To solve or not to solve? : effects of problem solving therapy and graded activity in non-specific low back pain

Citation for published version (APA):

van den Hout, J. H. C. (2002). *To solve or not to solve? : effects of problem solving therapy and graded activity in non-specific low back pain*. [Doctoral Thesis, Maastricht University]. Universiteit Maastricht. <https://doi.org/10.26481/dis.20020419jh>

Document status and date:

Published: 01/01/2002

DOI:

[10.26481/dis.20020419jh](https://doi.org/10.26481/dis.20020419jh)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

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

[Link to publication](#)

General rights

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

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

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

www.umlib.nl/taverne-license

Take down policy

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

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

TO SOLVE OR NOT TO SOLVE?

EFFECTS OF PROBLEM SOLVING THERAPY AND GRADED ACTIVITY
IN NON-SPECIFIC LOW BACK PAIN.

U n i v e r s i t e i t M a a s t r i c h t

COLOPHON

To solve or not to solve? Effects of problem solving therapy and graded activity in non-specific low back pain./ Anja van den Hout. – Thesis Maastricht University. – With references. – With summary in Dutch.

Cover: 'The Thinker' by Auguste Rodin, 1880.

Cover design: Paula Sprengers
Lay-out: Jacky Rinket
Drawings: Wil van den Hout
Fotography: Anja van den Hout

Printed by: Datawyse, Maastricht

© A. van den Hout, Maastricht 2002
ISBN 90-9015711-5

The studies in this thesis have been supported by Grant No. 940-31-004 of the Council for Medical and Health Research of the Netherlands (NWO-MW).

Financial support for the printing of this thesis has been kindly provided by the Hoensbroeck Rehabilitation Center, the Institute for Rehabilitation Research (iRv), Elsevier Gezondheidszorg, Dutch Arthritis Association (Nationaal Reumafonds), the Dutch Chapter of the International Association for the Study of Pain (de Nederlandse Vereniging ter Bestudering van Pijn), Rug AdviesCentra Nederland b.v., Anna Fonds, Leiden, and Hoensbroeck Centrum voor Arbeids- perspectief.

TO SOLVE OR NOT TO SOLVE?

EFFECTS OF PROBLEM SOLVING THERAPY AND GRADED ACTIVITY
IN NON-SPECIFIC LOW BACK PAIN.

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit Maastricht,
op gezag van de Rector Magnificus,
Prof. Dr. A.C. Nieuwenhuijzen Kruseman
volgens het besluit van het College van Decanen,
in het openbaar te verdedigen
op vrijdag 19 april 2002 om 14.00 uur

door

Johanna Henrica Cornelia van den Hout

PROMOTOR:

Prof. dr. M.A. van den Hout

CO-PROMOTOR:

Dr. J.W.S. Vlaeyen

BEOORDELINGSCOMMISSIE:

Prof. dr. A. Arntz (voorzitter)

Prof. dr. G. Crombez (Universiteit Gent, België)

Dr. S.J. Linton (Örebro Medical Center Hospital, Sweden)

Prof. dr. C.P. van Schaijck

Dr. A. Schmidt

CONTENTS

CHAPTER 1	
Introduction.	7
CHAPTER 2	
The effects of problem solving therapy in mental health care and behavioral medicine: a systematic review.	31
CHAPTER 3	
Does failure hurt? The effects of failure feedback on pain report, pain tolerance, and pain avoidance.	59
CHAPTER 4	
The effects of failure feedback and pain-related fear on pain report, pain tolerance, and pain avoidance in chronic low back pain patients.	77
CHAPTER 5	
Functional disability in non-specific low back pain: the role of pain-related fear and problem-solving skills.	97
CHAPTER 6	
Has problem solving therapy supplemental value when added to behavioral graded activity in non-specific low back pain patients? A randomized clinical trial.	111
CHAPTER 7	
Secondary prevention of work-related disability in non-specific low back pain: Does problem-solving therapy help? A randomized clinical trial.	143
CHAPTER 8	
General discussion.	163
REFERENCES	187
SUMMARY	205
SAMENVATTING	211
DANKWOORD	217
ABOUT THE AUTHOR	221
PUBLICATIONS	222

CHAPTER 1

Introduction

*"Prob...lem...sol...ving?" Pooh said slowly, scratching his head.
"That sounds like a 'What'. The 'Whats' are easy.
It's the 'Hows' that are difficult."*

*From: Winnie-the-Pooh on Problem Solving,
Allen & Allen, 1995*

INTRODUCTION

Interpersonal problem-solving, low back pain, pain-related fear, failure experiences, and secondary prevention, five themes that appear to be scarcely connected. The introductory chapter will sketch the hypothetical relationships between these themes, which conceptualize the contents of this thesis. Problem solving, the starting point of this chapter, has frequently been applied in health problems, but not in the field of pain. In this thesis, it is hypothesized that problem-solving skills may have important value in the treatment and prevention of chronic pain. In the second section, low back pain, a target health problem in most western industrialized countries, is defined and described in order to work out and evaluate possible solutions in this thesis. Contemporary theories of pain, stress the importance of a biopsychosocial approach to the problem. In this context the role of pain-related fear will be discussed. The third section will describe how problem solving and low back pain may hypothetically be related. The role of failure experiences is one of the issues introduced in this context. The fourth section shows how secondary prevention of low back pain incorporates several ideas presented in the foregoing sections and applies these to the problem of pain. Finally, research questions will be formulated to be dealt with in the subsequent chapters of this thesis.

PROBLEM SOLVING

Problem solving theory

Problems rarely tend to be fixed and static, but rather are conceptualized as "the normal ebb and flow of living life", a constantly fluctuating process that needs to be handled with flexibility and adaptability on the side of the problem solver (D'Zurilla & Goldfried, 1971). From the early 1970s the focus of research in problem solving began to shift from non-social and abstract problems to interpersonal and actual problems (Platt, Prout, & Metzger, 1987). In this thesis 'problem solving' refers to the latter. Initially, D'Zurilla and Goldfried (1971) suggested that a person's ability to cope with the world be largely determined by the effective use of problem solving skills. According to their theory, the process of problem solving consists of 5 steps: a) problem orientation, b) problem definition and formulation, c) generation of alternatives, d) decision making, e) and solution implementation and verification (Figure 1). Later, the problem solving theory was extended by the idea that originally anticipated goals, desires and consequences may alter as well. Thus, the effective problem solver must not only recognize that problems ask for a flexible use of problem solving skills, but also that a situation may ask for several interdependent solutions in which a feedback process is critical in guiding later decisions in an evolving process (Kanfer & Busemeyer, 1982). Subsequent to this initial theory, a tradition of research was started regarding the problem-solving process as it relates to human adjustment and development of social compe-

tence. Especially the identification of thinking processes that underlie the successful solution of intra- and interpersonal problems, and the development of programs for groups with difficulties in these processes were main directions in the social-competence tradition (Spivack, Platt, & Shure, 1976). It was hypothesized that: a) interpersonal cognitive problem-solving skills (ICPS) are learned; b) differences in ICPS have different significance for the social adjustment and mental health of the individual, depending on age; c) manifest deficiency in ICPS in an individual may reflect: i) insufficient learning of these skills, ii) blocking of such thinking due to interfering emotions, or iii) deterioration of such thinking; and finally d) therapeutic and preventive programs that enhance the operation of ICPS will increase social adjustment and prevent the individual from developing psychological dysfunction (Platt et al., 1987). The social-competence theory instigated the application of ICPS-therapy in (mental) health care settings, and especially in those disorders in which there was substantial interference of emotions, among others in depression (Areal et al., 1993; Hussian & Lawrence, 1981; Lynch, Tamburrino, & Nagel, 1997; Mynors Wallis, Gath, Lloyd Thomas, & Tomlinson, 1995; Nezu & et al., 1986; Nezu & Perri, 1989; Stark, Reynolds, & Kaslow, 1987), anxiety (DiGiuseppe, McGowan, Simon, & Gardner, 1990; Jannoun, Munby, Catalan, & Gelder, 1980), and aggressive behavior and conduct disorder in children (Kazdin, Bass, Siegel, & Thomas, 1989; Kazdin, Esveldt Dawson, French, & Unis, 1987; Kazdin, Siegel, & Bass, 1992).

FIGURE 1. Problem-solving theory (D'Zurilla & Goldfried, 1971).

Several studies were interested in the more fundamental question concerning the role of problem solving in the development and maintenance of mainly psychological and stress-related disorders. Chiefly, two roles were considered:

red: a mediating and a moderating role. First, support was found for a mediating model regarding depression and anxiety symptoms (Kant, D'Zurilla, & Maydeu Olivares, 1997; Nezu & Ronan, 1985). In this model problem solving coping occupies a position in a causal chain and links the antecedent, here current problems, to the outcome variable, i.e. depression or anxiety (Baron & Kenny, 1986). The model assumes that problem solving is an intervening variable that accounts, to a certain degree, for the causal relations between everyday problems and depression and anxiety. Nezu and Ronan (1985) hypothesized that important sources of stress emanate from experiencing negative life stress (disease, divorce, death of a family member), which, in turn, result in the increase of current daily problems that may function as minor-stressors. However, the impact these sources of stress have on the level of either depressive or anxious symptoms, would largely depend on the activation of efficient problem solving coping when confronted with current stress.

Second, the moderator model is based on both research and clinical experience that clearly indicate that most individuals do not necessarily develop psychological symptoms despite the severely stressful circumstances they may be confronted with (Nezu & et al., 1986). In this model problem solving ability is conceptualized as an independent trait-like variable (Baron & Kenny, 1986), i.e. a protective factor. It is now hypothesized that high levels of negative stress will have minimal effects on depressive or anxious symptoms when the individual is classified as an effective problem solver. Regarding depression, the moderator-hypothesis was supported on several occasions (Nezu & et al., 1986; Nezu & Ronan, 1988), however, later studies did not find supportive evidence (D' Zurilla & Sheedy, 1991; Kant et al., 1997). The different results may have been caused by different measures of stressful life events. In the Nezu studies, major negative life events were measured, whereas in the more recent studies (Kant et al., 1997), everyday problems and "daily hassles" were assessed.

Problem solving therapy

Analogous to the development of theoretical models regarding the role of problem solving in stressful life events, everyday problems and emotional disorders, treatment programs have been designed to intervene in these processes. Problem solving programs have varied widely in terms of target population, program goals, setting, amount of training in other coping skills within the overall problem solving framework, instructional methods used, and overall program structure (D'Zurilla, 1988). D'Zurilla (1988) suggests that the most appropriate case for problem solving therapy is one in which negative behavior or negative stress effects, such as depression, anxiety, or psychophysiological symptoms (e.g. pain), are associated with deficits in problem solving ability and performance. In line with this, the applications of problem solving therapy have been numerous: social phobia, agoraphobia, obesity, coronary heart disease, schizophrenia, psychiatric problems, HIV risk behaviors, drug abuse,

suicidal behavior, childhood aggression, conduct disorder, and depression (Nezu, Nezu, Friedman, Faddis, & Houts, 1998). Therapy programs may either focus directly on the specific target problem or on the training of problem solving skills according to the model of D'Zurilla and Goldfried (1971). In the former case, the problem-solving model is merely applied as general therapy structure, whereas in the latter case, target problems are suggested to change indirectly as a result of improved problem solving skills. Problem-solving therapy may be applied in combination with other techniques in order to accomplish multi-faceted treatment goals. In these combinations, problem-solving therapy merely serves as a basis to increase self-control in order to maintain initial treatment effects in the long term. As for training methods, it appears that, as compared to self-instructional training and social learning intervention, the most fully supported time-efficient and cost-effective approach to improve social problem solving skills simply is the sequential training of component behavioral skills (Tisdelle & St Lawrence, 1986).

Next, a protocol of problem-solving therapy according to Nezu et al. (Nezu et al., 1998; Nezu & Perri, 1989) will be briefly described. In this approach the five steps of cognitive problem solving are sequentially trained: problem orientation, problem definition and formulation, generation of alternative solutions, decision making, and solution implementation and evaluation.

Regarding efficacy, the problem-solving therapy has been claimed to be effective in a wide range of clinical disorders and psychological problems (Larson, 1990; Nezu et al., 1998). This claim is supported by individual outcome studies as well as reviews regarding specific domains: children and adolescents with internalizing disorders (Grossman & Hughes, 1992), children and youth with learning and behavioral problems (Coleman, Wheeler, & Webber, 1993), conduct disorder in children and adolescents (Kazdin, 1997), college students at risk (Heppner, Neal, & Larson, 1984), and patients with major depression and more broadly defined emotional disorders in primary care (Gath & Catalan, 1986; Mynors Wallis, 1996). Although results are mostly in favor of the hypothesis that problem-solving therapy is effective, some studies reported that problem-solving therapy is: i) no more effective than other cognitive and behavioral treatment conditions in social anxiety (DiGiuseppe et al., 1990), ii) effective regarding a limited number of outcome variables in learning and behavioral problems (Coleman et al., 1993) and even iii) less effective than programmed practice and exposure in agoraphobia (Jannoun et al., 1980; Jansson & Oest, 1982). Two critical reviews (Larson, 1990; Tisdelle & St Lawrence, 1986) regarding the efficacy of problem-solving therapy call for follow-up measurements of therapy outcomes, and discuss future directions regarding among others 'real-life problem-solving', i.e. a complex, dynamic, highly interactive and intermittent process, that differs very much from the stage-sequential model. Moreover, the need for adequate (behavioral) assessment of actual problem-solving skills is discussed (Tisdelle & St Lawrence, 1986).

TABLE 1. Problem solving therapy, an overview.

- 1) **PROBLEM ORIENTATION:** a) facilitate a positive and constructive orientation towards problems confronted with in everyday life, b) emphasize the idea that problem-solving is a viable method of effectively coping with problems, c) reduce cognitive distortions or belief patterns that might interfere with effective problem solving, e.g. "I'll fail, whatever I do", d) facilitate the acceptance of emotional reactions and "use" them as tool to identify current problems, e) use the "stop and think" method to prevent impulsive responses or avoidance behavior.
- 2) **PROBLEM DEFINITION AND FORMULATION:** a) gathering information about the problem, b) describe the facts in clear language, c) separate facts from assumptions, d) identify obstacles and conflicts, e) set realistic goals.
- 3) **GENERATION OF ALTERNATIVE SOLUTIONS / BRAINSTORMING:** a) think of as many possible solutions as possible, b) defer judgement until later, c) combine ideas.
- 4) **DECISION MAKING:** a) identify the consequences (long and short term, regarding self and others), b) evaluate the consequences of alternative solutions and conduct a cost-benefit analysis, c) estimate the likelihood that a solution will successfully solve the problem, d) estimate the likelihood that the implementation of a solution is practicable.
- 5) **SOLUTION IMPLEMENTATION AND VERIFICATION:** a) optimally carry out a solution plan, b) monitor the actual consequences of the implemented solution, c) evaluate whether the solution has solved the problem, d) engage in self-reinforcement if the problem is solved, e) troubleshoot if the outcome is unsatisfactory and recycle the model if necessary.

NON-SPECIFIC LOW BACK PAIN

As D'Zurilla (1988) suggests, the most appropriate case for problem solving therapy is one in which negative behavior and negative stress effects are associated with problem solving ability and performance. Pain problems may very well fit in this approach. Therefore, in this paragraph, the problem of non-specific low back pain will be described in more detail.

Epidemiology of non-specific low back pain

Non-specific low back pain is a common disease in most western industrialized countries (Andersson, 1999; Picavet, Schouten, & Smit, 1996). About 60-80% of western populations report back pain at some time in their life, and in most cases it concerns low back pain (Andersson, 1999; Borenstein, 1998). According to a review by Andersson (1999), point prevalence of low back pain as reported in cross-sectional studies varies between 12-18%, whereas period prevalence varies from 25-42%. Annual incidence was reported to be 5% (Andersson, 1991). In the Netherlands, half of the population reports low back pain in the year before the study took place, and 40% of them states that the pain has been

present more than 12 weeks, or is almost always present (Picavet et al., 1996). In Dutch General Practice, the incidence of low back pain is about 30 new episodes per 1000 registered patients per year, whereas prevalence is 9.4 episodes per 1000 registered patients in 3 months (Lamberts, 1991; Van den Velden, De Bakker, Claessens, & Schellevis, 1991). In approximately 85% of all low back pain patients, no specific diagnosis (vertebral fracture, infectious disease, rheumatoid arthritis, ankylosing spondylitis, or herniated disc; Waddell & Turk, 1992) can be objectified and are therefore classified as non-specific low back pain (Deyo, Rainville, & Kent, 1992).

The natural course of low back pain is generally favorable. Seventy-five to ninety percent of patients with acute low back pain recover within four to six weeks (Frymoyer, 1988). Of the remaining 10-25%, another 50% recover within six to twelve weeks (sub-acute phase). Unfortunately, in still 5-10% of the total population back pain lasts longer than 12 weeks (chronic phase). The natural course of back pain is rather whimsically, what makes it hard to make a solid prognosis. Although most back pain patients recover within a short period of time, complaints relapse in 70-80% of the population (Andersson, 1998).

Chronic low back pain (CLBP) has very negative consequences for the patient. Forty percent of CLBP-patients ceased normal activities because of low back pain, 32% were on the sicklist at the workplace, and 60% sought for medical help (Picavet et al., 1996). Next to the individual burden of illness, the relatively small group of chronic pain patients is responsible for 75% of the economic burden, i.e. direct costs (medical consumption) and indirect costs (absenteeism and disability pensions; Cats-Baril & Frymoyer, 1991). About 90% of the economic burden of low back pain are caused by indirect costs. Overall, musculo-skeletal disorders account for about 40% of costs of absenteeism, 60% of all persons assigned a disability pension, and about 30% of total medical costs, and is therefore among the five most expensive diseases in the Netherlands (Meerding, Bonneux, Polder, Koopmanschap, & van der Maas, 1998).

In conclusion, the problem we have to deal with is of a serious magnitude, and has large consequences, as well on the individual, as on the social (e.g. family, work) and the economic level. The next hypothetical case-example may illustrate some individual consequences.

A case example

A 42-year-old man is referred to the rehabilitation center. His main complaint is a continuous pain in the lower parts of his back, as well as a radiating pain in his right leg. He had back pain before, but never as sharp and disabling as this time. The pain started very unexpectedly while he was operating a machine at the workplace. He felt a shooting agony, as if a nerve got pinched, and he had the feeling he could not move his back. This was about 4 weeks ago. He is on the sick list ever since, the shooting pain has gone, but a persistent moaning pain remains. During the first interview with this patient, the rehabilitation physician could not make a specific diagnosis of the back pain, neither did X-rays

show any particularities. Nevertheless, the patient reported alarming disability: Besides his work, he has not felt able to resume some of his daily activities. Neither can he participate in family life as much as he wants to. He feels he has let down his wife and two young children, but can not think of a solution as long as the pain is constantly present. Mainly, he lacks confidence regarding movements and activities he should and should not do. Therefore he avoids any activity because he is afraid that specific movements could aggravate his situation. He worries a lot about the future, his job, his family, and lasting disability, leaving him seriously depressed as his situation seems rather hopeless at this moment.

The provisional diagnosis of the rehabilitation physician is non-specific low back pain. Next to a serious level of functional disability, he observes comorbid depressive features, and moreover he remarks that the patient has an irrational fear of movement and activity since the onset of his pain.

The present case not only leads to several clinical questions about how to treat this patient, but may also raise questions regarding the mechanisms underlying non-specific low back pain: Why do some patients develop serious and sometimes chronic problems, whereas others can resume their normal lives within a few days? What are the underlying mechanisms of chronic low back pain? What is the influence of psychological factors on pain disability, especially that of pain-related fear, pain behavior, negative affect, negative stress, and coping style? What treatments are (most) effective? Some of these questions will be addressed in this thesis. The next paragraph will give an overview of the development from a traditional model of pain towards a more contemporary one, the biopsychosocial model.

A biopsychosocial model of back pain

In the traditional Cartesian model of specific pain pathways, Descartes (1664) hypothesized that pain is merely a signal of tissue damage. In this model, the perception of pain was compared to "pulling at one end of a rope one makes to strike at the same instant a bell, which hangs at the other end". The Cartesian model was the most influential model for three centuries, however, since the early 1970s there was a growing awareness that the biomedical model could not give a satisfactory explanation for the extent in which chronic pain patients were disabled in functioning and daily lives. In fact, there was no relationship at all between biomedical findings (MRI-scans, X-rays) and pain symptoms and disability (Boden, Davis, Dina, Patronas, & Wiesel, 1990; Jensen et al., 1994). The gate control theory of Melzack and Wall (1965) broke with the biomedical dualistic tradition through their hypothesis that neurophysiological and psychological activity can modulate pain and help to explain the complexity of clinical pain. In this theory, pain is not only the product of information ascending from peripheral sources to the brain, but also of information descending from the brain. According to this theory, the so-called "gate" is located in the dorsal horn of the spine, in which peripheral and central sources of information are

integrated what eventually results in the experience of pain (Melzack & Wall, 1965).

About ten years after the publication of the gate-control theory, Fordyce (1976) made an important contribution to the biopsychosocial model of pain by introducing operant conditioning principles to the treatment of chronic pain. Essential was his assumption that factors maintaining the pain problem did not necessarily correspond with those initiating the pain. The maintaining factors are described in terms of pain behavior and the consequences this behavior has for the individual and the environment. When pain behavior is immediately followed by a situation that is reinforcing, e.g. the attention of a spouse (positive reinforcement) or pain relief (negative reinforcement), the chance that this particular behavior is maintained increases. According to Fordyce (1976), the consequences of pain will chiefly determine whether the problem develops into a chronic problem or will eventually dissolve.

The developments described above have resulted in a new clinical model of pain: the biopsychosocial model (Waddell, 1987; Waddell, 1998). The biopsychosocial model includes at least four levels: 1) biomedical, 2) individual psychological, 3) interpersonal and interactive, and 4) social and cultural. The levels are interactive, in that a change at one level will, as a matter of course, bring about changes in other levels. The model has successfully been applied in the case of low back pain (Waddell, 1987). During the last decades, research has mapped the role of various variables in the light of this model, including pain behavior (Fordyce, 1976), psychophysiological reactivity (Flor, Turk, & Birbaumer, 1985), cognitive processes (Heyneman, Fremouw, Gano, Kirkland, & et al., 1990), expectations (Arntz, Van den Hout, Van den Berg, & Meijboom, 1991b; Council, Ahern, Follick, & Kline, 1988), perceived control (Arntz & Schmidt, 1989; Philips, 1987; Spinhoven & Linssen, 1991; Turner, Jensen, & Romano, 2000), coping (Jensen, Turner, Romano, & Karoly, 1991b; Rosenstiel & Keefe, 1983; Turner & Jensen, 1993), personality (Bigos et al., 1991; Gallagher et al., 1989), and depression (Leino & Magni, 1993; Main, Wood, Hollis, Spanswick, & Waddell, 1992; Pincus & Williams, 1999). One of the most recent models that was deduced from the biopsychosocial model, describes the hypothesized role of pain-related fear (Asmundson, Norton and Allardings, 1997; Crombez et al., 1999; Lethem, Slade, Troup, & Bentley, 1983; Vlaeyen, Kole Snijders, Boeren, & van Eek, 1995b; Vlaeyen, Kole-Snijders, Rotteveel, Ruesink, & Heuts, 1995c; Vlaeyen and Linton, 2000). In the next section, the role of pain-related fear and avoidance in the development and maintenance of chronic pain will be further elaborated.

The role of pain-related fear

In figure 2, a biopsychosocial model on the role of pain-related fear in low back pain is outlined (Vlaeyen et al., 1995b; Vlaeyen et al., 1995c). The hypothetical model illustrates how acute low back pain can develop into chronic pain as a result of several interacting cognitive and behavioral processes.

FIGURE 2: Biopsychosocial model of low back pain (Vlaeyen et al., 1995b; Vlaeyen and Linton, 2000).

According to the model, immediately after the onset of pain, catastrophizing thoughts may develop (e.g. "There is something terribly wrong in my back", "My situation is deteriorating, I will end up in a wheelchair"; Vlaeyen, 1991). Pain catastrophizing is considered an exaggerated negative orientation towards noxious stimuli, and was found to modulate the experience of pain (McCracken & Gross, 1993; Sullivan, Bishop, & Pivik, 1995) and to predict pain-related fear (Vlaeyen et al., 1995b; Vlaeyen et al., 1995c). Next to and overlapping with pain catastrophizing, negative affectivity, as a mood dispositional dimension (Watson & Clark, 1984), may moderate the relation between pain and pain-related fear. In general, persons with high negative affectivity are more likely to experience intense states of distress, even in the absence of any overt stress (Watson & Pennebaker, 1989), and are therefore considered more vulnerable to the development of specific fears. As the instigation of pain may be a salient threat, pain-related fear may emerge especially in those cases patients are characterized as highly negative affective.

