chronische pijnpatiënten meer geneigd zijn om hun pijn te onderschatten. Als gevolg hiervan is de totale toename van pijn als gevolg van de onderpredicties over de dag heen groter dan de totale afname van pijn als gevolg van overpredicties. Verondersteld mag worden dat de mismatch effecten in deze studie een werkelijk psychologisch mechanisme vormen in het dagelijkse leven van chronische pijnpatiënten. De studie laat ook zien hoe pijnklachten in stand kunnen worden gehouden door middel van misschattingen tussen pijnverwachting en pijnervaring.

In hoofdstuk 6 wordt de invloed van het CBT-R programma op het MM mechanisme bepaald op basis van dezelfde dataset, waarbij alleen die data worden meegenomen van patiënten die zowel voor als na behandeling een valide meting hebben (n=43). De resultaten van deze studie tonen dat de patiënten na afloop van het CBT-R programma een significante afname laten zien van het aantal, de grootte en het effect van mismatches. Het effect van overpredicties blijkt meer resistent te zijn voor de invloed van het CBT-R programma. Op basis van deze resultaten kan geconcludeerd worden dat het CBT-R programma, hoewel niet beoogd, een positief effect heeft op de verandering van pijnverwachtingen.

Hoofstuk 7 presenteert een aanvullende studie naar het MM model. In deze studie wordt onderzocht waarom pijnpatiënten meer onderpredicties dan overpredicties maken. Het is van belang de oorzaken hiervoor te achterhalen, aangezien onderpredicties negatieve gevolgen kunnen hebben zoals een toename van angst voor pijn en beweging, alsook vermijdingsgedrag en toename van pijn. Om deze reden worden verklaringen gezocht voor de discrepantie tussen het aantal onderpredicties versus het aantal overpredicties. Verondersteld wordt dat de variabelen pijnintensiteit, vermijding als gevolg van angst (fear-avoidance), pijncontrole en emotionele 'distress' verklarende factoren kunnen zijn voor deze discrepantie. Gebruik wordt gemaakt van de ESM data van 83 patiënten, alsook de resultaten van deze patiënten op de MPI-DLV en de Tampaschaal voor Kinesiofobie. De resultaten van een padanalyse tonen dat er een trend is voor fear-avoidance als directe en onafhankelijk voorspeller voor de discrepantie tussen het aantal onderpredicties en overpredicties. Aangezien de data cross-sectioneel zijn, kunnen geen causale interpretaties van de gemodelleerde relaties worden gemaakt. Toekomstige experimentele, met name prospectieve studies zijn nodig om de causale mechanismen tussen de onderzochte variabelen te verhelderen.

Op basis van de resultaten van de studies in dit proefschrift kan geconcludeerd worden dat, ondanks dat een aantal aangrijpingspunten voor verbetering van het CBT-R programma gevonden zijn, het probleem van de variabiliteit van dit programma nog niet is opgelost. Om deze variabiliteit, alsook die van andere pijnmanagementprogramma's te verbeteren, is het ten eerste noodzakelijk dat het behandelprogramma beter in kaart wordt gebracht ten aanzien van de behandeling zelf (componenten, duur, frequentie en intensiteit), de doelstellingen, alsook de implementatie-omgeving. Dit betekent dat al deze domeinen en hun onderliggende theoretische assumpties duidelijk, objectief en geoperationaliseerd moeten zijn. Ten tweede moeten instrumenten gekozen worden, aansluitend op de (sub)doelstellingen van het programma, die geschikt zijn voor het meten van het resultaat van het CBT-R programma. Ten derde moet het behandelproces uitgebreid worden geëvalueerd. Aangezien (in)effectiviteit van behandeling het gevolg kan zijn van programmafouten dan wel theoriefouten, moeten mogelijke incongruenties tussen geplande en daadwerkelijke behandeling eerst worden uitgesloten, voordat de theoretische achtergrond van het programma in twijfel kan worden getrokken. Zodra programmafouten of theoriefouten kunnen worden

uitgesloten, kan het behandeleffect worden vastgesteld door middel van de geselecteerde instrumenten.