"Kinesiophobia" (Kori, Miller, & Todd, 1990), one of the constructs within the pain-related fear model, refers to an irrational fear that physical activity would cause (renewed) injury (e.g. "I wouldn't have this much pain if there weren't something potentially dangerous going on in my body", "Someone in my position should better refrain from any physical activity"). Kinesiophobia, parallel to fears in general, may result in the avoidance of various activities (e.g. work, sports, social life; Philips, 1987). In turn, avoidance behavior will result in the maintenance or exacerbation of fear, as irrational and negative beliefs about the consequences of movement and activity are never put to trial (fear-avoidance; Lethem et al., 1983). In the long term, avoidance behavior is found to result in reduction of both physical and psychosocial activities

(Crombez, Vervaeke, Lysens, Baeyens, & Eelen, 1998; Crombez, Vlaeyen, Heuts, & Lysens, 1999; Vlaeyen et al., 1995c). This reduction in turn would lead to an increase of disability (Crombez et al., 1999; Linton & Buer, 1995; Vlaeyen et al., 1995b; Waddell, Newton, Henderson, Somerville, & Main, 1993), to detrimental changes in the musculoskeletal system referred to as 'disuse syndrome' (Bortz, 1984), and comorbid depression (Leino & Magni, 1993). During the last two decades, the mechanisms of pain-related fear have been supported by numerous investigations establishing the importance of fear-avoidance in the development of chronic pain (Vlaeyen & Linton, 2000). In three independent cross-sectional studies described by Crombez, Vlaeyen, Heuts, and Lysens (1999), pain-related fear was superior to pain intensity, pain duration and negative affect in predicting self-reported disability and poor behavioral performance in chronic low back pain patients. Moreover, in two prospective studies evidence was found that supports the idea that pain-related fear is a precursor of disability, rather than a consequence (Klennerman et al., 1995; Linton, Buer, Vlaeyen, & Hellsing, 2000). The first study (Klennerman et al., 1995) included employed acute back pain patients in primary care, and found fear-avoidance variables to be successful predictors of disability 2 months after pain onset (R^2 fear avoidance = .25, $p < .001$), and even more predictive than variables regarding pain severity and pain history. Moreover, fear-avoidance and psychosocial variables, assessed 2 months after pain onset, were found to be predictive of the level of disability after one year (R^2 fear avoidance = .23, $p < .001$). In the latter study (Linton et al., 2000), a large sample of pain-free individuals from a general population was included. It turned out that individuals who scored above the median score regarding fear-avoidance beliefs, had twice the risk of developing a new episode of pain during the following year. A third prospective study in low back pain patients presenting in primary care (Burton, Tillotson, Main, & Hollis, 1995), however, did not find fear-avoidance beliefs to be predictive of disability after one year, when analyzed together with psychosocial factors such as coping (catastrophizing, praying and hoping) and distress. Close relationships between fear-avoidance and psychosocial measures may offer an explanation for the inconsistent findings.

In conclusion, although prospective studies are scarce, there is some evidence to support the hypothesis that pain-related fear is an important feature in the development of chronic pain and disability. Pain-related fear may not only be of theoretical interest, regarding chronic pain and disability, but may also be a powerful mediator of therapy outcome. To change pain-related fear variables in an early stage, as to prevent chronicity, may be one of the challenges of clinical practice.

Cognitive-behavioral therapy in low back pain patients

Cognitive-behavioral treatments of low back pain have adopted the concepts of the biopsychosocial model. Several cognitive-behavioral techniques have found their way to clinical pain practice, especially respondent techniques (e.g. bio-

feedback, progressive and applied relaxation; Flor & Birbaumer, 1993; Linton & Gotestam, 1984; Turner, 1982), operant techniques (e.g. graded activity; Fordyce, 1976; Vlaeyen, Haazen, Schuerman, Kole Snijders, & van Eek, 1995a), and cognitive techniques, e.g. Beckian therapy (Turner & Jensen, 1993), and pain coping skills training (Kole Snijders et al., 1999).

The aim of operant approaches is to increase activity levels and decrease pain behaviors by changing behavior consequences. In graded activity, activity levels are increased using time-contingent quota systems, including baseline determination, treatment contract, positive (self) reinforcement for activity increments, and workplace or home visits in order to generalize new behavior in a true to life environment (Fordyce, 1976). The operant approach has been successfully applied in both chronic pain patients (Kole Snijders et al., 1999; Nicholas, Wilson, & Goyen, 1991; Turner & Clancy, 1988; Turner, Clancy, McQuade, & Cardenas, 1990; Turner & Jensen, 1993) and sub-acute and acute low back pain patients (Fordyce, Brockway, Bergman, & Spengler, 1986; Lindström et al., 1992; Linton & Bradley, 1992; Linton, Bradley, Jensen, Spangfort, & et al., 1989). In the literature, the term cognitive treatment in pain patients is used for divergent forms of treatment such as education, hypnosis, thought challenging, and coping skills training. Cognitive treatment is often used in combination with other forms of therapy, and can either address the reduction of pain (e.g. Kole Snijders et al., 1999), or the reduction of stress (e.g. Turner & Jensen, 1993).

Multidisciplinary teams (physiotherapy, occupational therapy, psychology, rehabilitation medicine, and social work) carry out most of the cognitive-behavioral treatments described above. In a review by Flor, Fydrich, & Turk (1992) it was concluded that multidisciplinary pain clinics are effective, and at long term even more effective than no treatment, waitlist-control conditions, and monodisciplinary treatments with regard to mood, behavior and pain ratings.

In a meta-analysis, Turner (1996) selected a small sample of randomized controlled trials regarding educational, behavioral and cognitive interventions in chronic low back pain patients in a primary care setting. Except for the change in mood (i.e. depression), researchers found cognitive and behavioral treatments superior to control conditions immediately after treatment with regard to pain behavior and disability. However, in comparison to other active treatments, cognitive behavior treatments were not found to be more effective, although only few studies were available for this comparison. In a meta-analysis of 25 randomized clinical trials (Morley, Eccleston, & Williams, 1999), cognitive-behavior therapy and behavior therapy were compared to wait-list control conditions and alternative active treatments in chronic pain patients (excluding headache). It was concluded that cognitive-behavior therapy was significantly more effective than wait-list control conditions regarding all domains of outcome (median effect size of cognitive-behavior therapy across domains = 0.5). Moreover, in comparison to alternative active treatments, cognitive-behavior therapy produced significantly greater changes for the domains of pain experience, cognitive

coping, and appraisal, and reduced behavioral expression. A systematic review by Van Tulder and colleagues (2000), concluded that there is strong evidence that cognitive-behavioral treatments (cognitive-behavioral, operant, respondent, and cognitive techniques) have small to moderate positive effects in chronic low back pain patients when compared to no treatment, placebo or waiting list control conditions. However, there was also moderate evidence that addition of behavioral treatment to another treatment (i.e. physiotherapy and back education, multidisciplinary treatment, medical treatment) did not have a supplemental value, neither in the short, nor in the long term (van Tulder et al., 2000).

Finally, cognitive-behavioral treatment of chronic pain is complex, lengthy and variable, and raises methodological issues that need to be considered in clinical trials evaluating them, such as treatment credibility, attrition, the clinical relevance of change, and selection bias (Turk, Rudy, & Sorkin, 1993). The challenge for the future lies in the identification of patients at an early stage, development of effective techniques in the long term, and finding out what treatment is most effective for whom (van Tulder et al., 2000).

THE ROLE OF PROBLEM SOLVING IN PAIN

There are several reasons that make the application of problem solving therapy in pain patients appealing. In this section some hypothetical mechanisms are described in order to link problem solving to low back pain and stress. Moreover, the potential importance of problem solving therapy will be explained. Subsequently, the role of 1) stress-related avoidance and failure experiences, 2) consequences of pain for disability and the quality of life, and 3) problem solving therapy, are described.

The role of stress-related avoidance and failure experiences in pain

In figure 3, the operant model of pain as described by Fordyce (1976) is graphically represented. According to this model, stress not only increases the probability of physical symptoms (e.g. pain; Leino & Magni, 1993), but moreover pain may provoke solicitous reactions from the social environment (Flor, Kerns, & Turk, 1987; Lousberg, Schmidt, & Groenman, 1992). For example, a patient may be protected from all kinds of daily hassles and stressful events at home and at the workplace since he has pain and makes this noticeable to the environment. Being guarded against stressful situations may be reinforcing when the patient is feeling not capable to deal with these problems anyway. In operant terms this may be defined as negative reinforcement (-C-), which might result in increased pain behavior and prolonged pain duration or periods of absenteeism at the workplace. The introduction of problem-solving skills could offer an alternative coping-style when pain patients are confronted with daily hassles and stress. Being able to come up with some effective solutions, for instance when pain relapses, might enable a patient to carry on despite problems resulting from pain.

FIGURE 3: Stress-related avoidance and pain

Stress and emotional states that result from it, such as depression, have shown to increase clinical pain in humans (Brown, 1990; Sullivan & D'Eon, 1990; Wade, Price, Hamer, Schwartz, & Hart, 1990), and to decrease tolerance for experimentally induced pain (Romano & Turner, 1985; Zelman, Howland, Nichols, & Cleeland, 1991). The influence of anxiety on pain is not unequivocally, but seems to depend on mediating processes, like attention (Arntz, Dreessen, & Merckelbach, 1991a). Moreover, quality and intensity of the stressor, as well as of the induced mood may be important. Failure experiences can be an important source of emotional distress and therefore directly or indirectly influence pain. Levine, Krass, & Padawer (1993), showed that perceived failure of dealing with a non-pain-related stressor (a reading comprehension task) resulted in increased pain report. Levine and colleagues (1993) speculate on the vocational implications of their findings and assume that vocational failure (job dissatisfaction, (the risk of) being fired), may lead to increased pain report. Indeed, Bigos et al. (1991) found that job satisfaction was the best predictor of back problems, in that employees with low job satisfaction were almost 2.5 times as likely to report back pain symptoms to the company medical department or personal physician. In addition, people with low job satisfaction were found to line up for disability compensation more often (Kormann, 1977). In operant learning terms one would say that disability compensation offers the employee an opportunity to escape from unsatisfactory job conditions. On the other hand, Gallagher et al. (1989) found that health locus of control and the ability to do daily activities were significant predictors of return-to-work at six months. This finding was consistent with the clinical experience which suggests that helping patients develop a sense of control or mastery over their pain and the basic tasks of daily living, can build confidence and improve outcome

(Reeson & Craig, 1988). In this context, problem-solving skills may not only prevent possible negative effects from failure experiences, but also increase a sense of control and mastery in low back pain patients.

Consequences of pain for disability and quality of life

Pain interrupts physical, cognitive and affective functions, demands attention, and has negative consequences for the quality of life of the individual (Aldrich, Eccleston, & Crombez, 2000; Eccleston & Crombez, 1999). Aldrich et al. (2000) hypothesize that chronic pain patients tend to persevere in their attempts to solve an unsolvable problem, that is pain, despite the experience of repeated failure. This perseverance may keep them hypothetically stuck in a vicious circle of chronic pain. In order to intervene in this process it may be more important to help patients to identify and cope with the consequences of pain in everyday life. Problem solving skills are hypothesized to play an important role in the relation between everyday problems, negative affect and quality of (everyday) life (D'Zurilla & Nezu, 1982; Nezu et al., 1999).

FIGURE 4: The role of problem solving in disability and quality of life, based on Nezu et al. 1999.

Figure 4, illustrates how negative problem orientation and ineffective problem solving skills would hypothetically result in increased disability and decreased quality of life. As hypothesized, negative problem orientation can lead to negative affect, impulsive behavior and avoidance motivation, which might inhibit or disrupt subsequent problem-solving attempts. Indeed, patients with negative orientations tend to worry and complain about their health more frequently (Dugas, Letarte, Rheaume, Freeston, & et al., 1995; Godshall & Elliott, 1997). Next to problem orientation, problem-solving skills are hypothesized to moderate the impact of daily problems on affect and behavior. Evidence in favor of this moderating role was found regarding depressive affect (Nezu et al., 1986; Nezu & Ronan, 1988). Furthermore, one can imagine that daily problems increase as a consequence of hampered physical ability or diminished quality of life. For example, problems of everyday life, such as

childcare and planning of daily activities may complicate when functioning is hampered because of back pain. Though plausible, the model suggested above (based on Nezu et al. (1999)) has not been tested extensively. Regarding pain, Shaw, Feuerstein, Haufler, Berkowitz & Lopez (2001) recently found problem solving (positive problem orientation; impulsivity / carelessness; avoidant style) to moderate the impact of symptom severity on functional loss as measured by the SF-12, a general health status measure. Moreover, they found that with regard to avoidant style, the aforementioned moderation-effect was only found when patients experienced symptoms longer than 12 months. The study by Shaw and colleagues (2001) gives first indications that problem-solving skills, as well as problem orientation may contribute to the biopsychosocial model of chronic pain. This leaves us with some interesting empirical questions to be solved: What is the contribution of problem solving to a model of chronic pain? How relates problem solving to pain and disability? What is its importance to clinical practice?

Problem solving therapy in pain patients

Although the effects of problem solving therapy have been studied with regard to several populations within the field of mental health care and medical psychology, application in pain populations is scarce. Nevertheless, there are some studies that suggest that the treatment might be applicable in this area. First, Linton and colleagues (Linton & Bradley, 1992; Linton et al., 1989) evaluated the effects of a multi-modal program including the training of problem solving strategies in nurses at risk for developing a chronic pain problem. The authors found a significant improvement with regard to pain intensity, mood, sleep, pain behaviors, activity level and days of sick leave at 6-month follow-up. The multi-modality of the program prevented to detect which treatment ingredients were responsible for the improvement. Philips, Grant and Berkowitz (1991), added a problem solving-like counseling session to graded exposure in acute low back pain patients. At 6-month follow-up, the counseling approach showed a trend towards an earlier return to normal activities and reduced the likelihood of persisting pain difficulties as compared to the same approach without counseling. It was suggested that it would be worthwhile to elaborate counseling sessions in future research. Finally, Wilkinson and Mynors-Wallis (1994) found problem solving therapy to be feasible and acceptable in a pilot study regarding 11 patients with unexplained physical symptoms. In conclusion, the studies above all indicate that problem-solving therapy could be a valuable addition in the treatment of pain patients in an early phase, however, none of the previous studies investigated problem-solving therapy as an isolated ingredient. Consequently, questions may be formulated regarding the specific contribution of problem solving therapy in terms of 1) effects: i) changes in problem solving skills, ii) improvement of effect, iii) reduction of (work) disability; 2) maintenance of behavior; and 3) increased self-control.

Predictors of disability

Identifying predictors of chronic pain may be important in two ways. First, predictors may facilitate the screening of patients who are most at risk for developing a chronic pain problem. Secondly, predictors may give new directions to pain management as they point to potential mechanisms to be addressed in subsequent treatment. Psychosocial factors and factors related to work (e.g. job task enjoyment) have been identified to play an important role in the course of pain (Bigos et al., 1991; Burton & Tillotson, 1991; Gallagher et al., 1989; Gatchel & Turk, 1996; Linton et al., 1994; van der Giezen, Bouter, & Nijhuis, 2000). Moreover, there are strong indications that fear-avoidance, referring to the avoidance of movements and activities based on fear, is important with regard to the prediction of a long-term back pain problem (Crombez et al., 1999; Klennerman et al., 1995; Vlaeyen & Linton, 2000). Kendall, Linton and Main (1997) therefore included pain-related fear as an essential aspect of a broader early assessment of the so-called psychosocial 'yellow flags', providing guidelines for simple assessments already at the first consultation, in order to identify patients at risk in primary care.

In a recent review on psychosocial factors in back and neck pain, Linton (2000) concluded that psychological features, such as stress, mood, emotions, cognitive functioning (passive coping, pain cognition, and fear-avoidance beliefs), and pain behavior played a significant role in the transition from acute to chronic pain. Personality factors, however, resulted in mixed outcomes. Remarkably, psychosocial factors generally have more impact than biomedical and biomechanical factors on back pain disability (Linton, 2000). Linton (2000) stresses the need to sort out which variables are important at what time point, as to map variables involved in the development of pain over time. Furthermore, he stresses the challenge of designing interventions to deal with these factors to provide better and more effective care and prevention.

The effects of secondary prevention of low back pain

Secondary prevention of low back pain is defined as initiatives taken from the point of the first reported injury until a full chronic problem has developed, that is pain that continues after 6 months of health care or sick leave (Linton & Bradley, 1996). One of the first initiatives in this field was the Swedish back school, which is based on a health education model and merely consists of back-education, ergonomics and fitness. Although education and information seem to be a precondition for treatment, there is insufficient evidence to recommend group education only, as an effective treatment in low back pain patients (Cohen et al., 1994; Linton & van Tulder, 2001b). At the same time, a shift took place towards more cognitive-behaviorally oriented interventions to prevent chronicity.

In a recent review by Linton & Van Tulder (2001b), the effects of subsequent educational efforts, lumbar supports, exercises, ergonomics and risk factor modification in non-healthcare settings, i.e. at the workplaces and in the general population, were evaluated. The effects reported in the 27 investigations were very sobering in that only exercises provided sufficient evidence to conclude this form of prevention was effective. In line with this finding, Waddell (1998) concludes in "The back pain revolution", that "...trends of low back pain disability show that current management and health care have failed to solve the problem" (pp. 404). According to Linton and Van Tulder (2001b), the reasons for this failure may be found in the methodology of the studies involved, as well as in the way interventions are designed. As for methodological causes, small sample sizes as well as short follow-up periods may prevent detection of positive effects. Moreover, the preventive effect may be masked by the whimsical nature of the natural course of pain. The main benefit of early interventions will therefore be found in reduced disability levels and days of sickleave, rather than other variables, such as pain. As for the interventions, a multidimensional approach and a proper risk analysis may be needed to cover the wide range of risk factors and to tailor the intervention to the needs of the workplace and the individual, as to increase the program's success. Finally, the authors stress to make efforts to increase compliance rate, as to maintain effects in the long term.

There are several indications that a multidisciplinary approach is effective in the prevention of long-term disability and work loss. In sub-acute work-related back pain, the combination of a clinical and an occupational intervention resulted in a 2.4 times faster return-to-work, the occupational intervention accounting for the most important part of the effect (Loisel et al., 1997). Di Fabio (1995) found back schools to be most efficacious when coupled with comprehensive rehabilitation programs (e.g. work-site visits, general physical conditioning, work hardening, or operant conditioning), regarding pain, physical impairment and compliance outcomes in low back pain patients. This outcome, however, was not supported by outcomes regarding disability and work, which might be caused by power-problems of the studies included in this meta-analysis (Di Fabio, 1995). In contrast with the foregoing, Feuerstein, Menz, Zastowny and Barron (1994) found multidisciplinary rehabilitation successfully facilitating return-to-work for chronic back pain patients. Results of the seven studies included in their systematic review indicated that 71% of the work-disabled patients were working or involved in vocational rehabilitation efforts after 1 year, whereas only 44% were in corresponding comparison groups. Recently, a cognitive-behavioral return-to-work program was found to be more effective than treatment as usual in reducing days of sick leave and increasing general activity levels (Marhold, Linton, & Melin, 2001). However, these effects were only found with regard to patients with short-term sick leave (2-6 months) and not for those with long-term sick leave (>12 months), underscoring the importance for an early return-to-work intervention.

In conclusion, challenges for the prevention of chronic back pain disability and vocational rehabilitation lie in the identification of patients at risk, as early as reasonably possible.

OUTLINE OF THE THESIS

As highlighted in the introductory sections, an alternative focus is introduced in the treatment and prevention of pain, in that therapies aim at the consequences of pain in daily living, rather than at the pain itself. Dealing with complicated problems as a result of pain disability, and at the same time finding satisfactory solutions for the hassles of everyday life, asks for efficient use of considerably complex skills. It is hypothesized that problem-solving skills may prevent failure, by helping patients to cope with the consequences of pain in everyday life. Problem solving has the potential to gain new insight into the mechanism of pain and bring new options for secondary prevention. Whether problem solving is an issue in the etiology and treatment of pain problems, non-specific low back pain in particular, will be the main theme of the research questions addressed in this thesis. Next, main research questions will be explicitly described and shortly commented on. Besides this introductory Chapter, this thesis includes a systematic review, 2 experimental laboratory studies, 2 clinical studies (a cross-sectional study, and a randomized clinical trial), and the general discussion.

1. Is problem solving therapy an effective therapy within the field of mental health care and behavioral medicine, and do outcomes differentiate as a function of methodological quality, therapy form, (moment of) assessment, and study population?

Problem-solving therapy is introduced as a new technology in the treatment of low back pain patients. Although innovative regarding the treatment of pain, problem-solving therapy has a long history of multiple applications within mental health care and behavioral medicine. Moreover, problem-solving therapy is claimed to be an effective treatment, especially with regard to stress-related disorders. This claim, however, has never been investigated systematically. Therefore, Chapter 2 describes the outcomes of a systematic review that includes a structured search of the literature, a systematic methodological quality assessment, and an extensive and controlled data-analysis. The review adds to our knowledge of the efficacy of problem-solving therapy, in that previous reviews were descriptive in nature and referring to one specific problem area, whereas the current review is conducted systematically according to the method guidelines for systematic reviews of the Cochrane Collaboration (Van Tulder et al., 1997). Moreover, the review refers to outcomes of problem solving therapy regarding several problem areas within mental health care and behavioral medicine.

2. Does non pain-related failure result in higher pain reports, lower pain tolerance and higher pain avoidance in a pain-free population?

Does Negative Affectivity affect the impact of failure feedback on pain report, either as a mediator, in the case of negative state affect, or as a moderator when Negative Affectivity as a personality trait is considered?

Stressful events, including major life events or daily 'hassles', are found to be strongly associated with increased pain. Failure experiences, as a non-pain-related source of stress, resulted in increased pain report in a pain-free population in which acute pain was induced (Levine et al., 1993). In Chapter 3 of this thesis, the laboratory experiment of Levine is replicated. Confirmation of the outcomes in the original study are of importance in consideration of the more fundamental question whether failure to solve general life problems, directly or indirectly influences pain and pain behavior. In addition to the original study, the replication wants to find answers to the role of negative affectivity. First, it is hypothesized that failure induces negative mood, and that it is mood and not failure that directly influences pain. Second, negative affectivity is hypothesized to have a moderating role, in that pain is especially influenced when subjects are characterized as 'negative affective'. In addition to pain report, behavioral indices are tested in the replication (Chapter 3).

3. Does non pain-related failure result in higher pain reports, lower pain tolerance and higher pain avoidance in chronic low back pain patients?

Does Negative Affectivity affect the impact of failure feedback on pain report, pain tolerance and pain avoidance either as a mediator, in the case of negative state affect, or as a moderator when Negative Affectivity as a personality trait is considered?

Does pain-related fear enhance pain report and pain avoidance, and decrease pain tolerance?

In analogy with the experimental study described in Chapter 3, a laboratory experiment is executed in chronic low back pain patients (Chapter 4). The influence of failure and success experiences was never studied directly in a chronic pain population. The underlying idea of this experiment was that feelings of failure and frustration about not having control over pain consequences and life in general, might have become a part of everyday life in chronic pain patients. If failure modulates the experience of acute pain, it may as well contribute to the maintenance of pain over time in chronic pain. Again, the hypothesized mediating and moderating roles of Negative Affectivity are tested. Next to failure and negative affectivity, the role of pain-related fear is investigated. In previous studies, pain-related fear was found an important predictor of chronic pain, and is therefore added to the statistical model in this experimental design.

4. Are pain intensity, pain catastrophizing, pain-related fear, daily stress, and problem solving individually related to functional disability in low back pain patients?

Does problem solving have additional value in a model explaining functional disability?

Does problem solving moderate the effects of daily stress on functional disability?

Why do some pain patients succeed in resuming their daily activities within a few days after pain onset, whereas others fail to return to work and gradually develop a chronic pain syndrome? In chapter 5, results of a cross-sectional study are presented. On the one hand, pain-related variables are included in the model to test the hypothesized role of pain-related fear. On the other hand problem solving is introduced as an individual and additional predictor of disability. Problem solving, or the lack of skills to cope with the stressful events of daily living, is now seen as a mechanism in the development of chronic pain that operates independently and next to pain-related variables. Data of a sample of low back pain patients with a new episode of sick leave are analyzed. Functional disability is measured by means of the Quebec Back Pain Disability Questionnaire, a self-report measure. Next to the direct relationship of problem solving with disability, it is tested whether problem solving modulates the influence of daily stress on disability as a trait-like variable.

5. Does problem-solving therapy have a supplemental value when added to behavioral graded activity in employees who have a new episode of sick leave because of low back pain, with regard to negative affect, functional disability, medical consumption, days of sick leave and work status?

Do high-educated patients gain more from the supplemental problem-solving therapy, whereas low-educated patients scarcely profit?

A secondary preventive intervention was developed against the background of a biopsychosocial model of pain and focuses on factors that hypothetically contribute to the chronification of pain. The intervention consists of 2 components: behavioral graded activity and problem-solving therapy. Behavioral graded activity (GA) is hypothesized to be of influence in the process of pain-related fear, whereas problem-solving therapy (PST) merely intervene in the mechanisms of (daily) stress and stress-related avoidance. In a randomized clinical trial patients are assigned to either a GA + PST condition, or to an attention-control condition in which GA is combined with Group Education. Patients are all employees having a recent new episode of sick leave as a result of non-specific low back pain. To control for possible biases, several checks are included in the design, among which a check of the manipulation and of treatment credibility. Outcomes are described in 2 parts. The first part describes effects regarding self-report measures, observational measures, and outco-

mes of a cost-diary. This part is described in Chapter 6. The second part mainly goes into the effects on days of sick leave and work status. The data were obtained by means of sick leave registration of the occupational health services. Outcomes on sick leave are described in Chapter 7.

 

CHAPTER 2

The effects of problem solving therapy in mental health care and behavioral medicine: a systematic review.

This chapter was submitted as:

Van den Hout, J.H.C., Kerckhoffs-Hanssen, M.R., Vlaeyen, J.W.S The effects of problem solving therapy in mental health care and behavioral medicine: a systematic review.

ABSTRACT

To give an answer to the question whether problem-solving therapy (PST) is an effective therapy within the field of mental health care and behavioral medicine, a systematic review was conducted. Twenty-seven randomized controlled trials were systematically selected to be included in a methodological quality assessment, from which 19 studies were included in a structured data-extraction and analysis. Fifty-nine percent of the studies had high methodological quality, whereas only 25% of the studies was defined 'high quality' when internal validity criteria were regarded. Selected studies covered a wide range of populations and measures. Generally, the PST turned out to be most effective in 72% of the studies at both post-treatment and follow-up evaluations. Moreover, the positive effects of the PST were maintained at follow-up in 80% of the studies including both post-treatment and follow-up assessments. The chance of finding positive effects was higher when methodological quality was high. Furthermore, highest success-rates were found regarding problem-solving skills, quality of life, domain specific outcomes and outcomes referring to general improvement. The PST was more effective when including more than 5 sessions in the therapy protocol, and ideally more than 9. The review gives an indication that PST is an effective treatment in depressed patients and children with conduct disorders. In the discussion recommendations are given regarding future research and applications in clinical practice.

INTRODUCTION

In a real-life social context, problem solving (p-s) may be defined as a cognitive-affective-behavioral process through which an individual attempts to identify, discover or invent effective or adaptive means of coping with problems encountered in everyday living (D'Zurilla & Goldfried, 1971; D'Zurilla & Nezu, 1982). By defining problem-solving as the normal ebb and flow of living life, D'Zurilla and Goldfried (1971) postulated a theory that proved to be applicable to a wide variety of problems as useful self-management technique and versatile coping strategy. More specifically, the p-s theory describes problem solving as a cyclic process that consists of 5 steps: 1) problem orientation; 2) problem definition- and formulation; 3) generation of alternative solutions; 4) decision making; 5) solution implementation and verification. In answer to the p-s theory, Spivack (Spivack & Shure, 1974; Spivack, Platt, & Shure, 1976) and Shure (Shure & Spivack, 1981) proposed the so-called 'social competence theory' in which problem solving was linked to deficiencies in social adjustment or psychopathology. According to this theory, deficiencies in interpersonal cognitive p-s skills were hypothesized to result from insufficient learning of such skills due to deficient child rearing or neurological defects, and from 'blockage' of these skills due to interfering emotions. Furthermore it was hypothesized that therapeutic or preventive programs that enhance operation and performance of p-s skills could prevent psychological dysfunction. The 'social adjustment theory' has resulted in multiple applications of p-s, particularly in the field of psychopathology. D'Zurilla (1988) suggested that the most appropriate case for p-s therapy is one in which maladaptive behavior or negative stress effects, such as anxiety and depression, or psychophysiological symptoms (e.g. pain), are associated with deficits in problem solving ability, and/or performance. The emphasis in p-s therapy, however, should always be in the identification and resolution of the current antecedent problems that are causally related to the maladaptive behavior, in that only then the strategy is more likely to produce durable and generalized effects.

Regarding efficacy, the p-s therapy has been claimed to be effective within a wide range of clinical disorders and psychological problems (Nezu, Nezu, Friedman, Faddis, & Houts, 1998). Nezu listed following problems and disorders to profit from p-s therapy: social phobia, agoraphobia, obesity, coronary heart disease, schizophrenia, mentally retarded adults with concomitant psychiatric problems, HIV risk behaviors, drug abuse, suicide, childhood aggression, conduct disorder and depression. Though this list might not be exhaustive, it gives an impression of the variety of problems p-s therapy was effectively applied to. However, some remarks must be made regarding how effectiveness was defined in the corresponding studies. In some studies, regarding social anxiety and agoraphobia, the p-s therapy turned out to be effective, but no more than other cognitive and behavioral treatment conditions in social anxiety (DiGiuseppe, McGowan, Simon, & Gardner, 1990), and even less effective

tive than programmed practice and exposure in agoraphobia (Jannoun, Munby, Catalan, & Gelder, 1980; Jansson & Oest, 1982). As was cautioned by D'Zurilla (1988), one has to recognize when other intervention methods are necessary or more appropriate for a particular case, as may be the case in anxiety disorders. The current systematic review therefore particularly focuses on the comparison of p-s therapy with other conditions, representing alternative treatment or no-treatment controls in order to give answer to the relative efficacy of p-s interventions.

A few reviews were executed to decide about the efficacy of the p-s intervention with regard to one specific domain: children or adolescents with internalizing disorders (mood disorders; Grossman & Hughes, 1992), children and youth with learning and behavioral problems (Coleman, Wheeler, & Webber, 1993), conduct disorder in children and adolescents (Kazdin, 1997), college students at risk (Heppner, Neal, & Larson, 1984), and patients with major depression and more broadly defined emotional disorders in primary care (Gath & Catalan, 1986; Mynors Wallis, 1996). All but one review, showing selective efficacy (Coleman et al., 1993), found the p-s intervention to be effective with regard to one specific domain.

Although there seems to be more evidence in favor of the p-s therapy than against it, two critical reviews of the literature (Larson, 1990; Tisdelle & St Lawrence, 1986) warn against premature conclusions and give directions where to go from here. First, the need for more behavioral assessment and evaluation of a treatment's clinical validity is called for, because not all achievements in the cognitive domain may generalize to behavior. Second, the lack of follow-up assessments was found a common deficiency in p-s literature, the more because more rigorous follow-up data could articulate the degree and duration of treatment effects. Furthermore, research on problem solving should take a closer look at the most effective and efficient ways to achieve desired cognitive and behavioral outcome goals. Studies comparing direct behavioral approaches to cognitive p-s would be helpful in this respect. Larson (1990) proposes 3 questions to be addressed by future research, these are the contrast between individual versus group p-s interventions, the search for individual difference variables (moderators of effect), and finally the focus on problem attributes (e.g. source, cause, onset, difficulty; Heppner & Krauskopf, 1987) and its relation to treatment effects. Both reviews (Larson, 1990; Tisdelle & St Lawrence, 1986) refer to the need for an operational definition of real-life p-s, and adequate assessment materials. Furthermore, systematic (dismantling) research is needed to isolate the most important components for specific populations (Tisdelle & St Lawrence, 1986).

In addition to the reviews mentioned above, which were mainly descriptive in nature, the current systematic review uses a structured search of the literature and a systematic evaluation of the methodological quality of the selected studies. Then, an extensive data analysis will be executed to answer the

main research question whether p-s therapy is effective within the field of mental health care and behavioral medicine. Particular issues that will be addressed are methodological quality of the studies, outcome measures, follow-up effects, number of p-s sessions, individual versus group interventions, choice of the reference treatment condition, and efficacy of p-s therapy in particular populations within mental health care or behavioral medicine. To give answer to these questions, only randomized controlled trials were selected as a design.

METHODS

SEARCH STRATEGY

Potentially relevant RCTs are identified by a computer-aided search of the MEDLINE database for the period 1971 - January 1999, and the PSYCHLIT database for the period 1971 - January 1999. The MEDLINE search is based on the first two stages of the MEDLINE search strategy recommended by the U.K. Cochrane Handbook, regarding the selection of randomized controlled trials (Dickersin, Scherer, & Lefebvre, 1994; van Tulder, Assendelft, Koes, & Bouter, 1997a). This search is in conjunction with a specific search for PROBLEM SOLVING THERAPY, -TREATMENT, -TRAINING, or -INTERVENTION. Some small adaptations are made to make the MEDLINE search applicable to the PSYCHLIT database.

INCLUSION CRITERIA

In order to answer the research questions, the following inclusion criteria are set for data selection: 1) The study is published later than 1971, the year D'Zurilla and Goldfried published their problem solving theory; 2) The study design is randomized and controlled (RCT); 3) The study population consists of patients in the field of mental health care or behavioral medicine, suffering from a medical illness, behavioral disturbances, psychopathology, or being a population at risk; 4) The total number of participants per treatment condition is higher than 10; 5) The study is an effect evaluation and outcome measures are quantified; 6) The problem solving therapy is evaluated as a single component, being evaluated individually or as the additional component in a dismantling study; 7) An underlying theory is specified in that 2 or more steps from the problem solving theory (D'Zurilla & Goldfried, 1971) can be identified; 8) Problem solving skills are actually trained. If the model merely functions as a structure but no skills are trained, the study is not included in the review.

STUDY SELECTION

The stepwise study selection procedure is presented in figure 1. Two reviewers (JV, JvdH), independently apply inclusion criteria 1 - 3 and 5 - 6, to select the potentially relevant trials from the references retrieved by the literature

FIGURE 1. Study selection.

search. Criteria 4, 7 and 8 are for practical reasons reassessed by the same reviewers executing the methodological quality assessment (MK, JvdH). Methodological quality assessment and data extraction are proceeded when criteria 8 and 9 are met. A consensus method is used to solve disagreements between the two reviewers regarding the inclusion of studies. In case disagreement is not resolved in the consensus meeting, a third reviewer (JV) is consulted. Moreover, if titles, abstracts or keywords provide insufficient information to assess whether criteria 1-6 are met, the original article is obtained.

METHODOLOGICAL QUALITY ASSESSMENT

The methodological quality of the studies is assessed according to the criteria listed and explicitly described in table 1. The list contains all methodological criteria considered important by the Editorial Board of the Cochrane Back Review Group. Criteria can be clustered in 4 groups: patient selection; interventions; outcome measurement; and statistics. The list has previously been used in reviews concerning medical interventions (Koes, Assendelft, van der Heijden, Bouter, & Knipschild, 1991a; Koes, Bouter, Beckerman, van der Heijden, & Knipschild, 1991b; van der Heijden et al., 1995; van Tulder, Assendelft, Koes, & Bouter, 1997b; van Tulder, Koes, & Bouter, 1997c). Some applications have been made to define criteria more suitable for psychotherapeutic interventions (van Tulder et al., 2000). Criteria about blinding the therapists and adverse effects of therapy were deleted from the original criteria list. It is difficult to blind the therapists because therapists are aware of the kind of therapy they are applying and whether the therapy is experimental or not.

TABLE 1. Methodological quality criteria

METHODOLOGICAL CRITERIA LIST			
a. Specification inclusion/exclusion criteria	+	-	?
b1. Randomization treatment allocation *	+	-	?
b2. Concealment of treatment allocation *	+	-	?
c. Similarity of baseline characteristics *	+	-	?
d1. Description experimental intervention(s)	+	-	?
d2. Description control intervention(s)	+	-	?
e. Co-interventions comparable *	+	-	?
f. Compliance rate *	+	-	?
g. Blinding of patients *	+	-	?
h. Blinding of outcome assessor *	+	-	?
i. Sort of outcome measures	+	-	?
j. Withdrawal/drop-out rate *	+	-	?
k1. Description short-term follow-up measurement	+	-	?
k2. Length of long-term follow-up measurement *	+	-	?
l. Identical timing of the outcome assessment	+	-	?
m. Adequate sample size for each group	+	-	?
n. Intention-to-treat analysis *	+	-	?
o. Presentation point estimates and distribution measures	+	-	?

Note: *internal validity criteria. + 'positive'; - 'negative'; ? 'don't know/unclear'.

It is unusual to report adverse effects in psychotherapy research because adverse effects can not be identified as such. This criterion is therefore regarded not relevant to our study. If it was not feasible to blind the patients, criteria *g* of the criteria list is scored positive when treatment credibility was adequately evaluated and treatments turned out to be equally credible and acceptable to the patients. With regard to co-interventions, the criterion is redefined in that co-intervention had to be described but did not necessary to be avoided to score a 'positive'. Criterion *d* is divided into two criteria regarding description of experimental and reference interventions, separately. Each criterion is scored as 'positive', 'negative', or 'don't know/unclear' when methodological data can not be retrieved from the article. Total scores for methodological quality and internal validity are calculated by summing the positive scores. In accordance with Van Tulder et al. (2000), a study is judged 'high quality' when more than 50% of the criteria are scored positive. In the current study this is a score higher than 5 regarding internal validity, and higher than 9 regarding total methodological quality (internal and external quality).

Two reviewers (MK, JvdH) perform the methodological quality assessment independently. A consensus model is used to solve disagreements about cri-

teria the reviewers disagree about and a third reviewer is consulted when disagreements persist. The articles are blinded for authors, institution, journal, and references, before quality assessment takes place.

DATA EXTRACTION

Data are extracted with regard to 1) population; 2) interventions; 3) outcome measures and outcomes. Two reviewers (MK, JvdH) extract outcome measures and outcomes independently. Because study populations are various, outcome measures consequently vary widely. Therefore, outcome measures are clustered in 5 categories: problem-solving skills (PS), negative affect (NA), functional status/quality of life (QOL), domain specific outcomes (DS), and general improvement (GI). Definitions of these 5 clusters can be found in table 2¹. Reviewers cluster outcome measures independently and agreement is checked according to the consensus model as previously applied during the methodological quality assessment.

DATA ANALYSIS

Because study populations and outcome measures are very divers, describing different kinds of effects, it was decided that it would not be appropriate to conduct a meta-analysis. Instead, reviewers independently judge whether the experimental condition, including problem-solving therapy, is less effective than (<), more effective than (>), or equal to (=) the reference condition, at post-treatment and follow-up measurements (table 4). The reference treatment can either be no treatment, treatment as usual (standard treatment), or wait-list control (all indicated with an "R"), or another treatment or placebo-control (all indicated with an "I"). It was decided that treatments were subdivided in the latter category when the control treatment was individually protocolized. The condition including problem-solving therapy is indicated with an "I1". Data-interpretation consists of three stages. First, post-treatment outcomes and, if available, outcomes of the longest follow-up are extracted per outcome measure. Second, reviewers decide independently whether outcomes per cluster are favoring (+) or disapproving (-) the problem-solving therapy condition, or neither (0) (table 5). Only outcomes concerning the target population or the mediators are counted in this part of the evaluation. The rule is that the direction of the majority of outcomes per cluster is decisive. In case there is no majority, the direction (+ or -) is given the benefit of the doubt. The third step in this procedure is a general evaluation of the outcome-clusters at post-treatment and follow-up. Again the majority of directions within each study are decisive and in case there is no majority, the direction is given the benefit of the doubt. Moreover, outcomes on the PS-cluster are not counted in case of indecision, because PS-skills are seen as a process measure rather than an effect measure.

Note 1 The original outcome measures, and the clusters they were categorized into, can be obtained from the authors.

TABLE 2. Definition of outcome clusters.¹

CLUSTER (ABBREVIATION)	DEFINITION
Problem-Solving Skills (PS)	Measures assessing proper problem-solving skills: problem orientation, problem definition, goal setting, generation of alternatives, decision making, implementation, and evaluation skills.
Negative Affect (NA)	Measures assessing negative affects: depression, anxiety, tension, anger, self-esteem, -confidence.
Functional Status/ Quality of Life (QOL)	Measures assessing overall physical or mental well being / functioning. Items refer directly to quality of life or functioning.
Domain Specific (DS)	Measures that are only useful in a particular group of patients. The themes questioned in the measures refer to the specific domain the intervention is focussed at.
General Improvement (GI)	Measures that point at general improvement or effects, and cannot be classified in one of the other clusters.

During data analysis the following questions are of interest: 1) Does outcome differ as a result of methodological quality? 2) Is the p-s intervention more effective in one cluster of outcome measures than in others? 3) Are outcomes depending on the number of sessions invested in the p-s intervention? 4) Is there a difference between individual or group interventions? 5) Is the choice of the reference treatment determining the direction of the effect? 6) Is the p-s intervention effective in one particular field or population within mental health care and behavioral medicine?

RESULTS

STUDY SELECTION

The computer-aided literature search collected 35 references from Medline and 68 references from PsychLit databases. From these 103 references, 64 were included after a first selection that was based on titles, keywords and abstracts. 14 of these 64 references were collected in both Medline and PsychLit databases. Reviewers disagreed about 10 references, what was subsequently dis-

cussed and resulted in the rejection of 5 references. Another 11 references were excluded from the review, because both reviewers decided from the abstract that the study did not refer to the problem solving model by D'Zurilla and Goldfried (1971). Consequently, 34 blinded copies of the full papers were assessed regarding criteria 7 and 8 (reviewers JvdH and MK). In 3 studies there was uncertainty or disagreement between the reviewers about criterion 8 and in one study about criterion 2 (randomization). Therefore, a third reviewer (JV) was consulted who included 3 studies and excluded one. In the excluded study patients were not randomly assigned to the treatment conditions, what only became clear after reading the full paper. Seven studies were excluded after the full paper-screening procedure, resulting in a total of 27 studies. From these studies, another 8 studies turned out to include conditions with less than 10 participants. It was decided that these studies would be excluded from the final data analysis, but not from the methodological quality assessment. In the end, 19 studies were included in this review, in 2 cases the therapy was pointed at spouses or parents (Blanchard, Toseland, & McCallion, 1996; Graves, Meyers, & Clark, 1988). In one study (Kazdin, Siegel, & Bass, 1992) both children and parents were treated and took part in the assessment procedure.

METHODOLOGICAL QUALITY

Table 3 shows the results of the methodological quality assessment. After consensus, 22% of the 27x10 (internal validity criteria) and 12% of the 27x18 (total) quality assessments were scored unclear. Regarding total scores on the methodological quality scale (range: 0-18), 59 % of the studies shows methodological quality higher than 9 which was the preset threshold for high quality. However, when internal validity criteria were taken into account, only 26% of the studies were labeled 'high quality'. Remarkably, internal validity criteria regarding 'similarity of baseline characteristics' (c) and 'withdrawal/drop-out rate' (j), were the only criteria scored 'positive' in respectively 50 and 75% of the studies. On the other hand, criteria regarding 'concealment of treatment allocation' (b2) and 'compliance rate' were rarely described by the authors and therefore scored 'unclear' in 85 and 70% of the studies, respectively.

DATA EXTRACTION

In table 4 (page 12-17), characteristics of 19 studies meeting the inclusion criteria are displayed. The studies were published between 1975 and 1998, and covered a wide range of study populations within the field of mental health care and behavioral medicine. Dropout rate ranged from 3 to 53% (mean = 24%). One study did not report dropout rate.

With regard to the study's design, 8 studies compared problem-solving therapy with no treatment, treatment as usual, or wait-list control (a), 7 studies compared problem-solving therapy with another treatment or placebo-control (b), and 4 studies compared problem-solving therapy with no treatment (a), as well as with another treatment (b). In total, the 19 studies made 28 compar-

TABLE 3. Methodological quality assessment.

Study	a	b1*	b2*	c*	d1	d2	e*	f*	g*	h*	i	j*	k1	k2*	l	m	n*	o	internal validity	methodological quality
Arean et al., 1993	+	+	?	+	+	+	-	?	+	+	+	-	+	-	+	+	-	+	4	12
Atha et al., 1992	-	+	?	?	+	-	+	?	-	?	+	+	+	+	+	-	-	-	4	8
Blanchard et al., 1996	+	+	?	+	+	+	?	?	-	-	+	-	+	-	+	+	-	+	2	10
Catalan et al., 1991	+	+	+	+	+	-	+	?	-	+	+	+	+	+	+	-	-	+	7	13
Clarke & Greenberg, 1986	+	+	?	?	+	+	-	?	-	-	-	-	+	-	+	-	-	+	1	7
Coche & Flick, 1975	+	-	?	+	+	+	-	+	-	?	+	+	+	-	+	-	-	-	3	9
Coche & Douglas, 1977	+	-	?	+	+	+	-	-	?	-	+	-	+	-	+	+	-	-	1	8
DeVellis et al., 1988	+	+	?	+	+	-	-	?	?	+	+	+	-	-	+	+	-	-	4	9
Foster, Prinz, & O'Leary, 1983	-	+	?	-	+	+	-	+	-	+	+	+	+	-	+	-	-	+	4	10
Graves et al., 1988	+	+	?	+	+	+	-	?	+	?	+	+	+	+	+	-	-	+	5	12
Hussian & Lawrence, 1981	+	+	?	+	+	+	-	?	+	?	+	-	+	-	+	-	-	-	3	9
Kazdin et al., 1987	+	+	?	+	+	+	+	?	+	-	+	+	+	+	+	-	-	+	6	13
Kazdin, 1989	+	+	?	+	+	+	+	?	+	-	+	+	+	+	+	+	-	+	6	14
Kazdin et al., 1992	+	+	?	+	+	+	-	?	+	-	+	-	+	+	+	+	-	+	4	12
Lynch et al., 1997	+	+	?	+	+	-	+	?	?	+	+	-	+	-	+	-	-	-	4	9
McLeavey et al., 1994	+	+	?	+	+	+	-	+	+	+	+	-	+	+	+	-	-	-	6	12
Mendonca & Siess, 1976	-	+	?	-	+	+	-	?	?	?	+	+	+	-	+	-	-	+	2	8
Mueser, Valenti Hein, & Yarnold, 1987	+	+	?	+	+	+	-	?	+	?	+	-	+	-	+	-	-	-	3	9
Mynors Wallis et al., 1995	+	+	+	+	+	+	-	+	+	+	+	+	+	-	+	+	+	+	8	16
Mynors Wallis et al., 1997	-	+	+	?	-	+	+	-	-	+	+	+	+	+	+	+	-	+	6	12
Nezu, 1986	+	+	?	+	+	+	-	?	+	-	+	+	+	+	+	-	-	+	5	12
Nezu & Perri, 1989	+	+	?	+	+	+	-	?	+	+	+	+	+	+	+	-	-	+	6	13
Pfiffner, Jouriles, Brown, Etscheidt, & et al., 1990	+	+	?	+	+	+	-	?	+	+	+	-	+	-	+	-	-	-	4	10
Platt et al., 1993	-	+	?	+	+	+	-	?	-	-	+	-	+	+	+	+	-	-	3	9
Salkovskis, Atha, & Storer, 1990	+	+	?	+	+	+	-	?	-	-	+	+	+	+	+	-	+	-	5	11
Schwartz et al., 1998)	+	+	?	+	+	+	-	-	?	-	+	-	+	-	+	+	-	+	2	10
Yu, Harris, Solovitz, & Franklin, 1986	-	-	-	-	+	+	-	+	-	?	+	-	+	-	?	-	-	+	1	6

Note: *internal validity criteria. + 'positive'; - 'negative'; ? 'don't know/unclear'.

TABLE 4. Study Characteristics

STUDY/PARTICIPANTS	INTERVENTIONS	OUTCOMES
<p>1. AREAN ET AL., 1993</p> <p>•75 older adults with unipolar major depression, aged 55-80 years, recruited through media announcements. Sixteen subjects dropped out of treatment</p>	<p>I1: Problem Solving Therapy: general goal is to cope with depression and associated problems according to 5 problem-solving steps (n=28).</p> <p>I2: Reminiscence Therapy: Review life histories, discuss and interpret positive and negative life-events, future goal-setting (n=27).</p> <p>I1 & I2: 12 weekly group sessions of 1,5 hours, guided by advanced graduate students in clinical psychology.</p> <p>R: Waiting List Control: two individual contacts during waiting list period of 12 weeks (n=20).</p>	<ul style="list-style-type: none"> • (PS): I1 > I2 = R • (NA): I1 > I2 > R • (QOL): I2 > I1 = R <p>Effects were maintained at 3 months follow-up.</p>
<p>2. ATHA ET AL., 1992</p> <p>•33 patients attending a medical emergency department (A&E) for medical reasons and 1) felt the need to talk about worries after going home; 2) had difficulties coping problems of everyday life, aged between 16 and 65. One subject dropped out.</p>	<p>I: Treatment as usual + Psychological treatment based on the problem-solving approach, 5 individual sessions of 1 hour home-based treatment, guided by an experienced community psychiatric nurse, goal is to identify specific problems and apply 4 rational p-s skills (n=16).</p> <p>R: Treatment as usual: individual (n=17).</p>	<ul style="list-style-type: none"> • (NA): I = R; in women: I > R • (GI): I = R <p>All outcomes are based on 1-year follow-up assessments (repeated measures ANCOVA).</p>
<p>3. CATALAN ET AL., 1991</p> <p>•47 patients who had recent onset (<12 weeks) complaints of anxiety, tension, depressed mood, irritability, sleep disturbances or somatic symptoms, high risk, aged 18-65 years, referred by GPs. Three patients dropped out in the control treatment.</p>	<p>I: Four individual sessions (on the average 30 minutes per session) of problem solving, guided by a psychiatrist/therapist, contents: 1) assessment of the problem, 2) explanation symptoms, 3) p-s in 5 stages, 4) rehearsal of techniques (n=21).</p> <p>R: Treatment as usual: any treatment chosen by the GP during 6-weeks trial (n=26).</p>	<p>LRG: Low Risk Group (n=66)</p> <ul style="list-style-type: none"> • (QOL): I = LRG > R • (GI): I = LRG > R <p>Effects were maintained at 24 weeks follow-up.</p>
<p>4. CLARKE & GREENBERG, 1986</p> <p>•48 adults who voluntarily sought counseling to help them resolve an intrapersonal conflict over a personal decision, aged 16-73 years, fairly well educated. Twenty-two patients dropped out after randomization.</p>	<p>I1: Problem-solving treatment: 1) problem definition, generation of alternatives, deciding, 2) verification and evaluation (n=16).</p> <p>I2: Gestalt two-chair intervention: Client alternately assumes two sides of an intrapsychic conflict, the counselor guides in a dialogue between the two sides (n=16).</p> <p>I1 & I2: general orientation in a group, 2 individual sessions, debriefing, guided by advanced graduate students and registered psychologists.</p> <p>R: Waiting List Control: no treatment (n=16).</p>	<ul style="list-style-type: none"> • (DS): I2 > I1 > R <p>All post-treatment assessments, no follow-up.</p>

5. COCHE & FLICK, 1975

• 104 hospitalized psychiatric patients admitted to a small private psychiatric hospital, aged 21-65 years, divided into high- and low-schizophrenia groups (median MMPI, Sc-scale). Only patients who completed pre- and posttest are counted. Dropout rate: ?