Naast het afwisselend evalueren van het behandelproces en het behandelresultaat en daarbij het ‘ontrafelen’ van de actieve, effectieve componenten van het CBT-R programma, wordt aanbevolen om tegelijkertijd meer studies te doen naar de onderliggende theoretische mechanismen die, bij verschillende patiënten of patiëntgroepen, een verklaring bieden voor de ontwikkeling en instandhouding van chronische pijn. De kennis die met deze studies wordt verkregen, kan dan toegepast worden om ten eerste de intake- en selectieprocedures voor pijnbehandeling te verbeteren, en ten tweede om het behandelprogramma’s aan te passen (op maat maken) aan de specifieke kenmerken en profielen van pijnpatiënten. Uiteindelijk zullen de geoptimaliseerde selectieprocedures en de op maat gesneden behandelprogramma’s kunnen leiden tot een verbetering van het algehele effect van chronische pijnbehandeling.

DANKWOORD

Een groot deel van dit proefschrift gaat over verschillen, oftewel mismatches, tussen verwachtingen en ervaringen van pijn. Ook ten aanzien van de totstandkoming van dit proefschrift heb ik, waarschijnlijk net als vele andere promovendi, menige mismatch ervaren. Toen ik in 1999 met de RCT begon, was het plan om binnen 2 á 3 jaar te promoveren. Dit bleek de eerste mismatch te zijn. Een aantal jaar heb ik telkens opnieuw aangegeven dat het 'volgend jaar wel zou gaan gebeuren'. Nu ik dan straks op 12 januari 2007 mijn proefschrift ga verdedigen, ben ik blij dat het 'volgend jaar' is en dat het voor me liggende proefschrift een match vormt met mijn oorspronkelijke verwachting. Daarnaast ben ik blij dat ik met dit proefschrift als sluitstuk mijn arbeidsperiode bij Roessingh Research & Development (RRD) op een tevreden manier kan beëindigen.

Vele personen zijn, direct en indirect, in werk en privé, betrokken geweest bij mijn promotieonderzoek. Graag wil ik ieder van hen hier voor bedanken. Op de eerste plaats noem ik hier dr. R. Lousberg. Beste Richel, de afgelopen jaren heb je mij als assistent-promotor begeleid en aangestuurd bij de opzet van de RCT en de ESM studies, de analyse en interpretatie van data en het schrijven van het proefschrift. Je ideeën, je enthousiasme, het gevoel met elkaar op één lijn te zitten en je geloof in mij als onderzoeker (lees: niet wetenschapper) werkten daarbij aanstekelijk, inspirerend en vooral ook motiverend. Wat hebben we veel dagen samen doorgebracht, brainstormend en analyserend, in Enschede, Maastricht en ook in Bunde bij je thuis, waar ik zelfs nog met de laptop aan je bed heb gezeten toen je plat lag vanwege een hernia. Richel, ik had me geen fijnere assistent-promotor kunnen wensen. Mede dankzij jou is dit proefschrift geworden tot wat het is. Ontzettend bedankt voor je begeleiding, je inzet en je enthousiasme! Ook dank aan je vrouw Birgit en je kinderen Jelle en Annika, voor hun gezelligheid, de vele kopjes thee en de lekkere lunches.

Prof.dr. G. Zilvold dank ik voor zijn betrokkenheid als promotor. Beste Gerrit, de afgelopen jaren heb je als motor achter mijn promotie gefungeerd door middel van facilitering, de state-of-the-art gesprekken en je onvoorwaardelijke vertrouwen in en optimisme over mijn functioneren. Je bent ondertussen al met pensioen en ik voel me dan ook vereerd dat ik je laatste promovendus mag zijn.

Ook mijn co-promotor, prof.dr. D.C. Turk, wil ik danken voor zijn betrokkenheid. Dear Dennis, I have never seen someone responding to my

manuscripts and documents as quick as you did. When I sent you a manuscript to read in the afternoon, a complete revision was emailed back the same night. Thank you for all your comments and suggestions with regard to the contents of my thesis.

Mijn promotieonderzoek was niet mogelijk geweest zonder de faciliterende rol van het managementteam van RRD: prof. dr. M.J. IJzerman, dr. M.M.R. Vollenbroek-Hutten, prof.dr. H.J. Hermens, dr. M.J.A. Jannink en drs. B. Freriks. Beste Maarten, Miriam, Hermie, Michiel en Bart, met veel plezier en waardering kijk ik terug op de tijd die ik bij RRD heb doorgebracht, waarin ik mij als onderzoeker heb kunnen ontwikkelen en veel wetenschappelijke kennis en ervaring heb opgedaan. Ik dank jullie voor de kansen en mogelijkheden die jullie mij geboden hebben tijdens mijn arbeidsperiode bij RRD, alsook de welwillendheid om voor mij de weg terug naar het zuiden mogelijk te maken. Miriam, fijn dat ik de tijd kreeg om mijn proefschrift bij RRD af te ronden. Ook al zijn er de afgelopen jaren best wat tranen gevallen tijdens onze werkoverleggen, toch denk ik hier met plezier aan terug. Dank voor je inhoudelijke opmerkingen, je persoonlijke betrokkenheid en ondersteuning als clustermanager en het vertrouwen dat je in me hebt gesteld.

Prof.dr. W.H. Van Harten, beste Wim, je hebt mij in 1997 een functie geboden als wetenschappelijk medewerker van Revalidatiecentrum Het Roessingh. Als clustermanager ‘Kwaliteit van zorg’ binnen RRD gaf je me in 1999 de mogelijkheid om een RCT op te zetten om de effectiviteit van pijnrevalidatie te bepalen. Hiermee heb je tegelijkertijd de initiële stap gezet in de vormgeving van mijn promotieonderzoek. Hiervoor mijn dank, alsook voor je betrokkenheid toentertijd als clustermanager en je bereidwilligheid om in mijn promotiecommissie zitting te nemen en mijn proefschrift te beoordelen.

Dank ook aan dr. K.M.G. Schreurs, dr. P.A.E.G. Delespaul en prof.dr. M.I. Hasenbring voor de tijd en energie die zij gestoken hebben in het beoordelen van mijn proefschrift. Beste Karlein, ik ben blij dat je met mijn RCT data verder aan de slag bent gegaan. Dankzij jouw post hoc analyses en de veranderingen die de pijnprogramma’s ondertussen hebben ondergaan, weet ik dat de resultaten van mijn promotieonderzoek ook daadwerkelijk een toepassing in de praktijk gevonden hebben. Beste Philippe, mede dankzij jou heb ik ontdekt hoe waardevol experience sampling als onderzoeksmethode kan zijn. Dank voor je ondersteuning bij het analyseren en interpreteren van de data.

Een onderzoek voer je meestal niet alleen uit. Daarom wil ik hier ook alle patiënten bedanken die aan de metingen hebben deelgenomen en keer op keer vragenlijsten hebben ingevuld. Alle medewerkers en therapeuten van de pijndivisie van Het Roessingh dank ik voor hun medewerking en het praktisch mogelijk maken van de verschillende deelprojecten. Daarnaast hebben verschillende stagiaires meegeholpen om de data voor dit promotieonderzoek te verzamelen en te analyseren: Jitske Gulmans, Mandy Bergman en Idwine van der Blij, dank jullie wel voor jullie inzet en bijdrage. Dank ook aan Karin Groothuis voor de ondersteuning bij de (multilevel) analyses; Sarah Stutterheim voor de zorgvuldige correctie op het engels; en Wouter van Ek voor het mede vormgeven van de omslag van dit proefschrift.

Een speciaal woord van dank is voor Jan Otter. Helaas heb je de afronding van mijn proefschrift niet meer mee kunnen maken, Jan. Toch wil ik je graag bedanken voor de bijzondere gesprekken die ik met je mocht hebben. Met veel plezier en herkenning heb ik naar je geluisterd tijdens de sessies waarin jij over ‘jouw behandeling’ vertelde.

In negen jaar RRD tijd heb ik verschillende andere collega’s leren kennen met wie ik in meer of mindere mate lief en leed heb gedeeld en die maakten dat ik graag bij RRD aan het werk was. Zonder iemand hierbij te willen overslaan, noem ik hier: Victorien Erren, hoewel ik het erg jammer vind dat je mijn paranimf niet kan zijn, ben ik toch blij voor je dat je waarschijnlijk niet bij mijn promotie aanwezig bent. Anke Kottink en Yvette Bulthuis, mijn paranimfen en ex-kamergenoten, mijn andere (ex)kamergenoten Marc Nederhand, Anand Nene, Barbara Huijgen en heel even ook Margit Schlecht: ik dank jullie, alsook al mijn lieve vriend(inn)en voor jullie support, gezelligheid en luisterende oren.