6. COCHE & DOUGLAS, 1977

• 158 hospitalized psychiatric patients admitted to a small private psychiatric hospital, aged 21-65 years. Eighty-three patients dropped out after pre-test: 12 in I1, 36 in R1, and 35 in R2.

7. DEVILLIS ET AL., 1988

• 111 Rheumatoid Arthritis patients enlisted from the rheumatology hospital clinic, aged 23-80 years. Ten patients were lost to follow-up.

8. KAZDIN ET AL., 1987

• 56 psychiatric inpatient children, referred to the hospital for anti-social behavior (>98th percentile of the Aggression and Delinquency Scale of the CBCL), aged 7-13 years. Nine subjects dropped out during the study.

I1: Hospital program + problem solving group: 1) bringing up a problem, 2) clarifying, 3) bringing up alternative solutions, 4) discussing feasibility (n=41).

I2: Hospital program + Play-reading group (attention placebo):

Goal of the reading-play group is to find enjoyment and companionship, 15 minutes discussion (n=23).

I1 and I2: Eight group sessions of one hour, for 2 weeks.

R: Hospital program (n=40).

I1: Hospital program + problem solving group: 1) bringing up a problem, 2) clarifying, 3) bringing up alternative solutions, 4) discussing feasibility (n=37).

I2: Hospital program + Play-reading group (attention placebo):

Goal of the reading-play group is to find enjoyment and companionship, 15 minutes discussion (n=57).

I1 and I2: Eight group sessions of one hour / 2 weeks.

R: Hospital program (n=64).

I: Problem solving intervention: teach patients generic problem-solving in 6 major steps, one hour individual session at the clinic or patient's home + telephone follow-up 2 weeks after P-S intervention (n=56).

R: No intervention (n=55).

I and R: Psychosocial interview preceding the randomization to assess problems that arthritis has caused in patients' lives and to identify actual and potential resources.

I1: Problem-solving skills training: combines cognitive and behavioral techniques to teach P-S skills that the child can use to manage interpersonal situations, response cost approach (n=20).

I2: Relationship therapy: developing a close relationship with the child, providing empathy and unconditional positive regard, helping the child to express feelings, non-directive, response reward approach (n=19).

I3: Treatment-contact control: providing special treatment meetings with a therapist, but not provide therapy as in I1 or I2 (n=17)

I1, I2 and R: 20 individual sessions of approximately 45 minutes in I1 and I2, and 20 minutes in R, 2 to 3 sessions per week, guided by experienced (1-2 years) specialized therapists.

• (PS): I1 > I2 = R

• (GI): I1 = I2 > R

All post-treatment assessments 21 days after pre-testing, no follow-up.

• (PS): I1 = I2 = R

• (NA): I1 = I2 = R

• (QOL): I1 = I2 > R

All post-treatment assessments 3 weeks after pre-testing, no follow-up.

• (NA): I = R

• (QOL): I = R

• (DS): I = R

• (GI): I > R

All post-treatment assessments 4 months after pre-testing, no follow-up.

• (QOL): I1 > I2 = I3

• (GI): I1 > I2 = I3

All post-treatment assessments.

Effects were maintained 1-year follow-up.

9. KAZDIN ET AL., 1989

• 112 children with severe antisocial behavior (>90th percentile of either Aggression or Delinquency Scale of the CBCL), referred to the Child Conduct Clinic for in- and/or outpatient treatment, aged 7-13 years. Nineteen subjects were lost to one-year follow-up.

10. KAZDIN ET AL., 1992

• 97 children with aggressive and antisocial behavior (>90th percentile of either Aggression or Delinquency Scale of the CBCL), referred to the Child Conduct Clinic for outpatient treatment, aged 7-13 years. Twenty-seven subjects were lost to one-year follow-up.

11. LYNCH ET AL., 1997

• 29 mildly depressed patients (HRSD) in a family residency practice, mean age 48.5 years. Eight subjects dropped out during the procedure, while another 8 subjects filled out only the HRSD.

I1: Problem-solving skills training (PSST): combines cognitive and behavioral techniques to teach P-S skills that the child can use to manage interpersonal situations, includes opportunities to individualize the focus, response cost approach (n=37).

I2: PSST + in vivo practice: Stepwise practice outside the treatment setting with increasing difficulty of so-called "supersolvers" (n=38).

I3: Relationship therapy: developing a close relationship with the child, providing empathy and unconditional positive regard, helping the child to express feelings, non-directive, response reward approach (n=37).

I1/(PSST) and I3: 25 individual sessions of approximately 50 minutes, guided by experienced therapists (master's degree in a mental health-related field).

I1: Problem-solving skills training (PSST): combines cognitive and behavioral techniques to teach P-S skills that the child can use to manage interpersonal situations, token reinforcement, parental involvement in the training, in vivo practice, 25 individual sessions of 50 minutes once a week (n=29).

I2: Parent Management Training: treatment program originally drew on procedures described by Patterson et al. (1975), training behavioral techniques, reinforcement program, involvement of the child, 16 individual sessions of 1.5 to 2 hours (n=31).

I3: I1 + I2: (n=37).

I1, I2, and I3 all lasted between 6 and 8 months and were guided by experienced therapists (master's degree in a mental health-related field).

I: Problem-solving therapy: based on protocol by Nezu et al. (1986), training 5 steps of P-S with the goal of developing more useful methods for dealing with life problems, 6 individual sessions of telephone counseling for 20 minutes once a week, homework assignments, guided by supervised student therapists (n=15).

R: no treatment (n=14).

• (NA): I1 = I2 = I3

• (QOL): I1 = I2 > I3

• (DS): I1 = I2 > I3

• (GI): I1 = I2 = I3

Results at one-year follow-up comparable to post-treatment assessments.

• (QOL): I3 > I1 > I2

• (DS): I3 > I1 > I2

Post-treatment effects were maintained at one-year follow-up. I1 equalized I3 at one-year follow-up.

• (PS): I = R

• (NA): I > R

• (QOL): I > R

All post-treatment assessments, no pre-treatment assessment.

12. MCLEAVEY ET AL., 1994

• 39 patients admitted to the Casualty Department of a regional hospital following self-poisoning. Only patients who completed full pre- and post-assessments were included (39 of the 91 patients initially included). Average age was 24.5 years (SD=7). Follow-up data were available with regard to 33 patients.

13. MYNORS WALLIS ET AL., 1995

• 91 patients with major depressive disorder according to the research diagnostic criteria, HRSD > 13, referred by a GP, aged between 18 and 65 years. Only those patients completing 4 or more sessions were included in the analysis (n=82), of these 82 patients 17 were lost to 12-week follow-up.

14. MYNORS WALLIS ET AL., 1997

• 70 patients presenting at the GP with an emotional disorder of at least one month's duration, complaints of anxiety, tension, depressed mood, irritability, sleep disturbance, somatic symptoms not due to physical disorder, aged 18-65 years, GP decided about inclusion. Twelve patients were lost to follow-up.

I1: Interpersonal Problem-Solving Skills Training: progressive training of the 5 stages of D'Zurilla and Goldfried (1971), additional behavioral methods for skills training (i.e. role plays), 5 or 6 individual sessions of one hour guided by a clinical psychologist (n=19).

I2: Brief problem-oriented approach (crisis intervention): patient is helped to solve current problems in a practical and common sense manner, no specific skills training, treatment was terminated when presenting problems were satisfactorily solved, on the average 4.2 individual sessions guided by a clinical psychologist (n=20).

I1: Problem-solving treatment: rationale on relation between emotional symptoms and problems in daily living, choice of one target problem, instruction and application of PS-stages (n=30).

I2: Drug treatment: information on principles of drug treatment for depression, amitriptyline administration with an increasing dose up to 150 mg, non-specific interventions (listening, encouraging, sympathizing) without giving advice about how to manage problems (n=31).

I3: Placebo drug treatment: See I2, amitriptyline is replaced by a placebo drug (n=30).

I1, I2, and I3: 6 or 7 individual sessions of 30 minutes, first session 60 minutes, given in the patient's home or local health center, guided by a psychiatrist or a trained GP.

I: Treatment-as-usual + Problem solving treatment: a) usual GP-consultations (mean = 2.1), 9 patients started or continued antidepressant medication; b) explanation of the treatment and its rationale, application of 5 PS-stages, encouraging patients to formulate practical ways of dealing with problems, at maximum 5 individual sessions (mean = 3.6 sessions, initial session of 1 hour, subsequent sessions of half an hour), guided by trained nurse therapists (n=40).

R: Treatment-as-usual: usual GP-consultations (mean = 2.2), 13 patients started or continued antidepressant medication (n=30).

• (PS): I1 > I2

• (NA): I1 > I2

• (GI): I1 > I2

• Self-poisoning activities (GI):
1-year follow-up: I1: 10.5% vs I2: 20%

All post-treatment effects, At 6-month follow-up, effects were maintained regarding PS and GI. PS-effects were clinically significant whereas NA-effects were not.

• (NA): I1 = I2, I1 > I3

• (QOL): I1 = I2, I1 > I3

• Clinical relevant recovery at 12-weeks regarding NA: I1 = I2 > I3
All outcomes assessed 12 weeks after pre-treatment (= post-treatment).

• (QOL): I = R, at post-test and 26-week follow-up

• During eight-week treatment period: Costs (GI): I < R;
Disability days (GI): I = R

• Over 26 weeks from trial entry: Costs and Disability days (GI): I > R

15. NEZU & PERRI, 1989

• 43 subjects with unipolar depression for at least 4 weeks duration (BDI ≥ 20), recruitment by community newspapers, mean age 45.8 years (SD = 11.4). Four patients were lost to follow-up.

I1: Problem-solving therapy: entire problem-solving program as described by D'Zurilla and Nezu (1982; Nezu et al., 1989) a) problem-orientation: emotions as cues, inhibit impulsive responses; b) PS-skills: problem-definition and -formulation (1), generate solutions (2), evaluate potential consequences and select most optimal solution (3), monitor and evaluate solution (4), implementation of the model, homework assignments and practice (n=15).

I2: Abbreviated problem-solving therapy: training PS-skills (see I1 part b), therapists refrain from part a) (n=15).

I1 and I2: 10 weekly group sessions of 1.5 - 2 hours, guided by a pair of advanced clinical psychology graduate students.

R: Wait-list control: two telephone contacts during 10-week period to assess need for referral and 'support' (n=13).

• (PS): I1 > I2 = R

• (NA): I1 > I2 > R

• (GI): I1 > I2 > R

I1 vs. I2: all post-treatment effects were maintained at 6-month follow-up.

Clinically meaningful changes at post-treatment regarding NA-measures: I1 > I2 > R

16. PLATT ET AL., 1993

• 130 unemployed methadone clients from five methadone clinics, baseline characteristics are not described. Initially, 277 clients enrolled in the study, but data were only analyzed for the 130 clients who completed all data collection points.

I: Clinical services + Interpersonal cognitive problem-solving: a) daily methadone administration + weekly individual counseling sessions, on average 75 minutes per week; b) the cognitive process of analyzing and resolving interpersonal problems, teaches PS-skills, and application towards the concept of vocational problems, 10 group sessions, guided by a master's level leader (n=67).

R: Clinical services: see I, part a) (n=63).

• (DS: Employment rate): I1 > I2 at 6-month follow-up, and I1 = I2 at 12-month follow-up.

17. SCHWARTZ ET AL., 1998

• 341 women, who had a first-degree relative with a recent diagnosis of primary breast cancer, aged 19 - 75 years. Initially, 967 eligible women completed a baseline interview, 455 accepted and completed a visit at one of the cancer center sites, 341 returned questionnaires at 3-month follow-up. Data of these 341 women were reported.

I1 and I2: a) review and discussion of personal breast cancer risk factors, I1 and I2 consisted of one 2-hour individual session with a health educator.

I1: Problem-solving training: a) + introduction of 5 PS-techniques targeted towards problems associated with personal risk factors and reducing chances of getting breast cancer, recommendations, guidelines and literature (handouts), a 20-minute booster phone call to review PS-techniques (n=197).

I2: General Health counseling: a) + an interview assessment of current health practices, diet suggestions, recommendations for regular physical activity, encouragement to quit smoking, handouts with strategies for behavior change, booster phone call to review information (n=144).

• (NA): I1 = I2

• (DS): I1 = I2

Regarding pre-treatment to 3-month follow-up changes.

Regular practice of the PST-techniques resulted in significant improvement regarding DS-outcome: I1-high practice > I1-low practice = I2

18. BLANCHARD ET AL., 1996

•86 cancer patients and their spouses were recruited in the waiting room of a regional medical oncology clinic, patient's cancer had been diagnosed more than 3 months before recruitment, mean age of patients and their spouses was 52 years. Twenty-nine patients and/or spouses were lost to follow-up, analyses are based on patients and spouses who completed all measurements (n=57, and n=66).

I: Coping with cancer protocol: a short-term problem-solving approach, focus on 6 PS-strategies to reduce or manage specific problems underlying spouses' distress, apply on the most pressing problem(s), 6 one-hour individual sessions, guided by an experienced oncology social worker (n=30).

R: Treatment-as-usual: any services offered by the oncology practice including individual and group counseling (n=33).

• (NA): I = R (spouses),
I > R (patients)

• (QOL): I = R

• (GI): I = R

NA- and QOL-measures were assessed in spouses and patients. All outcomes are based on pre-treatment to 6-month follow-up assessments (repeated measures ANOVAs).

19. GRAVES ET AL., 1988

•40 obese children and their parents, 20% overweight of child regarding age, sex, and height, recruited through media announcements, aged 6 –12 years. Nine patients were lost to 6-month follow-up.

I1 and I2: a) eight weekly parent/child group sessions of 60 minutes: reviewing food-intake and physical activities (10 min.), weighting children (10 min.), b) presentation of several behavioral weight-reduction methods (20 min.),

I1: Problem-solving group: a) + b) + problem-solving exercises targeted towards problems associated with children's weight control, parents together with their children identify potential problems, list and evaluate alternative solutions, develop and evaluate plans (n=?).

I2: Behavioral group: a) + b) + exchange of recipes and exercise ideas among parents (20 min.) (n=?).

I3: Instruction-only group: see I1 and I2 part a) + exercising (15 min.) + exchange of recipes and exercise ideas among parents (20 min.), self-monitoring of food-consumption and exercise (n=?).

• (PS): I1 > I2 = I3

• (DS): I1 = I2 > I3

At 6-month follow-up, PS and DS (body weight) outcomes were reassessed and showed the same outcomes.

Note: PS: problem-solving skills, NA: negative affect, QOL: functional status/quality of life, DS: domain specific outcomes, GI: general improvement; I1: p-s intervention; I2, I3: other treatment; R: no treatment-reference; >: more effective than; <: less effective than; =: equally effective as.BDI: Beck Depression Inventory; CBLC: Child Behaviour Checklist; GP: General Practitioner; HRSD: Hamilton Rating Scale for Depression; MMPI-Sc: Minnesota Multiphasic Personality Inventory, Schizophrenia-scale.

isons of problem-solving therapy with a reference treatment, in 13 cases this was no treatment (a), and in 15 cases this was another treatment (b). Two dismantling studies (Kazdin et al., 1989; Nezu & Perri, 1989), compared 3 treatment conditions of which 2 included p-s therapy. In the general outcome review (table 5) the p-s conditions were compared to the reference treatment separately. In another study (Kazdin et al., 1992), the p-s intervention was compared to parent management training and a third intervention that included both. In the general outcome review, the third condition was left out of the analysis.

In 14 out of 21 p-s conditions, the problem solving therapy was given at an individual basis. The remaining 7 conditions were all group interventions. The number of sessions included ranged widely (1 to 25 sessions), as well as the time spent per session (20 minutes to 2 hours per session). Sessions of one-hour duration were most frequently reported. In 7 out of 19 studies, a health professional (i.e. GP, oncology social worker, health educator, and trained nurse therapist) guided the p-s intervention. In 5 studies a psychologist or psychiatrist guided the intervention, and in another 3 studies, graduate psychology students guided the intervention. In 2 studies, therapists were a combination of psychologist / psychiatrist and student therapists or trained GPs. In 2 studies, it was not reported what the therapist's background was.

Sixteen studies explicitly reported outcomes with regard to post-treatment assessments, and 13 studies included a follow-up measurement in the study. Follow-up duration ranged from 3 to 12 months. Sixty-seven different outcome measures were extracted from the RCTs and clustered in one of five outcome clusters. In 7 out of 19 studies, problem-solving skills were assessed. Most p-s assessments concerned self-report measures. Eleven studies reported results regarding negative affect. Most outcome measures in this cluster were self-report measures, with the exception of 4 studies using the Hamilton Rating Scale for Depression, which is a clinical interview. Quality of Life or Functional status (QOL) was measured in 11 studies. Again, most measures in the QOL-cluster represent self-report measures, although in some cases, 'important others' like parents or teachers do the report. Some QOL-measures, are clinical interviews. Eight studies report on Domain Specific measures. Measures in this cluster can be a self-report measure, an interview, a task, or an observational measure. The final cluster concerns effects that are indications of general improvement, most of them being primary outcome measures like costs, disability days, contacts with health care providers, frequency and intensity of problems, and clinical effects. Ten studies report on measures in the GI-cluster.

DATA-ANALYSIS

In table 5 (page 20-21), outcomes were evaluated per outcome cluster with regard to post-treatment and follow-up effects. An overall evaluation summarizes cluster-outcomes by counting the majority of signs. At post-treatment, 18 out of 25 comparisons (72%) were favoring the p-s intervention, 2 comparisons

were in favor of the reference treatment, and 5 comparisons failed to show efficacy of the p-s intervention. At follow-up, 18 comparisons were made, of which 13 (72%) showing the p-s intervention was most effective, and 5 failed to find evidence that the p-s intervention was more effective than the reference treatment. Fifteen comparisons described outcomes at both post-treatment and follow-up. In 12 comparisons (80%) positive outcomes of the p-s intervention were maintained at follow-up. One comparison failed to show effects at either post-treatment or follow-up. Finally, in one comparison positive effects faded away at follow-up, and in another, outcomes were favoring the reference treatment at post-treatment, but the p-s intervention at follow-up.

Do outcomes differ as a result of methodological quality of the studies?

When the total score of methodological quality was accounted for, 16 post-treatment comparisons were divided in the high quality group (> 9), and 9 in the low quality group. In the high quality group, all but two comparisons were in favor of the p-s intervention (88%). In the low quality group, on the other hand, only 4 comparisons were in favor (44%), 4 failed to show the p-s intervention was more effective than the reference treatment, and one showed the reference treatment to be most effective (table 5).

In the long term, 16 follow-up comparisons were divided in the high quality group and 2 in the low quality group. In the high quality group 13 out of 16 comparisons (81%) showed the p-s intervention to be most effective, whereas 3 found no effect. In the low quality group, both comparisons failed to find the p-s intervention more effective (table 5). In conclusion, for both post-treatment and follow-up comparisons, studies with a high methodological quality had a higher probability (in %) of finding the p-s intervention to be superior to the reference treatment.

Is the p-s intervention more effective in one cluster of outcome measures than in others?

In table 5 the percentage of comparisons favoring the p-s intervention per outcome cluster at post-treatment and follow-up is displayed. Highest success rates for the p-s intervention were found in the PS-cluster (67 and 100%, respectively), which might be expected as manipulation of p-s skills is supposed to be the mediator of treatment effect in most studies. In contrast, the lowest success rate was found in the NA-cluster (54 and 25%, respectively), which will be discussed later. Finally, in the clusters QOL, DS, and GI, success rates are more or less the same, ranging from 60 to 67% at post-treatment, and from 56 to 63% at follow-up. The levels of success as well as the maintenance of success at follow-up might be important in that most of the primary outcome measures are found in these clusters.

TABLE 5. General outcome review.

OUTCOME CLUSTERS														
STUDY	PS		NA		QOL		DS		GI		OVERALL EVALUATION		# SESSIONS	METHOD.
	POST	/ FU	POST	/ FU	POST	/ FU	POST	/ FU	POST	/ FU	POST	/ FU	GROUP/IND.	QUALITY
a.														
Atha et al., 1992				0						0		0	51	8
Catalan et al., 1991					+	+				+	+	+	41	13
DeVellis et al., 1988			0		0		0		+		0		21	9
Lynch et al., 1997	0		+		+						+		61	9
Mynors Wallis et al., 1997					0	0			-	+	-	+	51	12
Platt et al., 1993							+	0			+	0	10G	9
Blanchard et al., 1996				0		0		0			0	0	61	10
Nezu & Perri, 1989														
11vs.R	+		+						+		+		10G	13
12vs.R	0		+						+		+		10G	13
b.														
Kazdin et al., 1987														
11vs.I2					+	+			+	+	+	+	201	13
11vs.I3					+	+			+	+	+	+	201	13
Kazdin et al., 1989														
11vs.I3			0	0	+	+	+	+	0	0	+	+	251	14
12vs.I3			0	0	+	+	+	+	0	0	+	+	251	14
Kazdin et al., 1992					+	+	+	+			+	+	251	12
McLeavey et al., 1994	+	+	+	0					+	+	+	+	5/61	12
Mynors Wallis et al., 1995														
11vs.I2			0		0						0	0	6/71	16
11vs.I3			+		+						+	+	6/71	16
Graves et al., 1988														
11vs.I2	+	+					0	+			+	+	8G	12
11vs.I3	+	+					+	+			+	+	8G	12
Schwartz et al., 1998			0				0					0	11	10

c.

Areal et al., 1993												
a	+	+	+	+	0	0				+	+	12G 12
b	+	+	+	+	-	-				+	+	12G 12
Clarke & Greenberg, 1986												
a								+		+		2I 7
b								-		-		2I 7
Coche & Flick, 1975												
a	+								+	+		8G 9
b	+								0	0		8G 9
Coche & Douglas, 1977												
a	0		0		+					0		8G 8
b	0		0		0					0		8G 8
% favoring the p-s intervention	67	100	54	25	60	60	67	63	64	56	72	72

Note. a: Problem-solving therapy versus no treatment / treatment as usual (standard treatment) / wait-list control; b: Problem-solving therapy versus other treatment / placebo control; c: Problem-solving therapy versus no treatment and versus other intervention; I1: p-s intervention; I2, I3: other treatment; R: no treatment-reference; PS: problem-solving skills, NA: negative affect, QOL: functional status/quality of life, DS: domain specific outcomes, GI: general improvement; +: p-s intervention most effective; -: p-s intervention least effective; 0: p-s intervention equally effective / no effect; I: individual; G: group; Meth. Quality = Total methodological quality score (range: 0-18).

Are outcomes depending on the number of sessions invested in the p-s intervention?

In order to answer this question, it was decided that the number of sessions necessary to train p-s skills was at least 6. This minimum is based on the fact that problem solving consists of 5 individual skills or stages, and second, problem solving needs practice. In one study (McLeavey et al., 1994), the number of sessions could be increased to 6 if additional training was needed and was therefore included in the group of studies executing the p-s intervention in more than 5 sessions. Seven studies included no more than 5 sessions in the p-s intervention. In this group of studies, 5 comparisons were made at post-treatment of which 2 were favoring the p-s intervention, 2 were favoring the reference treatment, and one failed to show effect. Remarkably, although the number of comparisons is small, comparisons in favor of the reference treatment were only found when the p-s intervention was representing less than 6 sessions. At follow-up, only 4 comparisons were made in the "< 6 sessions" conditions, 2 favoring the p-s intervention and 2 failing to show any effect. In the group representing p-s interventions with more than 5 sessions, 16 out of 20 (75%), and 11 out of 14 comparisons (79%) were favoring the p-s intervention at post-treatment and follow-up, respectively. From this might be concluded that the chance that the p-s intervention is yielding a positive effect is increasing, as the number of sessions invested is at least 6. Moreover, and even more remarkable, p-s interventions with 10 or more sessions, with the exception of one follow-up comparison (Platt et al., 1993), were all yielding positive results (table 5). Thus, according to this selection of studies, the best chance to yield positive results is obtained when the p-s intervention consists of 10 or more sessions. As will be shown next, the number of sessions might have confounded the following comparisons, and its influence is therefore checked.

Is there a difference between individual or group interventions?

As indicated in table 5, at post-treatment, the p-s intervention was delivered in a group in 11 of 25 comparisons, and as an individual treatment in 14 of 25 comparisons. In 73% of the group treatments and 71 % of the individual treatments the p-s intervention showed to be superior to reference treatment. At follow-up, only 5 comparisons were made within the sub-group of group-interventions, of which 4 were favoring the p-s intervention (80%). When individual treatment was regarded, 9 of 13 comparisons (69%) were in favor of the p-s intervention. The positive results of the p-s intervention, might partly be the result of the number of sessions offered, that ranged between 8 and 12 in the group treatment and between 1 and 25 in the individual treatments. The conclusion might be that the p-s intervention is effective, as an individual as well as a group intervention, but the number of sessions offered within the therapy might be crucial to yield therapy success.

Is the choice of the reference treatment determining the direction of the effect?

In order to answer this question, reference treatments were split up in 2 subgroups: a) those representing no treatment, wait-list control, or treatment as usual, and b) those representing another treatment or a placebo control. At post-treatment, the number of comparisons favoring the p-s intervention was 8 out of 11 (73%) when compared to no treatment (a), and 10 out of 14 (71%) when compared to another treatment (b). At follow-up, only 6 comparisons were reported in the subgroup representing comparisons with no treatment (a), of which 3 (50%) showed the p-s intervention to be most effective. The subgroup representing comparisons with another treatment (b) reported 12 comparisons of which the p-s intervention was more favorable in 10 (83%). Somewhat surprising, the sub-group representing comparisons with another treatment (b) was favoring the p-s intervention more often at follow-up, whereas the contrasts between treatment conditions are expected to be greater when the p-s intervention is compared to no treatment. Noteworthy, the average number of sessions invested in the p-s intervention is smaller in subgroup a) than in subgroup b), namely 6.8 and 12.1, respectively. The outcome that the p-s intervention has a high success-rate at follow-up (83%), even when compared to another treatment, might be indicative of the positive effects of the p-s intervention in the long term.