Lieve papa en mama, zussen en zwagers, dank voor al jullie steun! Mede dankzij jullie aanmoediging en bemoedigende woorden “geef niet op, want nog even en je bent er” is het nu ‘volgend jaar’ en kan ik trots zijn dat ik mijn proefschrift heb afgerond. Ik ben blij te voelen dat jullie achter me staan, welke keus ook wordt gemaakt. Last, maar verre van least: lieve Ramon, wat fantastisch dat wij elkaar opnieuw hebben leren kennen. Dank je wel dat je er voor me bent! De afgelopen twee jaren waren niet altijd even makkelijk met het vele op en neer gereis, de twijfel om met het onderzoek door te gaan, de verhuizing en het zoeken naar een nieuwe baan. Nu kijk ik met veel plezier uit naar de nieuwe stappen in mijn en ons leven. (Aan)komend oud en nieuw ga ik mooi weer andere voornemens maken!

OVER DE AUTEUR

Carola Mes werd geboren op 17 januari 1973 te Eindhoven. De middelbare school volgde zij op het Strabrecht College te Geldrop, alwaar zij in 1992 haar VWO diploma behaalde. In 1992 begon zij met haar studie Bewegingswetenschappen aan de Rijksuniversiteit Groningen, waarbij zij koos voor de afstudeerrichting Revalidatie en Gehandicaptenzorg. Tijdens de studie Bewegingswetenschappen volgde Carola tevens de éénjarige opleiding 'European Master Degree in Adapted Physical Activity (APA)' aan de Katholieke Universiteit van Leuven, België. Tijdens deze opleiding specialiseerde zij zich in de richting Psychosociale Stoornissen. Voor beide opleidingen werd het afstudeerproject verricht bij Revalidatiecentrum Het Roessingh te Enschede. Dit project bestond uit een effect- en een procesevaluatie van de fysieke componenten van het 'Pijn de Baas' programma. In augustus 1997 behaalde Carola zowel het doctoraaldiploma Bewegingswetenschappen, alsook de master degree in APA, waarna zij per november 1997 in dienst kwam bij Revalidatiecentrum Het Roessingh voor de uitvoer van een behoefteonderzoek onder mensen met het Chronische Vermoeidheidssyndroom.

Sinds september 1998 is Carola werkzaam binnen het cluster Technology Assisted Pain Rehabilitation (voorheen Zorginnovatie) van Roessingh Research & Development (RRD), een onderzoeksinstituut gelieerd aan Revalidatiecentrum Het Roessingh ten Enschede. Gedurende deze periode voerde zij diverse onderzoeks- en implementatieprojecten uit op het gebied van chronische pijn en vermoeidheid. Daarnaast was Carola 5 jaar lid van de ondernemingsraad van RRD, waarvan 2 jaar als voorzitter, alsook van de centrale ondernemingsraad van de Roessingh organisatie. Ook leverde zij een bijdrage aan het opstellen van het beleid voor het pijnonderzoek binnen RRD, het Innovatiecentrum voor Pijnrevalidatie, de organisatie van symposia, alsook aan de organisatie van diverse personeelsfeesten. In 1999 startte zij met haar promotieonderzoek getiteld "Improving non-optimal results in chronic pain treatment: a tripartite approach". Op 12 januari 2007 promoveert zij op dit onderzoek. Per 1 januari 2007 is Carola werkzaam bij het PON Instituut voor Advies, Onderzoek en Ontwikkeling te Tilburg.

Publikaties

- Mes, C.A.J. Lousberg, R., Zilvold, G., Delespaul, P.A.E.G., & Turk, D.C. (2006). Application of the match-mismatch model in chronic pain treatment: an experience sampling study. *European Journal of Pain,* 10, S251.
- Schreurs, K.M.G., Mes, C.A.J., Witting, M., & Vollenbroek-Hutten, M.M.R. (2006) Measuring daily activity in fibromyalgia: Is an accelerometer a useful tool? *European Journal of Pain,* 10, S225.
- Mes, C. (2006). *Chronische vermoeidheid: implementatie van cognitief-gedragsmatige revalidatie in Oost- en Zuid-Nederland.* Abstractboek symposium 'Meer met minder moeite', Enschede.
- Torenbeek, M., Mes, C.A.J., Liere, M.J. van, Schreurs, K., Meer, R. Ter, Kortleven, G.C., & Warmerdam, C.G.M. (2006). Gunstige resultaten van een revalidatieprogramma met cognitieve gedragstherapie en gedoseerde fysieke activiteit bij patiënten met het chronische-vermoeidheidssyndroom*. Nederlands Tijdschrift voor Geneeskunde,* 150(38), 2088-2094.
- C.A.J. Mes e.a. (2005). *Chronische vermoeidheid: implementatie van cognitief-gedragsmatige revalidatie in Oost- en Zuid-Nederland.* Roessingh Research & Development i.s.m. Revalidatiecentrum Het Roessingh en Stichting Revalidatie Limburg, intern rapport.
- Mes, C.A.J., & Jillings, J.J. (2005). *Ontwikkeling van een op maat gesneden behandeling binnen de pijnrevalidatie*. Het Roessingh/Roessingh Research & Development: intern rapport.
- Visscher, C., Hofman, N., Mes, C., Lousberg, R., & Naeije, M. (2005). Is temporomandibular pain in chronic whiplash associated disorders part of a more widespread pain syndrome? *Clinical Journal of Pain*, 21, 353-357.
- Schreurs, K., & Mes, C. (2004). Depressie en chronische pijn. Twee kanten van een medaille? *Nederlands Tijdschrift voor Behavioral Medicine,* 15, 12-14.
- Mes, C.A.J. (2003). Meten in de praktijk. RAND 36-item Health Survey. *Nederlands Tijdschrift voor Fysiotherapie*, 113, 116.
- Mes, C.A.J. (2001). Het chronische whiplash syndroom: onderzoek naar de behoeften van patiënten aan zorg en behandeling. *Whiplash Rapport*, 3, 7-9.
- Mes, C.A.J., Wijnhoven, J.H., & Harten, W.H. van (2001). *Het chronische whiplash syndroom: onderzoek naar de behoeften van patiënten aan zorg en behandeling*. Roessingh Research & Development: intern rapport.
- Mes, C.A.J., & Harten, W.H. van (2000). Kwaliteit van leven en het Chronische Vermoeidheidssyndroom: de behoeften van patiënten als uitgangspunt voor een revalidatieprogramma. *TSG Tijdschrift voor Gezondheidswetenschappen*, 3, 161-164.
- Mes, C.A.J., Winter, F.A.M., & Wijck, R. van (1999). Programma-evaluatie in de pijnrevalidatie: analyse van het Pijn de Baas programma. *Gedrag & Gezondheid,* 27, 136-147.
- Mes, C.A.J., & Harten, W.H. van (1998). *Kwaliteit van leven en het Chronische Vermoeidheidssyndroom: de behoeften van patiënten als uitgangspunt voor de opzet van een revalidatieprogramma*. Revalidatiecentrum Het Roessingh: intern rapport.
- Mes, C.A.J. (1997). *Chronische pijn en fysieke activiteit: analyse van de fysieke componenten van het Pijn de Baas programma*. Rijksuniversiteit Groningen / EMDAPA / Revalidatiecentrum Het Roessingh: scriptie.

- Drost, G., Mes, C., Kuiper, D., & Stevens, M. (1996). Minitriathlon: op je sloffen of op je tandvlees? Duursport bij kinderen. *Lichamelijke Opvoeding*, 9, 393-397.
- Drost, G., Mes, C., Kuiper, D., & Stevens, M. (1996). Op je sloffen of op je tanden? Jeugdtriathlon volgens het Beijumprotocol. *Triathlon: Officieel Bondsorgaan van de Nederlandse Triathlon Bond*, 5, 22-23.