Is the p-s intervention effective in one particular field or population within mental health care or behavioral medicine?

It may be difficult to give an answer to this question because the number of studies selected within each field may be too small. In fact, only two fields were mentioned in more than 2 studies: depression in adults (6 studies) and anti-social behavior / conduct disorder in children (3 studies). With regard to depressive populations 7 out of 9 comparisons (78%) at post-treatment were favoring the p-s intervention, and 5 out of 6 comparisons (83%) at follow-up. Regarding the studies on anti-social behavior and conduct disorder in children, 5 comparisons were available at both post-treatment and follow-up, all showed the p-s intervention to be most effective. However, the positive results in children, might also be the result of the fact that all three studies were from the same research-group, had high methodological quality, and the p-s intervention consisted of 20 to 25 sessions.

DISCUSSION

In answer to the question whether problem-solving therapy is an effective treatment in mental health care and behavioral medicine, 27 randomized clinical trials were entered in a methodological quality assessment. Subsequently, 19 of these trials were included in a data extraction and analysis procedure.

Analysis showed that p-s interventions are effective in the treatment of several populations within mental health care and behavioral medicine, immediately after the intervention as well as at follow-up, 3 to 12 months after the intervention. The methodological quality was high in 59% of the studies. However, when only internal validity criteria were taken into account, 26% of the studies showed high quality. Methodological quality was important with regard to the success of the p-s intervention in that the probability of finding positive results regarding the p-s intervention was higher when the study was of high quality. The p-s intervention turned out to be most effective regarding the improvement of problem-solving skills. Surprisingly, outcomes referring to negative affect were least positive, especially at follow-up. The p-s intervention was more or less equally effective regarding outcomes referring to quality of life, domain specific effects and general improvements. The best chance of yielding positive effects was found when the p-s intervention consisted of more than 5 sessions, with highest success-rates when the p-s intervention consisted of 10 sessions or more. The current study did not find evidence that indicated superiority of either group or individual p-s interventions. Remarkably, studies comparing the p-s intervention to another treatment had better outcomes at follow-up than those comparing to "no treatment conditions". This outcome, however, might have been the result of a higher mean number of sessions in the former. As far as conclusions can be drawn about the effectiveness of the p-s intervention in certain populations, the intervention turned out to be favorable in depressed patients and children with anti-social behavior and conduct disorder. This study did not find evidence that the p-s intervention was contra-indicated in any of the populations included in this review. In only one study, the alternate intervention prevailed above the p-s intervention in volunteers with an intrapersonal conflict. However, the p-s intervention still showed better results than a waiting-list control.

Although study selection was executed in a systematic and objective manner, selection might have been introduced by the search strategy. First, as only MEDLINE and PSYCHLIT databases were searched, studies that were not included in these databases were ignored. Second, the specific search terms (problem solving therapy, -treatment, and -intervention) might have missed those treatments that were indeed executing a problem-solving intervention as intended, but used divergent terms to indicate this.

A second source of selection was deliberately introduced by the definition of inclusion criteria. First, only studies with 10 or more participants per condition were included in the review. Authors shared the opinion that with small Ns, the power of the study might have been impaired more easily what could have resulted in unjust decisions regarding the overall efficacy of the p-s intervention in the overall outcome assessment. Secondly, it had to be obvious from the full paper that the therapy was based on D'Zurilla and Goldfried's theory (1971), and moreover, problem-solving skills were actually trained. Especially the latter criterion was sometimes hard to interpret from the des-

cription in the paper. In these cases poor descriptions could have led to unjust exclusion of papers. Nevertheless, authors found it important to select only those studies where p-s skills were actually trained and served as a mediator of effect. When the theory served merely as a structure of therapy process, therapy might have been effective, but it would have been less plausible to hold increase of p-s skills responsible for it. Publication bias, might have been a third source of bias threatening the conclusions of this review, in that the international scientific press more easily accepted studies showing the p-s intervention to be effective.

Remarkably, only 12 out of 28 comparisons involved in our review evaluated whether p-s skills were actually changed in the p-s intervention. From a theoretical point of view, one would expect that if any, the PS-cluster would be changed as a result of the p-s intervention. So, when not assessed, it remains questionable whether effects might be attributable to the increase of p-s skills. On the other hand, when p-s skills were assessed but outcomes were not favoring the p-s intervention, other factors might have been responsible for the effects, whatever the direction of outcomes. Noteworthy in this respect might be the validity of p-s measures. As Tisdelle et al. (1986) and Larson (1990) discussed in their reviews, most p-s measures consist of self-report, which is by no means an objective method of assessing outcomes regarding actual p-s skills. One objection might be that subjects from the p-s intervention would recognize statements in the self-report measure more easily, because comparable statements were repeatedly stressed during therapy. In that case, we would call the assessment of p-s skills a process-measure (manipulation check), rather than a primary effect. Second, self-report measures may lack ecological validity as they do not measure real-life problem solving (Larson, 1990). Validation of self-report by more objective measures (e.g. observations) might be called for in future research.

Remarkably, effect-rates regarding NA-measures were lower than those in other clusters, especially at follow-up where only 25% of outcomes favored the p-s intervention. As the p-s intervention is claimed to be most effective regarding stress-related disorders as depression and generalized anxiety (D' Zurilla, 1988), one would expect success-rates in the NA-cluster to be at least comparable to other clusters. Floor-effects may be one explanation, in that initial levels of negative affect were already low at pre-treatment, so that it became difficult to make important changes. Indeed, in studies where depressive patients were included ($N = 9$), 6 out of 7 comparisons were in favor of the p-s intervention at post-treatment. Unfortunately, only 1 of the 'depression-studies' included NA at follow-up. It would be worthwhile to include follow-up assessments in future research, to investigate whether the positive effects of the p-s intervention in depressive populations are maintained in the long term.

The number of sessions invested in the p-s intervention turned out to be crucial regarding the success of the intervention. From this it may be deduced, that skills can not be trained within only a few sessions, but need rehearsal

and real-life practice. As some interventions did not include more than 1, 2, 4 or 5 sessions to introduce p-s skills, it is hard to believe that these interventions not only educated basic skills, but paid attention to practice and generalization of these skills as well. This analysis also showed almost perfect success when the number of sessions was 10 or more. This might be another indication that training and integration of skills in everyday life needs time (in- and outside the therapy setting) and practice.

Although the majority of studies succeeded to show efficacy of the p-s intervention, 28% did not. The reasons for this failure might be several. First of all, treatment contents, performance and number of sessions might be crucial. Especially in those studies p-s skills were assessed and not successfully changed, as compared to a reference treatment, there may be reason to doubt the contents of the p-s intervention. Moreover, as treatment integrity was rarely checked, doubt may be thrown on how the treatment was performed. In those studies treatment integrity was not checked it becomes even more important that treatment manuals could be obtained, and therapists were trained and supervised. Failure to show effect of the p-s intervention can also be due to the choice of the reference treatment. In some cases the reference treatment already proved effective in the treatment of a certain problem, and ~~was therefore chosen as standard of golden standard~~ in those cases, it might be more realistic to expect the p-s intervention to be as effective as the reference treatment, or to be partially effective regarding some specific outcome measures. Finally, there might have been some individual reasons why some studies did not succeed to prove the p-s intervention to be effective. In example, in a study of Hussian and Lawrence (1981), depressed elderly patients who participated in the p-s intervention, were for methodological reasons refrained from other daily activities held within the facility (like craft classes, college class, or artistic activities) during the study. In depressed patients, this refusal might have felt as punishment rather than reinforcement. Although the Hussian-study was not included in the final outcome review, it can not be ruled out that other studies sought comparable solutions for methodological problems and thereby put the experimental condition at a disadvantage.

The current study found methodological quality to be important with regard to the success of the p-s intervention. Poor methodological quality seems to reduce the chance of finding the experimental intervention successful and must therefore be considered as a serious source of bias. On the other hand, high methodological quality does not always guarantee that the p-s intervention was thoroughly performed.

It needs to be kept in mind that the current systematic review did not include studies conducted later than January 1999. Therefore, several studies that have been described since, e.g. problem solving therapy in oncology patients (Nezu et al., 1998) and non-specific low back pain (Van den Hout et al., Under review), were not involved in data-analysis. It would be preferable to have frequent follow-ups in order to include new RCTs and to suggest new

directions regarding evaluation of the p-s intervention. According to D'Zurilla (1988) the most appropriate case for p-s therapy is one in which negative stress effects, such as anxiety, depression, or psycho-physiological symptoms (e.g. pain), are associated with deficits in problem solving ability and/or performance. As the current review selected several studies regarding the treatment of depression, however, less with regard to anxiety and psychophysiology, these areas may need further addressing in future research. Another limitation of the current study may be the selection of RCTs, leaving out other valid methods of effect-evaluation. Especially when N was small, other methods might have been more appropriate to test treatment efficacy (e.g. single case experimental designs and multiple baseline designs; Heppner et al., 1984; Morley, 1989; Morley & Adams, 1989). Notably, those studies that were excluded because of small sample-size, all had low internal validity.

Some interesting findings were not discussed in the results section and are worthwhile remarking here. First, a dismantling study by Nezu and Perri (1989) gave a lot of insight into the working mechanisms of the experimental condition. An intervention in which both problem orientation and problem solving skills were discussed and trained, had more effective results than the intervention that left out the problem-orientation phase. Unfortunately, only one dismantling study was found in this data-search (Nezu and Perri, 1989). Moreover, from the description of the intervention it was not always clear whether problem orientation was actually included in the therapy or not. The latter prevented us from addressing the more specific question whether problem orientation would be a necessary part of problem-solving therapy. Future research may address this specific question more thoroughly, evaluating the contents of the intervention in particular, as well as how it affects therapy outcomes.

Finally, some studies yielded results that call for further exploration in future research. First, Kazdin et al. (1992), found evidence that at one-year follow-up, the p-s intervention alone was as effective as the p-s intervention in combination with a parent management training, and more effective than the parent management training alone. It appeared to be useful to investigate the optimal combination of interventions, which, as in this case, may not always be the most 'complete' one. Another characteristic of the Kazdin-study (Kazdin et al., 1992) is the involvement of parents in the p-s training of their children. Involving important others in the process of solving problems of everyday life, is an addition that deserves further attention of therapists and researchers. Second, the study by Mynors Wallis and colleagues (1997) investigated cost-effectiveness of a p-s intervention in primary care. Their outcomes are surprising because whereas treatment as usual turned out to be more cost-effective at post-treatment, the p-s intervention was more favorable a half year after the intervention. In other words, the investments in the short term are recovered within 6 months. Costs are rarely used as an outcome in studies included in our review, but could add interesting information to future effect-research of

the p-s intervention. Latest developments show a shift from the treatment of patients themselves towards applications of p-s principles in caregivers and (mental) health professionals who are now viewed as the actual problem solvers (Loscalzo & Bucher, 1999; Nezu & Nezu, 1993). In particular, the p-s theory might serve as an effective model of active clinical decision making regarding all kinds of clinical problems a health professional is professionally confronted with, such as differential diagnosis, selection of therapeutic interventions, and unanticipated problems occurring during the therapy process. Studies regarding the efficacy of this kind of interventions are quite unexplored territory, but might show very powerful especially when problem solving becomes a more integrated process between patient and clinician. Introduction of problem solving at multiple levels of care may be a most interesting issue of future research in this field.

CHAPTER 3

Does failure hurt?
The effects of failure feedback
on pain report, pain tolerance, and pain avoidance.

This chapter was published as:

Van den Hout, J.H.C., Vlaeyen, J.W.S., Peters, M.L., Engelhard, I.M., Van den Hout, M.A. (2000). Does failure hurt? The effects of failure feedback on pain report, pain tolerance, and pain avoidance. *European Journal of Pain*, 4 (4), 335-346.

ABSTRACT

In this study an experiment was conducted to examine whether failure experiences have an effect on pain report, pain tolerance and pain avoidance. Furthermore, it was investigated if negative affectivity (NA) affected the impact of failure feedback on pain report, either as a mediator in the case of negative state affect or as a moderator when NA as a personality trait was considered. Fiftyfour healthy female volunteers were included and randomly assigned to one of three conditions: 1) failure feedback; 2) success feedback; 3) neutral control task. After the manipulation, subjects were given a cold pressor task in order to obtain pain measures. Regarding the effects of failure feedback on pain report, it was found that, in comparison with success feedback, failure feedback led to increased pain report. With regard to pain tolerance, pain was tolerated longer when preceded by success feedback than when preceded by failure feedback. Differences between failure and control conditions did not reach significance. With regard to pain avoidance, no differences between the conditions were found. The hypothesized mediating role of negative state affect was not found. Though in the hypothesized direction, no significant effect was found for NA-trait to moderate the influence of failure on pain. The discussion focuses on a number of research questions that remain to be answered, and to the clinical relevance of the effects of failure and success experiences on pain report and pain tolerance.

INTRODUCTION

Common cerebral processes, involved in both emotion and pain, suggest that the experience of pain is always modulated to some extent by coexisting emotions (Melzack and Wall, 1965; Robinson and Riley, 1999). Because of this, it might not always be clear how both constructs relate to each other. Some studies support the hypothesis that negative emotion is a frequent psychological reaction to pain, i.e. emotional distress as a consequence of both acute and chronic pain (Craig, 1994; Gamsa, 1990). This study focuses on mechanisms by which emotions may instigate or increase pain. Although there is no empirical support for the proposition that emotional distress really causes pain (Gamsa, 1994), there is sufficient evidence that pain can be moderated by distress. Sternbach (1986) observed that stressful events, including major life events or daily life 'hassles', are strongly associated with increased pain. It can also be hypothesized that the influence of a stressor is mediated by a negative emotional state, such as depression or anxiety. Depressive states have been shown to increase clinical pain in humans (Brown, 1990; Sullivan and D'Eon, 1990; Wade et al., 1990), and to decrease tolerance for experimentally induced pain (Zelman et al., 1991; Romano and Turner, 1985). As for anxiety states, there is no clear empirical basis for the proposition that anxiety increases pain. A review of studies by Arntz et al. (1991) indicated that anxiety might enhance or relieve pain or not affect it at all. Their review concluded that anxiety has no unequivocal effect on pain, but that the anxiety-pain relationship is mediated by attention. With regard to positive mood states, there is some evidence which suggests that positive emotional states, induced by imagery or selfcontrolled distraction, enhance pain tolerance (Horan and Dellinger, 1974; Chaves and Barber, 1974; Zillmann et al., 1996). Evidently, the effect of emotional mood states on pain seems to depend on different dimensions, such as attention paid to the stressor, mood quality, and mood intensity. Moreover, ecological validity of the kind of mood inductions used in experimental studies might be of influence.

Failure experiences can be an important source of emotional distress and therefore directly or indirectly influence pain. Levine et al. (1993) examined whether pain experience can be influenced by a mental stressor and by the perceived failure or success in dealing with this non-painrelated stressor. Two dimensions of the experimental situation were manipulated: task difficulty and feedback quality (failure or success). Undergraduate students were randomly assigned to one of five conditions: an easy reading comprehension task with either success or failure feedback; a difficult reading comprehension task with either success or failure feedback; or a control condition. After false success or failure feedback was given in the comprehension task conditions, a cold pressor test was conducted. Subjects were asked to place their non-dominant hand in a bucket of circulating ice water, 0-2°C, up to their wrists for a maximum of 2 minutes. During immersion in the cold pressor, subjects

were asked to score their pain intensity every 30 seconds. In this experiment, feedback turned out to be more important than task difficulty in the explanation of pain report. Failure feedback led to higher pain reports than a non-feedback control condition. Success feedback, on the other hand, did not lead to decrease of pain report as compared to a control condition (Levine et al., 1993).

However, a number of methodological aspects may weaken the conclusion of this study. First, the impact of the manipulation was not checked, so it remains unclear whether the manipulation succeeded. In particular, the absence of effects of success feedback may have resulted from the manipulation not being strong enough. Second, it was not clear from this report whether the same experimenter executed both comprehension task/control task and the cold pressor task. If so, the experimenter was not blind for the condition. A third note concerns the pain measure. In order to prevent ceiling effects, subjects in the Levine study were told that they could choose numbers higher than the maximum of the scale (32). Because no word-anchors were given above 32, it becomes unclear what pain subjects were experiencing if their pain was rated worse than excruciating (>32). A fourth remark could be made about the way ratings were coded after withdrawal from the cold pressor: all ratings beyond tolerance time were assigned a rating of 64. When defined like this, the pain intensity measure is likely to be confounded with pain tolerance.

Besides these methodological aspects, there are three theoretical issues that are of interest: the comparison between failure and success condition, the role of Negative Affectivity (NA), and the effect of perceived failure on behavioral aspects of pain. Even though the success condition did not lead to a significantly lower pain report than the control condition, still the largest contrast may be expected between failure and success. It therefore might be interesting to compare these conditions directly.

In their discussion, Levine et al. (1993) proposed that NA might at least partly have accounted for the effect of failure on pain report. Watson and Clark (1984) originally described the NA construct as a mood dispositional dimension or trait. High-NA individuals tend to be distressed and have a negative view of themselves. Moreover, high-NA individuals are more prone to experience significant levels of distress and dissatisfaction at all times and in any given situation. High-NA individuals are more introspective and differentially dwell on their failures and shortcomings (Watson and Pennebaker, 1989). Thus, when failure was perceived in the Levine et al. study, an NA disposition (NA-trait) could have moderated the effect of failure on pain report, because high-NA individuals were more sensitive to the failure feedback. According to Watson & Pennebaker (1989) NA can also be an induced state, and it is quite possible that failure feedback induced such a negative emotional mood state (NA-state). This emotional state might have mediated between failure feedback (the stressor) and increased pain reports, or at least may have been partly responsible for increased pain reports in the failure condition. Hence, two diffe-

rent roles of NA can be hypothesized: as a moderator when referring to an NA-trait, or as a mediator when referring to an induced mood state (Baron and Kenny, 1986).

It would be worthwhile investigating whether, besides subjective pain report, behavioral aspects of pain (pain tolerance and avoidance) are affected by antecedent failure experiences as well. Moreover, measurement of pain behavior would provide a more objective outcome measure. In addition to the Levine replication, we therefore add the following research questions: are painrelated activities, when preceded by negative feedback, i.e. failure, less well tolerated or more avoided?

In summary, the aim of this study was to determine the effects of failure and success experiences on pain report, pain tolerance and pain avoidance. Furthermore, with regard to NA, the following question was examined: does NA influence the effect of failure on pain report, either as a mediator (NA-state) or as a moderator (NA-trait)? There were a few methodological modifications in comparison with the original study of Levine et al.: a) a manipulation check was included; b) intelligence test and cold pressor task were performed by two different experimenters, with the second experimenter being blind to the condition; c) to prevent subjects from becoming demotivated and reluctant during the experiment in which they just heard they failed an intelligence test, manipulation and cold pressor task were presented as two independent studies; d) the maximum score on the pain intensity scale was kept at 32.

METHODS

SUBJECTS.

Sixtythree first-year undergraduate students were recruited by means of posters at different faculties of the Maastricht University. The advertisement was designed such that subjects were led to believe that they were being invited for two independent experiments. The first would be a pencil and paper task. The second would be a laboratory experiment investigating the influence of pain and emotion on skin conductance levels (SCL). In case the subject asked how pain would be inflicted, he or she was told that this was by immersion of the hand in cold water. The candidates were told that they could decide themselves when to stop the test. Because only 6 male students were recruited and because it was not feasible to recruit equal numbers of male and female subjects in both conditions within time, we decided to include only female students. Fiftyfour of the 57 female students met the inclusion criteria (blood pressure lower than 140/90mmHg; no medical contraindications, no analgesics used within the last 12 hours). Two subjects were excluded because of medical reasons (blood pressure above 140/90; S did not feel well) and one subject was excluded because her housemate disclosed the rationale of the experiment to her. The remaining 54 subjects had a mean age of 19.1 (range: 18-22 years).

MEASURES

Manipulation check

The manipulation check consisted of a 0 to 100-percentile expectancy-scale. The subject was asked to indicate the expected percentile score on the intelligence test when she compared herself to a pre-university reference group. The scale was completed twice: before starting the intelligence test, and immediately after the manipulated feedback was given. On the second occasion, the subject was asked to give expectancy-scores with regard to completion of a similar test in the near future. Besides this check, two additional Visual Analog Scales (0–100) were used just before final debriefing. On these scale, the subject could indicate how disconcerted she was about the feedback given, and how satisfied she was about the test results, respectively. Both questions were checked to identify those subjects that were not disconcerted at all despite unsatisfying test results. In the case unsatisfying results did not seem to disconcert the subject, the experimenter questioned the subject in order to find out the reason. If the subject would declare that she was not disconcerted because she did not believe the test-results, the manipulation failed and the subject would be deleted from the analyses.

Pain intensity and unpleasantness.

During the cold pressor task, pain intensity and pain unpleasantness were both rated by means of two 32-point rating scales (Padawer and Levine, 1992). Pain words used in the original scales were translated and checked with the Dutch language. The scales were scored every 30 seconds with a pencil held in subject's non-immersed hand. Scales were presented in a booklet each page containing one pain intensity and one pain unpleasantness scale. Subjects were asked to turn the page immediately after scoring such that new scales became visible.

Pain tolerance

Pain tolerance was defined as the time elapsed (in seconds) between immersion and withdrawal from the cold pressor. A priori, it was decided that the maximum tolerance time was set at 5 minutes.

Pain avoidance

Pain avoidance was operationalized as the time elapsed (in seconds) between last instructions after withdrawal from the cold pressor and second immersion into the cold pressor.

Negative affectivity-trait (NA-trait)

NA-trait was measured with the Dutch translation of the Negative Emotionality subscale (NEM) of the Multidimensional Personality Questionnaire (Tellegen et al., 1989; Stegen, 1998). The NEM was taken before the experimental manipulation.

Negative affectivity-state (NA-state)

NA-state was measured by the Dutch translation of the shortened version of the Profile Of Mood States (POMS, Wald et al., 1990). The POMS has been shown to be a valid and responsive measure of negative affect. McNair (1971) reported effects on the POMS in the hypothesized direction with regard to short-term psychotherapies, drug trials and mood-inducing conditions. Because of high correlations among 3 subscales of the POMS (tension, anger and depression), they were joined into one composite measure. This composite measure was further used in this study. NA-state was measured twice, first before the manipulation task and before the measurement of NA-trait, and second just before the cold pressor task, after the manipulation.

PROCEDURE.

Failure manipulation

Subjects were randomly assigned to 1 of 3 conditions: a) an intelligence test with failure feedback; b) an intelligence test with success feedback; or c) an 'evaluation of art' control task. Experimenter 1 conducted the manipulation, and was therefore informed about the randomization scheme. After blood pressure was measured and medical condition was checked, two informed consent forms were signed, one for each experiment. After completion of the POMS and the NEM scales, instructions about the test were provided.

In the intelligence test conditions, the subject was told that the test was being conducted as a pilot for national implementation. The intelligence test was said to give a valid indication of her scientific qualities and would be applied as an entrance exam for university. After a brief introduction, the expectancy-scale (manipulation check) was administered, followed by the intelligence test. It consisted of 24 selected items of the 'Drenth Series' (Drenth et al., 1965, 1969, and 1970), an intelligence test developed to differentiate among subjects with elevated intellectual levels. To maximize failure experiences, 6 of 24 items were made unsolvable in the failure condition only. In addition, subjects in this condition were given only 25 minutes to complete the test, whereas in the success condition, subjects had 35 minutes to complete the test. In order to keep the second experimenter blind for the condition, a filler task of 10 minutes was given in the failure condition to make experiment duration comparable. When the subject finished the test, the experimenter immediately entered data into a computer program, after which the manipulated score appeared on the screen. In the failure condition the experimenter reported a fictitious and low percentile score (range: 21st to 28th percentile). It was made clear to the student that the score was very low in comparison with other pre-university students, and the experimenter gave several comments stressing the personal failure. In the success condition, the subject was given a fictitious high percentile score (range: 91st to 98th percentile), together with a couple of reinforcing comments, stressing the extremely favorable performance in comparison with other pre-university students.

In both conditions the experimenter checked the scores another time to convince the student that there was no mistake with data input and processing. Next, a computer printout was given to the student stating "Compared to the norms of a reference group of pre-university students....", followed by the pre-determined percentile score. The score forms as well as the forms used for the manipulation check were all given some official coating to make the feedback as credible as possible. Immediately after the feedback, the student was asked to complete the expectancy-scale once more (see manipulation check).

In the control condition the student was requested to give her opinion about abstract pictures, to tell something about her personal taste of art, and to select some favorite paintings out of several artbooks. It was stressed that there was no evaluating purpose in the test, and that it was one's personal evaluation that mattered. After completion of the manipulation (intelligence test or control task), the first experimenter thanked the student for her participation and handed her directly over to the second (blind) experimenter who waited for her in the laboratory.

Cold pressor task.