Congresbijdragen

- Mes, C.A.J. Lousberg, R., Zilvold, G., Delespaul, P.A.E.G., & Turk, D.C. *Application of the match-mismatch model in chronic pain treatment: an experience sampling study.* Posterpresentatie 5th Congress of the European Federation of IASP Chapters. Istanbul, september 2006.
- Schreurs, K.M.G., Mes, C.A.J., Witting, M., & Vollenbroek-Hutten, M.M.R. *Measuring daily activity in fibromyalgia: Is an accelerometer a useful tool?* Posterpresentatie 5th Congress of the European Federation of IASP Chapters. Istanbul, september 2006.
- Torenbeek, M., Mes, C.A.J., Liere, M.J. van, Meer, R. Ter, Kortleven, G.C., & Warmerdam, C.G.M. *Interdisciplinaire revalidatie bij chronische vermoeidheid. Resultaten.* Mondelinge presentatie symposium 'Meer met minder moeite'. Enschede, mei 2006.
- Torenbeek, M., Mes, C.A.J., Liere, M.J. van, Meer, R. Ter, Kortleven, G.C., & Warmerdam, C.G.M. *Interdisciplinaire revalidatie bij chronische vermoeidheid. Een pre-experimenteel onderzoek.* Posterpresentatie Jaarcongres Koninklijk Genootschap voor Fysiotherapie, oktober 2005.
- Mes, C.A.J., Lousberg, R., Oudshoorn, C.G.M., Zilvold, G., Delespaul, PA.E.G., & Turk, D.C. *Pain experiences and expectations: replication of the match-mismatch model in daily life of chronic pain patients*. Posterpresentatie 11th World Congress on Pain. Sydney, augustus 2005.
- Mes, C., & Wassink, B. *De ontwikkeling van een op maat gesneden revalidatieprogramma voor allochtone chronische pijnpatiënten.* Mondelinge presentatie symposium 'Allochtone pijnpatiënten: in beweging!'. Enschede, juni 2005.
- Mes, C.A.J., Lousberg, R., Zilvold, G., Delespaul, P.A.E.G., & Turk, D.C. *Experience sampling: the influence of cognitive-behavioral rehabilitation on pain expectations and experiences.* Mondelinge presentatie 8th International Congress of Behavioral Medicine. Mainz, augustus 2004.
- Mes, C.A.J., Lousberg, R., Gulmans, J., & Harten, W.H. van. *Treatment outcome in a Dutch multidisciplinary program for chronic pain*. Mondelinge presentatie 2st World Congress of the International Society of Physical and Rehabilitation Medicine. Praag, mei 2003.
- Mes, C.A.J., Lousberg, R., Blij, I. van der, & Zilvold, G. *The role of the (un)predictibility of pain in the maintenance of chronic pain*. Mondelinge presentatie 2st World Congress of the International Society of Physical and Rehabilitation Medicine. Praag, mei 2003.
- Visscher, C.M., Hofman, N., Mes, C.A.J., Lousberg, R., & Naeije, M. *Is temporomandibular pain in chronic whiplash associated disorders part of a more widespread pain syndrome?* Mondelinge presentatie 23e Store Kro Meeting of the Society of Oral Physiology. Leuven, mei 2003.
- Mes, C.A.J., Lousberg, R., Bergman, M. & Harten, W.H. van. *Experience Sampling in chronic pain rehabilitation: the relation between the fear-avoidance model of pain and the match-mismatch model of fear and pain*. Posterpresentatie 1st World Congress of the International Society of Physical and Rehabilitation Medicine. Amsterdam, juli 2001.
- Mes, C.A.J., Lousberg, R., Bergman, M. & Harten, W.H. van. *Experience Sampling in chronic pain rehabilitation: the relation between the fear-avoidance model of pain and the match-mismatch*

model of fear and pain. Posterpresentatie symposium Revalidatiegeneeskunde en revalidatietechnologie in Twente. Hengelo, september 2001.

- Mes, C.A.J., & Harten, W.H. van. *Rehabilitation and Chronic Fatigue Syndrome: patient needs as a starting point for designing a rehabilitation program*. Posterpresentatie ISOQOL 6th Annual Conference of the International Society for Quality of Life Research. Barcelona, november 1999.
- Mes, C.A.J., & Harten, W.H. van. *Revalidatie en het Chronische Vermoeidheidssyndroom: de behoeften van patiënten als uitgangspunt voor de opzet van een revalidatieprogramma*. Mondelinge presentatie Nederlands Public Health Congres. Groningen, maart 1999.