After completing the POMS for the second time, the subject was told that this experiment aimed at investigating the influence of pain and mood on skin conductance levels. Subsequently, and after the attachment of sham SCL-electrodes, a cold pressor task was conducted. The subject was asked to place the non-dominant hand up to the wrist in the cold pressor apparatus. This apparatus consisted of a container filled with ice water, 0-2°C, circulated by a fish-tank aerator. The experimenter scooped out the ice to prevent contact with the subject's skin. Water temperature was recorded for each subject before the cold pressor task, but could not be read by the subject. The subject was instructed to tolerate the cold pressor as long as possible. During immersion, the subject was given an audible signal every 30 seconds to rate pain intensity and unpleasantness on a 32-point rating scale. When the hand was withdrawn from the cold pressor, the tolerance time was clocked and an audiotape instructed the subject to complete both scales once more for the pain she felt just before she withdrew her hand from the cold pressor. After a few seconds, new audiotaped instructions were given to restart the same cold pressor task once more when the subject felt ready for it. As soon as the subject reimmersed her hand into the cold pressor, the avoidance time (the time the subject delayed reimmersing her hand) was clocked. The subject was now told she could withdraw from the coldpressor and that the task was finished. The cold pressor task was executed by means of the Micro Experimental Laboratory Professional software package (Schneider, 1988).

Immediately after the cold pressor task, the experimenter asked what kind of test was carried out in the preceding experiment. When this was an intelligence test, the experimenter requested the subject to complete the two VAS-

scales designed as two additional manipulation check measures. Finally, the subject was debriefed.

HYPOTHESES.

The following hypotheses were put forward: 1) In the failure condition, subjects report higher pain intensity ratings than subjects in the control condition (replication Levine et al., 1993), and than subjects in the success condition; 2) Subjects in the failure condition tolerate the cold pressor for a shorter period of time than subjects in the control condition, and than subjects in the success condition; 3) Subjects in the failure condition avoid immersing their hand in the cold pressor a second time for a longer period of time than subjects in the control condition, and than subjects in the success condition; 4) NA-state mediates between failure feedback and pain report; 5) NA-trait moderates the effect of failure on reported pain intensity: high NA-subjects are more prone than low NA-subjects to react with high pain reports in a failure condition.

STATISTICAL ANALYSES.

To test Hypothesis 1, repeated measures MANOVA on the successive pain ratings was conducted. As in the original study, only pain ratings at a 60-, 90-, and 120-second interval (measurement) were taken into analyses. The highest rating on the scale (32) was assigned to ratings when the subject's hand was withdrawn from the cold pressor. Though the procedures were identical to those in the Levine et al. study, we used a score of 32 instead of 64, for reasons named in the introduction. As differences were expected between the failure versus control and the failure versus success condition, but not between success versus control condition, posthoc planned comparisons between all three conditions were performed. An ANOVA and posthoc planned one-tailed *t*-tests for independent observations were conducted for testing hypotheses 2 and 3. Because it was hypothesized that tolerance time was positively correlated with avoidance time, tolerance time was taken into analyses as a covariate, when analyzing pain avoidance.

To construct a hypothetical model for the role of NA-trait and NA-state in reported pain intensity, multiple linear regression analyses were conducted. Statistical analyses were carried out according to Baron and Kenny (1986). To test for mediation, three regression equations were tested (Baron and Kenny, 1986): 1) regressing the mediator (NA-state) on the independent variable (condition); 2) regressing the dependent variable (pain report) on the independent variable (condition); 3) regressing the dependent variable (pain report) on both the independent variable (condition) and the mediator (NA-state). To establish mediation, the effect of the independent variable on the dependent variable should be less in the third regression equation than in the second (Baron and Kenny, 1986). Besides condition, NA-trait and NA-state (pre-manipulation) were added to the first regression equation, and NA-trait was added to the second and third regression equations.

FIGURE 1. The hypothetical model.

To test hypothesis 5, the product of condition and NA-trait was added to the regression equation. The moderator effect is indicated by the significant effect of the interaction while predictor and moderator variables are controlled (Baron and Kenny, 1986). With regard to hypotheses 4 and 5, pain reports on the 120-second time interval were taken into analysis and the last reported score was assigned when the subject had already withdrawn her hand from the cold pressor. The hypothetical model is presented in figure 1.

RESULTS

MANIPULATION CHECK.

As expected, the pre-post expectancy-scores increased by 27.39 percentiles in the success condition, whereas a decrease of 27.44 was found in the failure condition. This difference was significant ($t(34)=-14.19, p<0.001$). Furthermore, Ss in the failure condition were significantly more disconcerted about the intelligence test feedback than Ss in the success condition ($t(33)=10.59, p<0.001$), whereas Ss in the success condition were significantly more satisfied about their results than Ss with failure feedback ($t(33)=-17.02, p<0.001$). Only one subject reported not to be disconcerted despite unsatisfying test-results. However, she did not doubt the genuineness of the test-score, but attributed the bad score to a temporary state (a hangover) instead. It was therefore decided to keep the data of this subject in the analyses. Finally, it was checked if failure induced negative affect. Change scores pointed out that failure Ss increased NA-state, whereas success Ss decreased NA-state ($t(19.36)=3.38, p<0.003$). In sum, experimental manipulation appeared to be successful.

FIGURE 2. Pain intensity ratings over four 30-second time intervals, for each condition.

PAIN REPORT.

Reported pain intensity ratings over 4 time-intervals are shown in figure 2. The pain unpleasantness measure significantly correlated with the pain intensity measure ($r>0.90$; $p<0.001$, for each of the 60-, 90- and 120-second time-intervals), and results were identical for both measures. Therefore, as in the Levine et al. study, only the results of the pain intensity measure are presented.

Because pain report was not normally distributed, scores were reflected (maximum score + 1) minus (observed pain intensity score) and transformed (square root). Means and standard deviations after transformation can be found in Table 1. Due to the reflection of data before transformation, highest scores become lowest and lowest become highest. Repeated measures MANOVA (measurement X feedback condition) was conducted with 'Measurement' as the within subjects factor with three levels: pain report at 60, 90, and 120 seconds. There was a significant effect for measurement ($F(2, 102)=16.16$; $p=0.000$), which means that pain report increased as a function of immersion time. Neither condition effect ($F(2,51)=1.68$; $p=0.196$), nor measurement X condition interaction ($F(4,102)=0.63$; $p=0.645$) reached significance. As planned, post hoc pairwise comparisons were done. Unlike the original study, only a trend was found with regard to comparison of failure and control conditions (estimated difference=-0.64; s.e.=0.462; one-tailed $p=0.086$). As in the original study, there was no significant difference between success and control conditions. The additional comparison of failure and success conditions resulted in a significant effect (estimated difference=-0.80; s.e.=0.462; one-tailed $p=0.045$), which means that during a 60-, 90-, and 120-second interval, subjects in the failure condition reported higher pain than subjects in a success condition.

TABLE I. Means and standard deviations of pain report (after transformation), pain tolerance, and pain avoidance by condition. N=18 for each condition. F=Failure condition; C=Control condition; S=Success condition.

Dependent variable Condition	Mean	Std. Deviation
pain report (VAS)		
60-sec. F	2.75	1.34
C	3.24	1.51
S	3.39	1.34
90-sec. F	2.24	1.32
C	2.88	1.50
S	2.99	1.47
120-sec. F	1.95	1.37
C	2.74	1.66
S	2.96	1.55
Pain tolerance (sec.)		
F	120.84	88.89
C	160.32	103.94
S	178.49	112.05
Pain avoidance (sec.)		
F	51.12	59.62
C	35.22	32.57
S	45.59	46.89

PAIN TOLERANCE.

Means and standard deviations of pain tolerance (in *s*) are displayed in Table I. Subjects in the failure condition, on average, withdrew their hand from the cold pressor almost one minute earlier than subjects in the success condition. ANOVA showed no statistically significant differences between the three conditions ($F(2,53)=1.501$; $p=0.233$). Post hoc planned pairwise comparisons, however, reached significance for the difference between failure and success conditions ($t(34)=-1.71$, one-tailed $p=0.048$). The difference between failure and control conditions was not significant ($t(34)=-1.22$, one-tailed $p=0.115$). Figure 3 shows the number of subjects (per condition) who withstood the cold pressor task. The number of subjects continuing to do so is given for every 30 seconds.

PAIN AVOIDANCE.

Table I gives means and standard deviations of pain avoidance (in seconds). No differences were found between the conditions ($F(2,53)=0.248$; $p=0.781$). Pairwise comparisons did not reach significance either (Failure vs. Success:

Figure 3. Number of subjects continuing the cold pressor task at ten 30-second intervals, for each condition.

$t(34)=0.161$, one-tailed $p=0.437$; Failure vs. Control: $t(34)=0.686$, one-tailed $p=0.249$). Analyses were performed after data transformation (square root). The outcomes of the ANCOVA were comparable to those acquired when tolerance time was not taken into analysis as a covariate.

THE MEDIATING ROLE OF NA-STATE.

To test the hypothesized mediating role of NA-state, three multiple regression analyses were performed. One subject was left out of the analyses because the score on negative affect (post-manipulation) was identified as an outlier. Final results of the regression analyses are presented in table II. To simplify interpretation of reflected and transformed data, the original direction of the relationship as expressed in the Standardized Beta was restored. The Variance Inflation Factors were small enough (range 1.01 to 1.74), which suggests that there is no problem of collinearity. First, predictors of NA-state (post-manipulation; mediator) were analyzed. A model was entered with NA-trait, NA-state (pre-manipulation) and condition (failure vs. control/success) as predictors of NA-state (post-manipulation; Table II, model 1). The model predicted 58% of the variance and feedback condition was one of the predictors. Secondly, the predictive value of feedback condition on pain report was tested (Table II, model 2). Results showed that this was not the case, and therefore the mediation hypothesis was rejected. Nonetheless, the third analysis, in which the mediator was added to the regression equation, is reported because, in contrast to the second equation, condition significantly predicted pain report in this equation (Table II, model 3a). Although accounting for only 13% of the variance, the third equation, with NA-trait, condition (failure vs. control/success) and NA-state (post-manipulation) as predictors, was significant. Surprisingly, in the

TABLE II. Regression analyses of NA-state (post-manipulation) and pain report. Analyses 1 to 3a refer to the test of the mediator effect of NA-state; Analysis 3b refers to the test of the moderator effect of NA-trait.
* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Dependent variable	<i>N</i>	Adj. R^2	Explanatory variables	Standardized Bèta
1) NA-state (post-manipulation)	53	0.581***	1) •NA-state (pre-manipulation) •NA-trait •Condition (failure vs. success and control)	0.518*** 0.304** 0.438***
2) Pain report	53	-0.009	2) •NA-trait •Condition (failure vs. success and control)	0.097 0.141
3a) Pain report	53	0.132*	3a) •NA-state (post-manipulation) •NA-trait •Condition (failure vs. success and control)	-0.492** 0.331* 0.325*
3b) Pain report	53	0.158*	3b) •NA-state (post-manipulation) •NA-trait •Condition (failure vs. success and control) •Interaction: NA-trait x condition	-0.571** 0.496** 0.364* 0.248

third equation, NA-state contributed to the predictive model of pain report, but in the opposite direction: The higher NA-state (post-manipulation), the lower the pain report. Rather than mediating the effect of failure on pain report, NA-state seemed to confound the effect. Only by controlling for it, the predictive model, in which condition was one of the predictors, became significant.

THE MODERATING ROLE OF NA-TRAIT.

With regard to the moderator effect, the interaction term (NA-trait X condition), was added to the third model (Table II, model 3b). Although the Bèta coefficient appears to be in the expected direction, the interaction term did not reach significance. Hence, it cannot be concluded that there was a moderator effect of NA on pain report.

DISCUSSION

After successful manipulation of success and failure feedback, this study showed that, compared to success, failure leads to increased pain report and decreased pain tolerance during a cold pressor test. However, differences between failure and control conditions, which reached significance in the Levine et al. study, did not reach significance in the current study. Although findings were in line with the original study, they only partly supported the "failure hurts" hypothesis: In a female student population, pain report was increased only when failure was compared to success feedback, not when compared to a control condition. Furthermore, in contrast to what Levine et al. (1993) proposed, we did not find evidence for the mediating role of NA-state in the failure-pain report association. Neither did we find support for the moderator effect of NA-trait on the influence of failure experiences on pain report, although results point in the hypothesized direction.

One of the reasons for the weaker support might be the methodology used. First, in the Levine-study subjects were told that they could choose ratings higher than the maximum of 32 to prevent ceiling effects. This could have increased differences, especially when those subjects who already scored high on the rating scale used this option. A second methodological difference is the way missing ratings were recoded. In the original study, after withdrawal from the cold pressor, ratings were recoded with the value of 64, whereas in our study the maximum value on the pain intensity scale was used, that is 32. Recoding missing values (after withdrawal from the cold pressor) as 64 may have increased the mean pain report in the condition where subjects withdrew their hand early and tolerated less. Unfortunately, Levine et al. (1993) did not report for which conditions the 25 missing ratings were recoded, but based on our tolerance data, it is likely that missings occurred mostly in the failure condition. Another reason why significance levels differ between the two studies might be the impact of the manipulation. Although the manipulation check in our study proved that the feedback manipulation was successful, the manipulation in the Levine et al. study could have been more powerful because of non-specific factors such as impact of the experimenter or the verbal feedback given with the computer output. In general, it might be difficult to find an experimental false feedback manipulation that is strong enough to impress subjects, and being ethically acceptable as well.

Another possibility is that the art control task could have led to unnoticed feelings of frustration and 'failure' if subjects were not interested in or ignorant of art. This could have hampered the intended contrast between both conditions. Not only pain report was influenced by failure feedback, we also found that pain tolerance was influenced. This would suggest that, as compared to success conditions, pain-eliciting activities are more likely to be quitted earlier as a result of failure in non-pain related domains. So far, pain tolerance has been reported to be influenced by psychological variables, such as pain-related

fear (Vlaeyen et al., 1995b), dissociative tendencies (Orbach, 1994), pain-incompatible imagery (Neumann et al., 1997), distraction, and even humor and laughter (Weisenberg et al., 1995). Furthermore, Keefe et al. (1997) found that in arthritis pain, self-efficacy relates positively to pain tolerance in experimental pain. In our study there are some indications that failure experience might be an additional determinant of pain tolerance. Pain tolerance has clinical importance in that reduced tolerance may generalize to daily activities of which the reduction leads to increased disability levels in chronic pain patients. The role of non-painrelated failure experiences in the development and maintenance of chronic pain deserves further examination.

With regard to pain avoidance, the expected effect was not found. This is surprising, because from literature on learning foundations (Kanfer & Philips, 1970) one would expect that processes (i.e. aversive stimulation) influencing avoidance, would be comparable to those influencing tolerance. On the other hand, in a study by Cipher and Fernandez (1997) it was found that expected pain tolerance significantly predicted actual tolerance, whereas expected danger significantly predicted pain avoidance. These findings suggest that, pain avoidance and pain tolerance might not have the same predictors. As we decided to operationalize pain avoidance by the time subjects delayed the start of a second cold pressor task, *this measure is likely to be influenced by the preceding cold pressor task*, what might have biased results on pain avoidance. Tolerance time, pain on stopping the first task, and anticipation of pain, are likely sources of variance with regard to the moment subjects decide to immerse their hands into the cold water a second time. Furthermore, based on the previous task, subjects may have decided to start the next one soon, but to withdraw their hand quickly, whereas others, avoid longer as to prepare for another long immersion time. It therefore seems more prudent to examine the effects of failure feedback on tolerance and avoidance in separate experimental set-ups.

The mediating role of NA-state was not confirmed. In contrast, the regression analysis pointed in the opposite direction, which raises new questions. Analysis of the separate regressions within each condition reveals that the negative relation between NA-state and pain report only holds for the control condition. In both feedback conditions, no relation was found between NA-state and pain report. There are no clear explanations for these findings. It still remains an interesting question which factors could possibly mediate between failure feedback and pain report. We suggest that attributional processes might be one of the alternative possibilities: Whether failure is attributed internally ("It's my fault") or externally ("I'm not responsible for this failure"), stable and global ("always and on any occasion") or unstable and specific ("just today on this particular task"), is likely to make the difference. Mikulincer (1986), for example, found that performance in a dissimilar situation was impaired when failure (unsolvable problems) was attributed to global and stable causes, and when it was attributed to global and internal causes. In a later

study, it was found that exposure to unsolvable problems worsened subsequent performance only when subjects attributed failure to stable causes (Mikulincer, 1988). Likewise, Klein et al. (1976) found that only failure that leads to decreased belief in personal competence is sufficient to produce helplessness deficits in men. Whether attributional style rather than NA mediates the effects of failure to pain report and pain behavior, might be a topic for future research.

There was no evidence for the moderating role of NA-trait. Still, it might be of interest to investigate whether negative effects of failure feedback might be attributed to the absence or presence of specific traits. Relevant in this respect might be a comparable experiment by Ciota et al. (1998), which found that self-esteem modulated the influence of feedback on pain in that low self-esteem plus negative feedback participants reported the highest levels of pain during a cold pressor task. In line with this finding it seems very plausible that sortlike traits like NA-trait have comparable outcomes.

What is the external validity of this study? The reader is reminded that only female subjects were included which might have hampered comparison with the Levine study. In a study by Riley (1998), a meta-analysis was done on sex differences in the perception of noxious experimental stimuli. They concluded that females are more pain-sensitive than males. In case of induced thermal pain, the effect was smaller and more variable. It might be interesting to compare prediction models for males and females in future research. At present, it remains unclear if the findings are comparable between the sexes. Our prudent conclusions should therefore be restricted to female populations.

It would be worthwhile to study the effects of failure in chronic pain patients especially because this group might have been confronted with their own shortcomings and disability for a longer period of time. Responding to (non-painrelated) failure experiences with pain behavior may have become a habitual way of coping. We are currently investigating these hypotheses in chronic low back pain patients (Van den Hout et al., 2001a). If failure indeed holds pain patients in a vicious circle of trying, failing and avoiding in future situations, prevention of and dealing with failure experiences might be of interest in the management of chronic pain.

CHAPTER 4

The effects of failure feedback and pain-related fear
on pain report, pain tolerance, and pain avoidance
in chronic low back pain patients.

This chapter was published as:

Van den Hout, J.H.C., Vlaeyen, J.W.S., Houben, R.M.A., Soeters, A.P.M., Peters, M.L. (2001). The effects of failure feedback and pain-related fear on pain report, pain tolerance, and pain avoidance in chronic low back pain patients. *Pain* 92, 247-257.

ABSTRACT

The aim of this study was to investigate the influence of non pain-related failure experiences and pain-related fear on pain report, pain tolerance and pain avoidance in chronic low back pain (CLBP) patients. Moreover, the mediating and moderating role of Negative Affectivity (trait-NA) in the relationship between failure experiences and pain was examined. Seventy-six patients were divided into high and low pain-related fear groups and within each group they were randomly assigned to failure or success feedback-condition. In the first part of the study patients completed a "social empathy test" and experimenter 1 subsequently delivered false failure or success feedback. A second experimenter, who was blind for the condition, subsequently administered two lifting tasks in order to obtain measures of pain report, tolerance and avoidance. Failure feedback did have an effect on pain avoidance but unexpectedly, and not as hypothesized, pain avoidance was reduced instead of enhanced. With regard to pain report and pain tolerance similar patterns were found, but these were not statistically significant. The effect of failure feedback on pain avoidance was moderated by trait-NA: Only in the sub-group of patients who scored low on trait-NA, failure feedback decreased pain avoidance. State-NA did not mediate the effects of feedback. In line with previous findings, pain-related fear resulted in lower pain tolerance. Moreover, this study was the first to show that pain-related fear predicted higher pain report in CLBP-patients. Pain-related fear did not predict pain avoidance when pre-lifting pain and gender were controlled for. Finally, pre-lifting pain turned out to be the strongest predictor with regard to all pain measures. The role of pain-related fear and unexpected findings with regard to feedback are discussed as well as some clinical implications.

INTRODUCTION

Previous studies demonstrated that pain may be modulated by failure and success experiences. In a study described by Levine et al. (1993), failure feedback following a difficult reading comprehension task was found to increase pain intensity ratings when compared to a control condition or a condition in which success feedback was given. The study was carried out with undergraduate psychology students. A cold pressor task was conducted in order to induce pain. In a replication by Van den Hout et al. (2000), female undergraduate students executed an intelligence test and received manipulated failure or success feedback, or were assigned to a non-evaluative control task. Immediately after the manipulation, a second experimenter performed a cold pressor task in order to measure pain intensity as well as pain tolerance and avoidance. Results from this replication were comparable to those in the original study by Levine et al. (1993), but with smaller effect sizes. As an extension to the original study, it was found that not only pain report increased but also pain tolerance decreased when failure feedback was given. In an earlier study (Mayerson and Rhodewalt, 1988) it was shown that the experience of failure on an intelligence test produced significantly higher posttest pain ratings and also, performance impairment was attributed to pain to a greater extent than after success experiences. It was suggested that pain could serve a self-protective role in the event of failure.

Experimental studies have shown that perceived pain control can reduce negative effects of laboratory pain and increase pain tolerance (Arntz and Schmidt, 1989), whereas the opposite holds when pain control fails (Staub et al., 1971). In the before mentioned studies (Levine et al, 1993; Van den Hout et al., 2000; Mayerson and Rhodewalt, 1988) it was found that failure, but with regard to non-pain-related events, had outcomes that were comparable to those resulting from failed pain control. Outcomes in healthy populations suggest that failure modulates the experience of acute pain. In a chronic pain population, however, failure experiences could also contribute to the maintenance of pain over time. So far, the influence of failure and success experiences on pain and pain behavior has not been studied in a population of chronic pain patients directly. It would be interesting to speculate on the effects of non-pain-related failure on pain in chronic pain patients. Helplessness resulting from feelings of failure and frustration about not having control over pain consequences and life in general, might have become a part of everyday life. If pain has started to play a self-protective role and failure experiences indeed lead to pain increase and decreased pain tolerance, as it does in a non-patient population, failure experiences might enhance the development and maintenance of chronic pain. The effects of failure experiences on pain in chronic pain patients will be examined in this study.

A variable that is assumed to influence the impact of failure experiences is Negative Affectivity (NA, induced state or trait). Two different roles of NA in

the relation between failure and pain experiences may be considered: That as a mediator and as a moderator. Theoretically, the mediator explains how or why an independent variable affects a dependent variable, whereas the moderator specifies when and under which conditions certain effects will occur (Baron and Kenny, 1986). Levine (1993) suggested that failure might induce elevations in NA (state-NA) which, subsequently, lead to increased pain report. In other words, the induction of negative affect by failure, is the main reason why pain report increases, not failure itself. However, Van den Hout et al. (2000) did not find this mediation-effect in their replication. When NA is referred to as a mood dispositional dimension (Trait-NA; Watson and Clark, 1984), a moderator-effect of NA may be hypothesized. High-NA individuals are more prone to experience significant levels of distress and dissatisfaction at all times and in any given situation, and are said to dwell more on their failures and shortcomings (Watson and Pennebaker, 1989). An NA-disposition can hypothetically moderate the effects of failure on pain in that failure feedback affects pain experiences particularly when trait-NA is high, and much less when trait-NA is low. Although results pointed in the expected direction, the study by Van den Hout et al. (2000) did not support the hypothesis that NA moderates the effect of failure feedback on pain. In the present study, two questions are postulated and addressed regarding the interference of NA in the relation between failure and pain: 1) Does trait-NA moderate the effects of failure to pain? 2) Does state-NA mediate the effects of failure to pain?

Although no studies have been conducted with regard to the influence of failure experiences on clinical pain, other psychological variables, such as pain-related fear, have been reported to be of influence in clinical pain populations. Pain-related fear, such as fear of movement or fear of (re)injury (Kori et al., 1990), refers to an irrational fear that physical activity would instigate pain or lead to (renewed) injury. Philips (1987, p.277) suggests that "chronic pain and chronic fear - both aversive experiences which result in avoidance behaviour - may share important characteristics" and further on (p. 278) "...theoretical notions that have been developed to explain pathological fears may prove useful in explaining the characteristics of chronic pain". According to Lethem et al. (1983), avoidance, as a response to pain-related fear, results in maintenance or exacerbation of fear, and in reduction of both physical and psychosocial activities. This reduction in turn would lead to an increase of disability. During the last two decades, the fear-avoidance model described above has been supported by numerous investigations establishing the importance of fear-avoidance in the development of chronic pain (Philips et al., 1987; Waddell et al., 1993; Vlaeyen et al., 1995b,c; Asmundson et al., 1999; Vlaeyen and Linton, 2000). Pain-related fear proves a very salient predictor of pain disability in a chronic pain population, even more predictive than biomedical status and pain intensity (Waddell et al., 1993; Linton and Buer, 1995; Linton et al., 2000; Vlaeyen et al., 1995c). As in other fears, pain-related fear is found to be closely related to avoidance behavior. In two studies by Crombez

et al. (1998, 1999), avoiders reported more fear and performed worse on a behavioral task, than confronters. These results are in accordance with earlier findings by Vlaeyen et al. (1995b) that CLBP-patients with a high degree of fear of movement/(re)injury show more escape/avoidance, when exposed to a simple lifting task. In this study, pain-related fear will be measured in order to replicate the findings mentioned above. Moreover, baseline pain will be measured in order to control for possible influence on experimentally induced pain

In sum, the following hypotheses are put forward: 1) Failure feedback leads to higher pain report and pain avoidance, and to lower pain tolerance in a sample of CLBP patients; 2) Trait-NA moderates effects of failure feedback; whereas 3) state-NA mediates the effect of failure feedback on pain report, pain tolerance and pain avoidance; 4) Pain related fear enhances pain report and pain avoidance, and decreases pain tolerance.