PROGRESS RANGE

The following publications have also been published in the progress range by Roessingh Research & Development, Enschede, the Netherlands. Copies can be ordered as long as available via info@rrd.nl.

1. Pot, J.W.G.A., Boer, H., Harten, W.H. van, Hermens, H.J., & Seydel, E.R. Comprehensive need-assessment. Ontwikkeling van een meetinstrument voor zorgbehoeften en kwaliteitsbeoordeling door patiënten. Roessingh Research & Development, the Netherlands, September 1994, ISBN 90-25452-01-2.
2. Leerdam, N.G.A. van, & Hermens, H.J. Revalidatietechnologie in Euregio. Roessingh Research & Development, the Netherlands, July 1995, ISBN 90-75452-02-0.
3. Duda, L., Noort, L.O. van, Röseler, S., Greitemann, B.O.L., Harten, W.H. van, & Klazinga, N.S. Rehabilitation in Germany and the Netherlands, a comparison of two rehabilitation-systems. Roessingh Research & Development, the Netherlands, August 1995, ISBN 90-75452-03-9.
4. Hermens H.J., Nene, A.V., & Zilvold, G. Electrophysiological Kinesiology. Proceedings of the 11th congress of the International Society of Electrophysiologie and Kinesiology in Enschede, The Netherlands 1996. Roessingh Research & Development, the Netherlands, October 1996, ISBN 90-75452-04-7.
5. Harten, W.H. van. Bouwen aan een kwaliteitssysteem in de revalidatiezorg. Een poging tot constructieve technology assessment van een kwaliteitssysteem in een gezondheids-zorginstelling. Roessingh Research & Development, the Netherlands, December 1997, ISBN 90-75452-07-1.
6. Baardman, G., & IJzerman, M.J. Design and evaluation of a hybrid orthosis for people with paraplegia. Roessingh Research & Development, the Netherlands, November 1997, ISBN 90-75452-08-X.
7. Hutten, M.M.R. Lumbar Dynamometry: a useful method for assessment of patients with chronic low back pain? Roessingh Research & Development, the Netherlands, November 1999, ISBN 90-75452-13-6.
8. Salm, A. van der, Harten, W.H. van, & Maathuis, C.G.B. Ketenkwaliteit cerebrale parese zorg. Een beschrijving van de cerebrale parese zorg en mogelijke verbeteringen hierin. Roessingh Research & Development, the Netherlands, April 2001, ISBN 90-75452-19-5.
9. Nederhand, M.J. Muscle activation patterns in post traumatic neck pain. Roessingh Research & Development, the Netherlands, March 2003, ISBN 90-75452-27-6.
10. Jannink, M.J.A. Usabillity of custom-made orthopaedic shoes in patients with degenerative disorders of the foot. Roessingh Research & Development, the Netherlands, September 2004, ISBN 90-75452-28-4.
11. Blokhorst, M. State-dependent factors and attention in whiplash associated disorder. Roessingh Research & Development, the Netherlands, January 2005, ISBN 90-365-2111-4.
12. Buurke, J.H. Walking after stroke co-ordination patterns & functional recovery. Roessingh Research & Development, the Netherlands, February 2005, ISBN 90-365-2140-8.

13. Salm, A. van der. Spasticity reduction using electrical stimulation in the lower limb of spinal cord injury patients. Roessingh Research & Development, the Netherlands, October 2005, ISBN 90-365-2253-6.
14. Snoek, G.J. Patient preferences for reconstructive interventions of the upper limb in tetraplegia. Roessingh Research & Development, the Netherlands, December 2005, ISBN 90-365-2255-2.
15. Kroon, J.R. de. Therapeutic electrical stimulation of the upper extremity in stroke. Roessingh Research & Development, the Netherlands, December 2005, ISBN 90-365-2269-2.
16. Dijk, H. van. Motor skill learning: age and augmented feedback. Roessingh Research & Development, the Netherlands, March 2006, ISBN 90-365-2302-8.