METHODS

SUBJECTS

Seventy-seven CLBP-patients were recruited from 4 different sources: an outpatient pain center within the university hospital, a rehabilitation center, a center for physical therapy, and a group of patients who participated in a previous treatment outcome study. Inclusion criteria were: non-specific low back pain for at least 12 weeks; age between 18 and 65. Exclusion criteria were: illiteracy, pregnancy, or specific back complaints (vertebral fracture, infectious disease, rheumatoid arthritis, ankylosing spondylitis, or hernia nuclei pulposi (HNP)). In case of HNP or vertebral fracture, the patient could be included if at least 3 or 4 months respectively had elapsed since back surgery. After patients consented to be contacted for participation in the study, the researchers made an appointment and sent an informative letter. The letter was designed such that patients were led to believe that they were being invited to participate in two independent studies that, for practical reasons, were executed at the same time. The first study would be a paper and pencil task being part of a national survey on chronic illness. The second study was designed to measure the influence of emotion and pain on the performance of daily activities, for which the patients would be asked to perform a simple movement. An institutional ethics committee approved the study protocol, and all 77 patients gave written informed consent for both studies. One subject was excluded because he turned out to be older than 65. Another subject (low-TSK / failure feedback) dropped out because she refused to participate in the lifting task. Hence, 75 patients were included in the analyses.

MEASURES

Sociodemographics.

During a short interview the following data were collected: age, gender, education level, job status, duration of the complaints, medication and medical consumption.

Pain disability.

Pain disability was measured by means of the Dutch Version of the Roland Disability Questionnaire (RDQ, Roland and Morris, 1983). The RDQ is a 24-item 2-point scale measuring the extent in which the performance of daily activities was hampered due to back pain. Reliability, responsiveness and validity of the RDQ-DV were satisfactory (Gommans et al., 1996; Beurskens et al., 1996).

Manipulation check.

The manipulation check consisted of a 10-point expectancy scale. Patients were asked to indicate the expected score on a "social empathy test" (see procedure section) when compared to a general reference group. The scale was completed twice: before starting the social intelligence test, and immediately after the manipulated feedback was given. On the second occasion, the subject was asked to give expectancy-scores with regard to the completion of a similar test in the near future. The manipulation check was designed and successfully applied in a previous study (Van den Hout et al., 2000). The manipulation was considered to have succeeded if patients in the failure feedback-condition lowered their second expectancy-score and patients in the success feedback-condition increased their second expectancy score significantly. Moreover, after final debriefing, patients were asked to score on two 100-point visual analogue scales (VAS), how disconcerted and dissatisfied they were about the test results. Besides this check, patients were asked to score on a 100-point VAS how much they doubted the results on the social empathy test (0=no doubt about the authenticity of the test results; 100=being confirmed that the test results were fake), just before final debriefing. If subjects doubted the authenticity of the test results, the manipulation was assumed to have failed.

Pain-related fear.

Pain-related fear was assessed by the Tampa Scale for Kinesiophobia-Dutch Version (TSK, Kori et al., 1990). The TSK is a 17-item 4-point questionnaire that measures the fear of (re)injury due to movement. The Dutch Version has been shown to have good reliability and validity (Vlaeyen et al., 1995b,c).

Trait-NA.

Trait-NA was measured by the Dutch translation of the Negative Emotionality subscale (NEM) of the Multidimensional Personality Questionnaire (MPQ,

Tellegen et al., 1982). The NEM is a 14-item 2-point scale. High scorers on this sub-scale tend to be nervous, apprehensive, irritable, overly sensitive and emotionally labile. The NEM was preferred because it does not contain items reflecting bodily or somatic complaints which may index pain severity rather than trait-NA (Williams, 1995). Reliability of the NEM has shown to be satisfactory (Cronbach's alpha: 0.86; Stegen, 1998). Moreover, the NEM turned out to correlate significantly with Positive And Negative Affect Schedule (PANAS; Watson et al., 1988) and the Dutch version of the State Trait Anxiety Inventory (STAI; Van der Ploeg et al., 1980), Pearson r is .58 and .70 respectively (Stegen, 1998).

State-NA.

State-NA was measured by the Dutch translation of the short version of the Profile Of Mood States (POMS; Wald et al., 1990). The POMS has been shown to be a valid and responsive measure of negative affect. McNair (1971) reported effects on the POMS in the hypothesized direction with regard to short-term psychotherapies, drug trials and mood induction conditions. Because of high correlation among 3 sub-scales of the POMS (tension, anger and depression; Cronbach's alpha=0.95), these three scales were combined into one composite measure. This composite measure was further used in this study.

Pain intensity.

Pain intensity was measured by means of a VAS designed as a thermometer with a scale from 0 to 100, 0 indicating "no pain" and 100 indicating "the worst imaginable pain". The VAS was scored every 15 seconds during the lifting task (see pain tolerance).

Pain tolerance.

Pain tolerance was operationalized by means of a lifting task. For this task patients were asked to stand up and lift a 5,5-kg bag with the dominant arm and hold it until pain or physical discomfort made it impossible for the patient to continue. The experimenter registered the lifting time (in s) by means of a stopwatch. When a maximum score of 300 seconds was reached, the experimenter who indicated that the subject could place the bag on the floor and sit down, terminated the test. The lifting task was tried and successfully applied in previous studies (Vlaeyen et al., 1995b; Crombez et al., 1999).

Pain avoidance.

Pain avoidance was operationalized by means of a second lifting task. Eleven outwardly identical bags were presented to the patients. The patients were told that the bags were of an increasing weight-order and bag number 6 was the bag they had just been lifting. They were not informed about the weight of the bags. In fact weight increased with 0.5 kg in ascending order, excluding bag 6, starting with 3.25 kg (bag 1) and ending with 7.75 kg (bag 11). Subjects were asked to choose one bag from the row to perform the lifting task once more.

Bag number 6 could not be chosen, so that patients could only avoid (bag 1 to 5) or confront (bag 6 to 11). Patients might have decided to lift a higher weight, but for a short period of time. Therefore, the lifting time (in s) was clocked a second time. Pain avoidance was operationalized as the product of weight and tolerance time, hence, the higher the score on this measure, the less pain avoidance was shown.

The task used in this study to operationalize pain avoidance was not previously tested. From our data set, avoidance scores (tolerance time x weight) turned out to be normally distributed (K-S z-score = 1.064; $p > 0.2$), although frequencies were somewhat skewed to the left (skewness = 1.22). Furthermore, pain avoidance seemed to correlate in the hypothesized direction with pain report and pain tolerance (Pearson correlations: $-.59$ ($p < .001$); $.80$ ($p < .001$)) as well as with pain disability ($-.43$ ($p < .001$)).

PROCEDURE

Preparation, stratification, and randomization.

Experimenter 1 (A.S.) conducted the manipulation part of the experiment. First, two informed consent forms were signed, one for each experiment. Next, a short inventory about socio-demographic data was completed. Experimenter 1 handed out the TSK, POMS, NEM, RDQ and a VAS-scale in order to measure baseline pain intensity, and explained that these questionnaires were administered as part of the second experiment. While the subject completed the remaining questionnaires, the experimenter scored the TSK that had been completed first. Based on the TSK score, the experimenter grouped the subject in one of two strata: low TSK-score or high TSK-score. The cut-off score was 39. This had been shown to be the median in chronic back pain populations (Vlaeyen et al., 1995b; 1995c, Crombez et al., 1999). After stratification of patients in high or low TSK-groups ($n(\text{high-TSK})=32$; $n(\text{low-TSK})=44$), patients were randomly assigned to failure or success feedback according to a computer generated randomization scheme within each stratum.

Failure manipulation.

Experimenter 1 told the patients that they were going to complete a social empathy test that was supposed to measure a person's ability to empathize with social situations. The results of the test would be analyzed in the context of a national survey. The patients were led to believe that the survey was set up to test the hypothesis that chronic patients were more prone to social isolation than healthy people because of their limited social empathy skills. Before starting the test, patients were asked what score on a 0 to 10 scale they would give themselves with regard to social empathy when they referred themselves to a healthy reference group. They were also told that 70% of a Dutch population scored between 5 and 8. Subsequently, patients were given 15 minutes to fill out part 1 of the test that consisted of 29 multiple choice questions describing social situations and per item 4 possible reactions to these

situations. After completing the first part of the test, the experimenter picked up the manipulated false computer output from another room and handed it over to the subject. Patients in the success feedback-condition received a score of 8.5 and those in the failure feedback-condition were given a score of 4.3. The second part of the test consisted of 10 drawings taken from the Thematic Apperception Test (TAT, Murray et al., 1943). Per item patients could choose one out of four words that described the emotion involved in the drawing. Patients were given 5 minutes to complete this part. The rationale provided was that social empathy also depended on how fast a person could see what was going on in a social situation. After completion of this task the experimenter immediately scored the results and gave direct manipulated false computer output to the subject. Again, patients in the success feedback-condition were led to believe that they had a high score (8.9) on the test, and patients in the failure feedback-condition that they had also failed this part and scored 4.5. Besides the specific output on the results, an overall low (4.4) or high score (8.7) was given. Experimenter 1 asked patients to fill out the expectancy scale once more. This time patients were asked to give expectancy-scores with regard to completion of a comparable test in the near future. Finally, experimenter 1 handed patients directly over to experimenter 2 (R.H.) who was blinded for the feedback-condition.

Lifting task.

After a short introduction, patients were asked to fill out POMS and pain-VAS once more. Before starting the lifting task, if necessary, patients were asked to remove their watches so that they would not have any direct time reference. Patients were instructed to verbally report pain intensity on a 0 to 100-point VAS-scale every time they heard an audible signal during lifting (every 15 seconds, the first signal was given after 10 seconds). Patients could see the VAS-scale as an enlarged thermometer attached to the wall in front of them. Experimenter 2 gave the following information about the task: "The bag that you are going to lift in a moment, contains a certain weight. The intention of the task is to lift the bag with your dominant hand, as long as possible until you decide to stop because of pain or other discomfort. When you want to stop, say "stop" and put the bag back on the floor." Now the patients could start the lifting task. Immediately after they quit the lifting task, they were asked to score the pain-VAS once more, this time with regard to the pain they experienced just before quitting the lifting task. To measure avoidance behavior, a second task was added to this part of the experiment. The patients were asked to perform the lifting task once more with the bag they chose from the row. Again, tolerance time was clocked and taken into account when calculating the avoidance indices. Finally, experimenter 2 debriefed patients and administered the measures concerning disconcertion and dissatisfaction to check whether the experimental manipulation was successful.

STATISTICAL ANALYSES

To test the hypotheses formulated in the introduction, multiple linear regression analyses were conducted for all three dependent measures (pain report, pain tolerance and pain avoidance), separately. Pain reports after 100 s were taken into analyses. In case patients had already terminated the task, pain intensity scores patients retrospectively reported immediately after quitting the lifting task, regarding the moment just before quitting, were imputed. Independent variables taken into analyses were: gender (male=1, female=2), pain report before starting the lifting task (mean score of the two pre-lifting measurements), feedback (success=1, failure=2), pain-related fear (TSK), and trait-NA (NEM). For each dependent variable separately, a hierarchical backward elimination method was used to filter significant main effects ($P \leq 0.1$; Hypothesis 1 and 4). A priori, it was decided that pain-related fear and feedback always remained in the model. To test moderating and mediating effects, analyses were carried out according to Baron and Kenny (1986). In order to test moderating effects, an interaction variable was added to the model: trait-NA * feedback interaction. To prevent collinearity, the interaction-variable was centered. First, it was tested whether the interaction variable was significant when the main variables (trait-NA and feedback) were still in the model ($P \leq 0.05$; Hypothesis 2). When the interaction variable was significant, a median split on the independent variable (trait-NA) was used to analyze the model within each sub-group. To test mediation (Hypothesis 3), state-NA (post-manipulation) was regressed on the independent variables. If feedback added significantly to the prediction of state-NA (post-manipulation), it was tested whether the predictive value of feedback on the dependent variable decreased when state-NA (post-manipulation) was added to the model (assuming that feedback already added to the predictive model of the dependent variable). Finally, VIF-values were calculated to make sure that none of the regression analyses violated the assumption of collinearity.

RESULTS

SOCIO-DEMOGRAPHICS

Seventy-five patients were included in the analyses, 25 men and 50 women, with an average age of 46 years ($SD = 8.4$; range: 29-63 years). Mean pain duration was about 12 years ($SD = 8.9$ years; range: 4 months-39 years). Of all patients, 37% was employed and received no disability compensation, 14% received disability compensation, 13% was partly compensated and partly employed, about 12% was sicklisted at the time of the interview, 4% were registered as unemployed, and 20% did not have paid work. Almost all patients (91%) had been seeing a physical therapist while 61% had consulted a medical specialist for their back complaints. Forty-seven percent of all patients used analgesics and 20% used psychopharmacologic medication. Twenty-eight percent had been using analgesics or psychopharmacologic

medication on the day of the experiment. The mean RDQ-DV score was 12.6, and more than half of the patients (53%) had an RDQ-score >15, suggesting that the patients were moderately to highly disabled (Roland et al., 1983; Patrick et al., 1995).

MANIPULATION CHECK

According to the a priori agreed criteria, the experimental manipulation was successful. In the failure feedback-condition patients lowered their expectancy scores 1.42 points on a 10-point scale ($t(36)=7.49$; $P<0.001$), while in the success feedback-condition expectancy scores were increased 1.28 points ($t(37)=-6.67$; $P<0.001$). Moreover, patients in the failure feedback-condition were more disconcerted ($t(73)=-5.34$; $P<0.001$) and less satisfied ($t(58.78)=-7.81$; $P<0.001$) about their results than patients in the success feedback-condition. Remarkably, patients in the failure feedback-condition expressed more doubt about the test results than patients in the success feedback-condition (21.7 vs. 7.3; $t(53.82)=-3.26$; $P<0.01$). However, none of the patients had questioned the authenticity of the test.

RANDOMIZATION CHECK

The four subgroups (high/low TSK * failure/success feedback), did not differ on sociodemographics, previous health care utilization and current medication use (Chi-square tests, NS), suggesting that the randomization procedure was successful. Fewer men were assigned to the low TSK*failure feedback condition, but this difference was not significant. ANOVA's showed that pain disability, pre-lifting pain report and trait-NA significantly differed between the groups. The differences were due to differences between high-TSK and low-TSK groups. Unequal means were not surprising because correlations were found between pain disability and pain-related fear ($r=.48$; $P<0.01$), pre-lifting pain and pain-related fear ($r=.35$; $P<0.01$), and trait-NA and pain-related fear ($r=.33$, $P<0.01$).

ANALYZING THE MODEL

Figure 1 displays pain intensity reports at 15-second intervals during the lifting task, in four conditions (high/low TSK * failure/success feedback). Although a steady increase of pain intensity is seen up to 100 seconds, curves of all four conditions flatten beyond this point. The most important reason for this is that at a 100-seconds interval 44.7% of all patients already quitted the lifting task and pain intensity scores reported immediately after quitting the lifting task, regarding the moment just before quitting, were imputed for the remaining intervals.

Table 1 and figure 1 illustrate that highest pain intensity was reported in the high TSK * success condition, whereas lowest pain intensity was reported in the low TSK * failure condition. A similar pattern was found with regard to pain tolerance and pain avoidance. Patients in the low TSK * failure condition,

FIGURE 1. Group means (feedback/painrelated fear) on pain intensity during the first lifting task. For this comparison, groups were created by median-split on basis of the Tampa Scale for Kinesiophobia (TSK): TSK-Low ≤ 39 ; TSK-High > 39 .

lifted the bag about 63 seconds longer than patients in the high TSK * success condition did. Furthermore, regarding pain avoidance, patients in the low TSK * failure condition approached the task more than patients in the other conditions did. The difference in pain avoidance score was particularly due to increased tolerance time, as the weight chosen in this condition did not differ from the weight chosen by the other conditions (table 1). Table 2 shows the results of regression analyses for all three dependent variables separately. Trait-NA was eliminated from the equation in all three analyses.

TABLE 1. Group means and standard deviations (SD) on pain report, pain tolerance, and pain avoidance.¹

Dependent variable	Low-TSK ² /success feedback (n=22)	Low-TSK/failure feedback (n=21)	HighTSK/success feedback (n=16)	HighTSK/failure feedback (n=16)
Pain report (100 s)	41.27 (27.15)	37.38 (26.95)	71.75 (25.19)	58.31 (17.91)
Pain tolerance (in s)	155.73 (102.35)	175.76 (117.85)	112.82 (93.88)	133.19 (95.37)
Pain avoidance ³	621.94 (426.57)	986.24 (717.76)	554.89 (390.93)	639.78 (490.38)
Weight (kg)	5.70 (1.30)	5.73 (1.51)	5.13 (1.36)	5.53 (1.57)
Tolerance time (in s)	109.32 (70.15)	169.14 (107.80)	106.56 (77.19)	122.88 (97.22)

¹ Values are given as the mean (SD)

² Tampa Scale for Kinesiophobia: pain-related fear.

³ A high score on this scale means less avoidance.

TABLE 2. Summary of multiple regression analyses of (1) pain intensity, (2) pain tolerance, and (3) pain avoidance as dependent variables, and feedback (success=1, failure=2), pain-related fear (TSK), prelifting pain (pain report (0-100) before starting the lifting task), gender (male=1, female=2), and trait-NA (NEM) as independent variables. 4) Summary of multiple regression analysis of pain avoidance as a dependent variable, and feedback, pain-related fear, prelifting pain, gender, trait-NA, and trait-NA*feedback interaction as independent variables, in order to test moderation of trait-NA.

Dependent variable	N	Adj. R^2	Variables in the model	Standardized Bèta
1) Pain intensity (after 100 sec.)	75	0.554***	- Feedback (Success vs. Failure) - Pain-related fear - Pre-lifting pain - Gender	-.136~ .216* .649*** .224**
2) Pain tolerance	75	0.237***	- Feedback - Pain-related fear - Pre-lifting pain - Gender	.146 -.225* -.340** -.344**
3) Pain avoidance	75	0.191**	- Feedback - Pain-related fear - Pre-lifting pain - Gender	.261* -.155 -.307** -.274*
4) Pain avoidance	75	0.212**	- Feedback - Pain-related fear - Pre-lifting pain - Gender - Trait-NA - Trait-NA*Feedback	.263* -.148 -.276* -.268* .005 .208~

~ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

ses (analysis 1 to 3). The remaining variables (gender, pre-lifting pain, pain-related fear and feedback), regarding the models of pain intensity, pain tolerance and pain avoidance, explained about 55%, 24% and 19% of the variance, respectively. In all models, pre-lifting pain was the most powerful predictor: the higher pre-lifting pain, the higher pain intensity report during the lifting task, the lower pain tolerance, and the higher pain avoidance. Gender was predictive in each of the models, in that females reported higher pain intensity, tolerated less and avoided more. Outcomes with regard to failure feedback, negative affect, and pain-related fear will be discussed separately in the following sections.

DOES FAILURE FEEDBACK INFLUENCE PAIN REPORT, PAIN TOLERANCE AND PAIN AVOIDANCE? (HYPOTHESIS 1).

Table 2 shows that feedback condition was not a significant predictor in the models of pain intensity and pain tolerance, whereas statistical significance was found in case of pain avoidance. Contrary to the a priori set hypothesis, failure feedback resulted in confronting behavior rather than avoidance. Although not significant, the same pattern was found with regard to pain intensity and pain tolerance, in that failure feedback resulted in decreased pain intensity reports and increased tolerance times.

THE MODERATING ROLE OF TRAIT-NA (HYPOTHESIS 2)

With regard to pain report and pain tolerance, the interaction variable (trait-NA * feedback) was eliminated from the regression model. However, with regard to pain avoidance, interaction between feedback and trait-NA showed significance ($P=0.05$; table 2, analysis 4). Figure 2 displays mean pain avoidance scores for high/low trait-NA * feedback. Patients were divided into high/low trait-NA groups by means of a median-split (median=3). Figure 2 suggests that failure feedback decreases avoidance, while success feedback enhances it, but only in the low trait-NA condition. Results of regression analyses for high and low trait-NA condition separately, reveal that within low trait-NA condition, feedback ($\beta=-.46$; $P<0.01$), and gender ($\beta=-.43$; $P<0.01$) predict pain avoidance, whereas in the high trait-NA condition none of the independent variables turned out to be predictive. In contrast with the main analyses, pre-lifting pain was not predictive in either of the sub-analyses. In sum, trait-NA moderates the relation between feedback and pain avoidance. Only if trait-NA is low, failure feedback is predictive of pain avoidance.

THE MEDIATING ROLE OF STATE-NA (HYPOTHESIS 3)

In order to investigate the mediating role of state-NA (POMS), it was first checked whether feedback significantly predicted the dependent variable (pain report, pain tolerance and pain avoidance). As described in the foregoing section, this was only the case with regard to pain avoidance. Secondly, it was tested whether feedback was predictive of post-manipulation state-NA. This turned out not to be the case. The only variable that had predictive value was pre-manipulation state-NA. From this analysis it can be concluded that the effect of feedback on pain avoidance was not mediated by state-NA. Results were not surprising because both in failure and success feedback-conditions negative affect decreased after the manipulation. The pre-post manipulation change on the POMS, did not differ between both feedback-conditions ($t(52.04)=-0.66$; $P>0.5$).

FIGURE 2. Moderation effect of Negative Affectivity (trait-NA) on pain avoidance (pain tolerance X weight, second lifting task). For the comparison, groups were created by median split on basis of the Negative Emotionality Scale (NEM; trait-NA): Low trait-NA ≤ 3 ; High trait-NA > 3 .

DOES PAIN-RELATED FEAR INFLUENCE PAIN REPORT, PAIN TOLERANCE AND PAIN AVOIDANCE? (HYPOTHESIS 4)

Table 2, analysis 1 to 3, shows that pain-related fear was predictive in the model of pain report and pain tolerance. Pain-related fear, as hypothesized, increased pain intensity reports and decreased tolerance times during the lifting task, even when pre-lifting pain was controlled for. The association between pain-related fear and pain intensity was already present at baseline, as can be seen in figure 1. Pain-related fear did not predict pain avoidance when pre-lifting pain and gender were controlled for. In all cases, pre-lifting pain turned out to be the most powerful predictor, and thus more influential than pain-related fear.

DISCUSSION

The aim of this study was to investigate the effects of failure versus success experiences on pain report, pain tolerance and pain avoidance in CLBP-patients. Moreover, it was examined whether the effects of failure feedback were moderated by trait-NA and/or mediated by state-NA. Finally, the effects of pain-related fear on pain report, tolerance and avoidance were analyzed.

Contrary to the a priori formulated hypotheses, failure feedback resulted in decreased pain avoidance behavior as compared to success feedback. With regard to pain report and pain tolerance similar patterns were observed in that failure feedback decreased pain intensity reports and increased tolerance times, but these results were not statistically significant. Trait-NA moderated the effects of failure feedback on pain avoidance in that pain avoidance only decreased if trait-NA was low. Although the same pattern was shown with regard to pain report and tolerance, moderation was not statistically significant in those cases. The relation between failure feedback and pain was not mediated by state-NA. In answer to the fourth hypothesis, it was concluded that pain-related fear resulted in higher pain report and lower pain tolerance, but did not predict pain avoidance when pre-lifting pain and gender were controlled for. Pre-lifting pain was the strongest predictor with regard to all pain measures, in that higher pre-lifting pain increased pain intensity report and pain avoidance behavior, and decreased pain tolerance.

How can contrary findings with regard to failure feedback be explained? First it might be questioned whether the manipulation succeeded. From the manipulation check it became clear that the success and failure manipulation were successful. Nevertheless, some questions might be posed with regard to the failure manipulation and whether feedback influenced patients as expected. Although not explicitly hypothesized, one would expect that the manipulation would affect mood as well. However, state-NA decreased in both failure and success feedback-conditions. Therefore, it might be questionable whether feedback impressed patients enough to differentiate not only on outcome expectancies, but also on outcome-compatible mood. Furthermore, it is not known how much importance patients attributed to being evaluated on social empathy. If importance was low, by the time patients arrived at the second experimenter to accomplish the lifting task, the impact of the manipulation might already have faded away. With regard to future research, it is suggested that manipulation impact may improve if failure and success experiences are of personal relevance.

Secondly, it must be noticed that conclusions only apply to half the sample. Subgroup analyses revealed that, at least for pain avoidance, the model only holds for low trait-NA patients (moderation effect). This might explain the opposite findings: Not being bothered by negative thoughts and emotions about the foregoing experience, failure feedback rather stimulated low-NA patients to do their best and prove that they were able to perform well in the subsequent task. In addition, patients were led to believe that the study wanted to investigate whether chronic patients were more prone to social isolation than healthy people because of their limited social empathy. This might have encouraged patients to show the second experimenter that this suggestion did not apply to their case. It might be interesting in future studies to investigate if small frustrations encourage rather than discourage a subgroup of pain patients to increase activities.

In support of the foregoing, post-hoc comparisons revealed that low-NA

patients doubted failure feedback more than success feedback, whereas for high-NA patients no difference was found. This might indicate that low-NA patients not only were more eager to compensate for their failure, but moreover were not willing to take the suggestion that they had impaired social empathy, perhaps because they were more convinced of their own abilities. The fact that low-NA patients rarely doubted success feedback might be another indication of positive and stable beliefs about self. Similar findings were found by Ciota et al. (1998). In their study, highest pain levels were reported after negative feedback, but only if participants had low self-esteem. A hypothesis that might be distracted from the foregoing is that pain patients, who show a low degree of negative affectivity and have moderate to high self-esteem, achieve best when they are somewhat frustrated about their achievements in a foregoing task.

A similar pattern as with regard to pain avoidance is seen for pain tolerance, but neither feedback as a predictor nor trait-NA as a moderator reached significance. If patients were indeed trying to compensate for their failure, the sequence in which the lifting tasks took place might explain different outcomes on pain tolerance and avoidance. Although the first lifting task (pain tolerance) was without reference about task norms, the second lifting task (pain avoidance) gave the opportunity to do at least better than during the previous task by lifting the bag even longer. In line with this, patients could better anticipate the second lifting task than the first. In addition, patients in the success feedback condition had already earned their credits, and could have been more hampered by feelings of fatigue and lack of motivation during the second task. This might have increased the contrast between feedback conditions from first to second lifting task.

In the high-NA subsample, feedback did not seem to matter at all. A possible explanation might be that in CLBP-patients other mechanisms have started to play a more important role. One single failure or success experience might do in healthy controls (Levine et al., 1993; Van den Hout et al., 2000), but might not make any difference in a CLBP population who tend to be high negative affective and may hold a negative view on their lives anyway.

Another interesting hypothesis, might be that failure experiences were attributed to external and specific causes ("this test was not right", "I might have failed on social empathy, but other social abilities are okay"), whereas success was attributed internally ("the test proved that I have got a talent for dealing with social situations"). In literature on attributions in failure and success situations, the phenomenon described above is known as ego serving bias (Green et al., 1994) or self-protective attributions (Mayerson and Rhodewalt, 1988), in that it protects the ego from devaluation after failure experiences. This kind of bias might have inhibited the impact of the manipulation by leading patients to think that it was not due to themselves in general that they failed on the social empathy test, but to something specific that was beyond their control.

Finally, it would be interesting to study the role of trait-NA in more detail.

Because scores on the NEM were skewed to the left (median=3), it should be questioned if the high trait-NA group had a NEM-score that was substantial ($M=8.03$; $SD=3.25$). The overall NEM-score turned out to be 10 ($SD=6.8$) in one of our previous studies (Crombez et al., 1999) with a comparable clinical back pain patient sample. The inclusion of more patients scoring high on trait-NA, would normalize the distribution and increase the likelihood of stronger moderation effects of trait-NA.

What are the clinical implications of these findings? Negative affectivity and failure experiences turned out to be of importance in clinical pain and pain behavior. If low trait-NA patients are confronted with non pain-related failure as compared to success, the absence of trait-NA seems to protect them from increasing pain and even stimulates them to perform better on a subsequent task. On the other hand, neither failure nor success seem to affect patients with more elevated NEM-scores, what might be translated as an overall shortcoming to feel reinforced by positive events. In line with the foregoing, it could be hypothesized that positive affectivity (trait-PA) has protective value when patients are confronted with failure. However, this remains to be examined¹.

Pain-related fear affected pain intensity and tolerance in the hypothesized direction. Pain-related fear being predictive of pain intensity in a clinical back pain sample was not reported before. How can the influence of pain-related fear on pain report be explained? Possibly, pain-related fear makes chronic pain patients to attend more to pain-related information and consequently experience more pain (Asmundson et al., 1997). Another explanation might be given by the fact that pain-related fear is related to increased muscle reaction which subsequently results in higher pain report (Vlaeyen et al., 1999). Recently, anxiety sensitivity appeared to be indirectly associated with pain through its contribution to anxiety during a cold pressor task (Schmidt and Cook, 1999). Asmundson and colleagues (1996) suggested that since anxiety sensitivity marks an individual's general propensity to develop fears, it might amplify a specific tendency to develop fear of pain. Hence, another hypothesis might be that pain-related fear is not directly associated to pain report, but both pain-related fear and pain report are positively related to anxiety sensitivity. Obviously, in the current study, pre-lifting pain already varied with pain-related fear beforehand, when no manipulation took place yet (Pearson correlation = .35; $P<0.01$). It would be of interest to investigate if both variables are causally related or influenced by some common factor.

Pain related fear did not predict pain avoidance when prelifting pain and gender were controlled for. Pain-related fear not adding to the predictive model of pain avoidance is somewhat surprising because previous studies (Vlaeyen et al. 1995b; Crombez et al., 1998; Crombez et al. 1999) confirmed the role of pain-related fear on avoidance. A plausible explanation might be the inclusion of pre-lifting pain, which has a high correlation with pain-related fear at baseline. When pre-lifting pain was deleted from the model, pain-related fear did indeed predict

Note 1 In our study positive affectivity was measured by means of the Positive Emotionality subscale (PEM) of the MPQ (Tellegen et al., 1982). PEM was significantly correlated with NEM (Pearson correlation = -.47; $P<0.001$). As expected, effects of failure feedback in the high PA subgroup were comparable to those in the low NA subgroup. Trait-PA however, did not significantly moderate the effect of failure feedback.

pain avoidance ($\beta = -0.26$; $P < .05$), in the sense that the higher pain-related fear, the more avoidance behavior was shown by the patients during the lifting task. From this might be concluded that obviously, pre-lifting pain is a variable of which the influence must not be underscored.

The pain avoidance measure has not been used previously, and its validity might be questioned. Post-hoc analyses of lifting time and weight as separate dependent variables revealed that the effects were mainly due to lifting time, and not to the weight chosen. Therefore, it might be reasonable to presume that the pain avoidance measure was in fact tapping pain tolerance instead of avoidance. On the other hand, an important contrast between tolerance and avoidance measures might result from the sequence in which the tasks are given. During the avoidance task, being the second task, cognitive processes like pain expectation and anticipation, might have started to play a more important role now the patient has the reference of the first lifting task. Still, it might be difficult to define both variables in one experiment independently. Although pain tolerance and pain avoidance are presented as separate constructs in this study, it might be more prudent to suggest that we possibly introduced a second measure of pain tolerance.

It can not be denied that pre-lifting pain had the most powerful and constant predictive value with regard to all three dependent variables. According to current findings, high pain reports in rest predict higher pain levels during activity, decrease pain tolerance and increase pain avoidance. These results might have clinical relevance in that initial high pain-levels during therapy might inhibit or prevent therapy progression when not attended to.

With regard to pain-related fear, and in line with the state-of-the-art literature (Crombez et al., 1999; Vlaeyen and Linton, 2000), we found support for its role on pain report and tolerance. What might be the clinical relevance of this finding? Several investigators have indeed introduced and adapted cognitive-behavioral treatments, and successfully applied these in the treatment of chronic pain (e.g. Fordyce, 1986; Lindström et al., 1992; Linton et al., 1989, 1992; Philips et al., 1991; Kole-Snijders et al., 1999). For patients reporting excessive pain-related fears, treatments have moved from graded exposure to movements in general, to exposure in vivo to very specific and idiosyncratic movements that are being feared (Vlaeyen et al., 2001). Indeed, from a cognitive-behavioral point of view, excessive fear is best treated by exposing the subject to the object or situation feared (Philips, 1987). Therefore, patients in whom high levels of pain-related fear are detected, might only profit from intervention if it includes physical exposure to the movements avoided, and challenge of the fearful thoughts that go along with it. Tracing personally relevant pain-related fears and related dysfunctional thoughts might be crucial to increase the success-rate of treatments in the field of pain psychology.

Finally, as this was the first study to investigate the influence of failure experiences in chronic pain patients, prudence with regard to the interpretation of unexpected and surprising results is called for.

CHAPTER 5

Functional disability in non-specific low back pain: the role of pain-related fear and problem-solving skills.

This chapter was published as:

Van den Hout, J.H.C., Vlaeyen, J.W.S., Heuts, P.H.T.G., Sillen, W.J.T., Willen, J.E.H.L. (2001). Functional disability in non-specific low back pain: the role of pain-related fear and problem-solving skills. *International Journal of Behavioral Medicine*, 8 (2), 134-148.

ABSTRACT

Previous studies have shown that pain disability in chronic back pain patients is closely associated with pain-related fear and avoidance. The current study was aimed at replicating this finding in a sample of low back pain patients with a new episode of sick leave. In addition, the supplemental value of problem solving in predicting pain disability was examined. It was further hypothesized that problem solving would moderate the effects of daily stress on pain disability. The results were partly as predicted. Pain-related fear, pain intensity, pain catastrophizing, daily stress, and problem solving were found to be individually related to functional disability, of which pain intensity and pain catastrophizing were the strongest predictors. Problem solving did not add predictive value regarding pain disability; neither could the moderating role of problem solving be confirmed. Surprisingly, and in disagreement with previous findings, pain intensity was found to be closely related to disability in this sample of low back pain patients.

INTRODUCTION

Why do some pain patients succeed in resuming their daily activities within a few days after pain onset, even when pain is still present, whereas others fail to return to work and gradually develop a chronic pain syndrome? This question can only be answered by considering a multidimensional process in which psychosocial factors play an important role (Gatchel & Turk, 1996). This study highlights two potential mediators between the onset of pain and chronic disability, namely pain-related fear and problem solving.

Pain-related fear has been frequently described over the last two decades (Kori, Miller, & Todd, 1990; Lethem, Slade, Troup, & Bentley, 1983; Philips, 1987; Waddell, Newton, Henderson, Somerville, & Main, 1993). Besides pain catastrophizing, avoidance behavior, disuse and depression, pain-related fear turned out to be an important process in a model predicting chronic pain disability (Vlaeyen, Kole Snijders, Boeren, & van Eek, 1995b). The processes mentioned above have advantages in that they describe a behavioral process that can be changed without invasive interventions, which makes them applicable in the cognitive-behavioral management of pain. Although a lot of theoretical work has been done so far, studies to find empirical evidence in favor of the model have started only recently. In three independent cross-sectional studies described by Crombez, Vlaeyen, Heuts, and Lysens (1999), pain-related fear measures were superior to pain intensity, pain duration and negative affect in predicting self-reported disability and poor behavioral performance in chronic low back pain patients. Moreover, a prospective study by Klenerman et al. (1995) found evidence for the idea that pain-related fear is a precursor, and not a consequence, of disability.

In addition to pain-related fear and avoidance behavior, pain coping and adjustment to pain have received considerable attention (Burton, Tillotson, Main, & Hollis, 1995). However, according to a review by Jensen, Turner, Romano, & Karoly (1991b), inconsistent results and methodological problems limit any conclusions regarding the strength and nature of the association between pain coping strategies and adjustment to pain in chronic pain patients. More recently, it has been suggested that attention should shift from coping with pain to coping with the consequences of pain (Aldrich, Eccleston, & Crombez, 2000). The authors hypothesize that chronification is the result of perseverance in ineffective and misdirected problem solving. In their opinion, chronic pain is an insoluble problem. Therefore, to prevent disability in the longer term, it may be even more important to focus upon coping with the consequences of pain in everyday life, and to take into account how people cope with problems, stressful events, and daily hassles, rather than coping with pain itself.

A number of hypotheses regarding the role of problem solving can be generated from the assumptions put forward by Aldrich et al. (2000). First, daily hassles and stressful life events, which might or might not result from

pain, influence disability in the longer term. Secondly, for patients with recurrent pain, coping with the consequences of pain may be hampered by the lack of adequate problem solving, resulting in increased disability. In particular, the appraisal of one's self-efficacy in overcoming problems (problem orientation) and the presence of rational problem solving skills to solve problems one is faced with (problem solving skills)¹ might influence the extent to which a person is disabled by pain in the long term. Moreover, the process of negative reinforcement (Fordyce, 1976) may give cause to pain behaviors to persist, as they inadvertently allow the successful avoidance of daily problems for which no useful solution can be generated. A third hypothesis might follow from the previous two hypotheses, namely that the influence of daily hassles on disability depends on the quality of individual problem solving.

Studies that did study the direct or indirect influence of problem solving on quality of life did not concern the pain-related domain. Nezu, Nezu, Friedman, Faddis, and Houts (1998) suggested that chronic illness (such as cancer) is accompanied by medical and psychosocial problems that affect the patient's quality of life. However, the impact that illness and psychosocial problems have on daily activities depends on how capable the patient is in using problem solving skills.

The model proposed by Nezu et al. (1998), mostly with regard to depression and cancer, may also apply to other chronic illnesses, such as back pain. The quality of life back pain patients experience might largely be reflected in the level of disability because of their back problem. If an association between problem solving and disability exists, it may give new directions to the management and prevention of chronic back pain disability and to addressing a growing socio-economic problem.

The current study aimed to replicate some of the findings by Crombez et al. (1999), by examining in particular the role of pain intensity, pain catastrophizing and pain-related fear on functional disability. Moreover, the study intended to expand the model by including daily stress and problem solving in analogy to the model proposed by Nezu et al. (1998). The following hypotheses were tested. 1) Pain intensity, pain catastrophizing, pain-related fear, daily stress and problem solving are individually related to functional disability. 2) Problem solving has additional value in a model incorporating pain, pain catastrophizing, pain-related fear, and daily stress to explain functional disability. 3) Problem solving moderates the effects of daily stress on functional disability.

METHODS

DESIGN AND SUBJECTS

The patients included in the present study had all been included in a randomized clinical trial investigating the supplemental value of problem solving therapy added to a graded activity program (Van den Hout et al., 1998). The

Note 1 The term 'problem solving' points at both problem orientation and problem solving skills.

baseline measurements of this trial were used in the current study. Patients were all employees referred to the study by General Practitioners, Occupational Physicians or Rehabilitation Physicians. Eligibility criteria were as follows: age between 18 and 65 years, low back pain (LBP) for more than 6 weeks, on sick leave with LBP but no longer than 20 weeks, no more than 120 days of sick leave during the last year, informed consent given to take part in a randomized controlled trial that would proceed after baseline measurements. Non-eligible were those patients who had demonstrable specific back complaints (vertebral fracture, infectious disease, rheumatoid arthritis, ankylosing spondylitis, or herniated disc; Waddell & Turk, 1992), were pregnant, were not proficient in Dutch, or had been referred to a medical specialist for diagnosis or treatment of their LBP at the time of referral. Furthermore, patients had to agree that they would stop other therapies they were receiving for their back complaints at the start of the intervention.

During the selection procedure the Symptom Check List 90 (SCL-90; Arrindell & Ettema, 1986) and the Dutch Personality Questionnaire (NPV; Luteijn, Starren, & Van Dijk, 1985) were used to check for psychopathology that would hamper individual or group processes. It had been decided in advance that patients would be excluded if scores on the SCL-90 were 'high', relative to those of ambulatory psychiatric patients, regarding "Insufficiency of thinking and acting" or the "Suspicion and interpersonal sensitivity" and "Hostility" subscales together, or if the total score on "Psychoneuroticism" was 'high' (Arrindell & Ettema, 1986). Furthermore, subjects would be excluded if scores on NPV were 'high', relative to those of a Dutch standard population, regarding the "Rigidity", "Discontentment", and "Self-satisfaction" subscales together (Luteijn et al., 1985). In the period between September 1996 and December 1998, 123 patients were included in the study. One subject did not complete the Tampa Scale for Kinesiophobia correctly, which was detected when the subject had already started the intervention. Baseline measurements of the remaining 122 subjects were included in the analyses.

MEASURES

Demographic variables.

Demographic variables were obtained by means of a semi-structured interview. Days of sick leave were measured by the number of calendar days the patient had been sicklisted because of back pain from the first day of sick leave up to the moment of intake.

Disability.

The Quebec Back Pain Disability Scale (QBPDS; Kopec et al., 1995) was included as a measure of pain disability. The QBPDS is a 20-item 6-point self-report scale describing activities commonly affected by back pain. Subjects have to rate the level of difficulty executing these activities. The QBPDS has been found to be a reliable, valid, and responsive outcome measure (Gommans,

Koes, & Van Tulder, 1996; Kopec et al., 1995; Schoppink, van Tulder, Koes, Beurskens, & de Bie, 1996). Secondly, the Dutch Version of the Roland Disability Questionnaire (RDQ; Roland & Morris, 1983) was administered to compare functional disability with that of a reference group. RDQ and QBPDS had a Pearson correlation of .76 in the current study. An advantage of the QBPDS as compared to the RDQ may be that the QBPDS has been specifically designed for patients with low back pain by selection and reduction of items, whereas the RDQ has been derived from the more generic Sickness Impact Profile (Bergner, Bobbitt, Carter, & Gilson, 1981).

Pain intensity.

Pain intensity was measured by the Dutch version of the McGill Pain Questionnaire (MPQ-DLV; Van der Kloot, Oostendorp, van der Meij, & van den Heuvel, 1995), a frequently used and reliable measure of pain intensity. The MPQ-DLV consists of 20 groups of 3 or 4 words describing pain qualities. The subject can choose those words that fit in best with his/her pain. The total Pain Rating Index (PRI) is calculated by adding the individual values of the words marked.

Pain Catastrophizing.

The Pain Catastrophizing Scale (PCS; Crombez & Vlaeyen, 1996; Sullivan, Bishop, & Pivik, 1995) was applied to measure catastrophizing thoughts about pain. The PCS is a 13-item 5-point scale that has proved satisfactory internal consistency for the total scale (Cronbach $\alpha = .93$; Osman et al., 1997), as well as good temporal stability (Pearson's $r = .92$; Crombez, Vervaeke, Lysens, Baeyens, & Eelen, 1998). The total score was used in the present study.

Fear of movement/(re)injury.

Pain-related fear was assessed by the Dutch version of the Tampa Scale for Kinesiophobia (TSK; Kori et al., 1990). The TSK is a 17-item 4-point scale that measures the fear of injury or re-injury due to movement. The Dutch version has been shown to possess good reliability and validity (Vlaeyen et al., 1995b).

Fear avoidance beliefs.

Fear avoidance beliefs were measured by the Dutch version of the Fear Avoidance Beliefs Questionnaire (FABQ; Waddell et al., 1993). The FABQ consists of two factors and focuses specifically on patients' beliefs about how physical activity (FABQ-act) and work (FABQ-work) affect their low back pain. Test-retest reproducibility and internal consistency of the FABQ have turned out to be satisfactory (Waddell et al., 1993).

Daily stress.

The Everyday Problem Checklist (EPCL; Vingerhoets & Van Tilburg, 1994) was used to measure daily stress. The EPCL consists of 114 items describing eve-

ryday problems from several domains of daily life. First, the question is asked whether the problem described in the item has occurred during the past two months. If the answer is affirmative, the experienced intensity of the problem is scored on a four-point scale. By counting the number of items answered in the affirmative, a total frequency score is obtained (FREQ). Tallying up the intensities of the items yields a total intensity-score (INT). The EPCL has been found to have good test-retest reliability (coefficients of .87 and .85 for FREQ and INT respectively) and validity, and has turned out to be predictive of health problems (Vingerhoets & Van Tilburg, 1994).

Problem solving.

Problem solving was measured by the Dutch translation of the Social Problem Solving Inventory-Revised (SPSI-R; D'Zurilla, Nezu, & Maydeu-Olivares, 1994). The SPSI-R is a 52-item self-report measure of social problem solving processes. The SPSI-R consists of five scales measuring two different problem-orientation dimensions (positive (PPO) and negative (NPO)), and three problem solving proper dimensions (rational problem solving (RPS), impulsive/carelessness style (IMP) and avoidance style (AS)). In the current study, Cronbach's Alpha ranged from .70 to .93.

STATISTICAL ANALYSES

Multiple linear regression analyses were conducted to test the hypotheses mentioned in the introduction. To test hypothesis 1, five blocks of independent variables were separately entered into the analysis: 1) pain intensity (MPQ-DLV); 2) pain catastrophizing (PCL); 3) pain-related fear (TSK, FABQ-work, FABQ-act); 4) daily stress (EPCL-freq, EPCL-int); 5) problem solving (SPSI-R: PPO, NPO, RPS, AS, & IMP). With every block of independent variables, sex, age, and days of sick leave were included in the analysis as covariates. A hierarchical backward elimination method was used to identify significant 'predictors'.² First, non-significant variables ($p > .1$) were eliminated in a stepwise process from each block of independent variables. It had been decided in advance that covariates would always remain in the model. Secondly, a model was constructed involving the stepwise addition of blocks of independents in order to test hypothesis 2. Only variables that were identified as significant predictors were added. At each step it was checked whether variables did add to the model ($p < 0.1$). If not, they were eliminated from the analysis in the next step. The order in which the blocks were added was fixed as indicated above. By adding the block of 'problem solving' last, it was tested whether this block would add to the model constructed thus far. Hypothesis 3 was tested using multiple regression analysis according to Baron and Kenny (1986). To test whether problem solving moderated the effects of daily stress on disability, the interaction between problem solving and daily stress was added to the main variables (SPSI-subscale and EPCL-subscale) and the covariates (sex, age, and days of sick leave). In order to prevent inflation of Type

Note 2 'Predictor' is here used in a statistical sense, not in a causal one. Authors are aware of the fact that the term 'predictor' in a causal sense would be inappropriate in a cross-sectional study.

I errors and because high correlations were found between NPO and AS (Pearson $r = .84$) and between the RPS and PPO subscales (Pearson $r = .68$), it was decided that interactions would only be analyzed with regard to NPO, RPS and IMP. Thus, six separate analyses were executed to test whether the interaction variables between daily stress and problem solving were significant. To prevent collinearity, interaction variables were centered. If the interaction variable turned out to be significant while the main variables were included in the model, this would indicate moderation. Collinearity (variance inflation factor > 10) and outliers (Cook's distance > 1 ; studentized residual < -3 or > 3) were checked at each step.

RESULTS

DEMOGRAPHIC VARIABLES

Eighty-eight male and thirty-four female patients entered the study ($N=122$). Mean age was 40 ($SD=9.02$) and mean sick leave duration was 59 days ($SD=52.75$), i.e., about 2 months. The current pain episode lasted an average of 89 weeks ($SD=234.02$), while the mean period of time elapsed since the first onset of the pain was 447 weeks ($SD=432$ weeks.). About 65% of the patients reported that this was not their first pain episode, implying that their pain was recurrent. In 70% of the cases, the current pain episode could be defined as chronic (pain duration >12 weeks), whereas in 25% it was sub-acute (4 weeks $<$ current pain episode ≤ 12 weeks). Only 5% of the patients reported the current pain episode to have lasted no longer than 4 weeks, but none of them were in their first episode. The mean RDQ-score was 13.2 ($SD= 4.96$), and 28% of the patients had an RDQ-score > 15 , suggesting that these patients were moderately to highly disabled (Patrick et al., 1995; Roland & Morris, 1983). Mean pain intensity was 17.8 ($SD= 9.1$), which is higher than that of a group of Dutch physiotherapy patients who reported a mean pain intensity of 11.6 ($SD= 7.0$) just before their treatment started (Van der Kloot et al., 1995). In sum, the population included in the present study can be characterized as mainly chronic, most of them suffering from recurrent pain episodes.

ARE PAIN INTENSITY, PAIN CATASTROPHIZING, PAIN-RELATED FEAR, DAILY STRESS, AND PROBLEM SOLVING INDIVIDUALLY RELATED TO FUNCTIONAL DISABILITY? (HYPOTHESIS 1)

Table 1 summarizes means, standard deviations, and Pearson correlation coefficients between functional disability (QBPDS), pain intensity (MPQ-DLV), pain catastrophizing (PCS), pain-related fear measures (TSK, FABQ-work, FABQ-act), daily stress (EPCL-freq, EPCL-int), and problem solving (SPSI-r: PPO, NPO, RPS, AS, and IMP). The highest correlations between functional disability and independent variables were found for pain intensity and pain catastrophizing. Other significant correlations were found with pain-related fear, daily stress, negative problem orientation and avoidance style. Table 2 displays the outcomes of multiple regression analyses for each block of independent vari-

TABLE 1. Means, Standard Deviations (*SD*), and Pearson Correlation Coefficients for Dependent and Independent Variables (*N*=122).

Variables	Mean	<i>SD</i>	2	3	4	5	6	7	8	9	10	11	12	13
Disability:														
1. QBQDS	41.47	15.3	.45***	.51***	.34***	.21*	.33***	.22*	.30***	-.01	.32***	.17	.08	.21*
Pain intensity:														
2. MPQ-DLV	17.82	9.12	-	.41***	.32***	.21*	.18*	.35***	.30***	.04	.17*	.18*	.05	.09
Pain catastrophizing:														
3. PCS	16.80	9.82			.62***	.10	.36***	.25**	.18*	-.09	.41***	.09	.07	.27**
Pain-related fear:														
4. TSK	37.02	7.52				.31***	.50***	.13	.07	-.14	.35***	.02	.11	.38***
5. FABQ-work	17.47	9.51					.44***	.04	.15	-.11	.19*	-.04	.01	.14
6. FABQ-act	10.62	4.96						.17	.11	-.09	.31***	.04	.09	.32***
Daily stress:														
7. EPCL: FREQ	25.59	19.50							.52***	.08	.25**	.13	.13	.05
8. EPCL: INT	26.14	31.45								.16	.13	.20*	-.09	-.08
Problem solving:														
9. SPSI-r: PPO	12.16	3.35									-.28**	.69***	-.14	-.37***
10. SPSI-r: NPO	9.02	5.75										.07	.37***	.68***
11. SPSI-r: RPS	41.65	12.69											-.14	.04
12. SPSI-r: IMP	12.65	4.78												.48***
13. SPSI-r: AS	6.39	3.89												

Note. QBQDS: Quebec Back Pain Disability Scale; MPQ-DV: McGill Pain Questionnaire-Dutch Version; PCS: Pain Catastrophizing Scale; TSK: Tampa Scale for Kinesiophobia; FABQ: Fear Avoidance Beliefs Questionnaire, work (work) and physical activity (act); EPCL: Everyday Problem Checklist, frequency (FREQ) and intensity (INT); SPSI: Social Problem Solving Inventory; PPO: Positive Problem Orientation; NPO: Negative Problem Orientation; RPS: Rational Problem Solving Skills; IMP: Impulsivity; AS: Avoidance Style.

* $p < .05$, ** $p < .01$, *** $p < .001$.

ables in analyses 1 to 5. The analyses showed that pain intensity as well as pain catastrophizing were strong predictors when analyzed separately. With regard to pain-related fear, FABQ-work was eliminated from the model, whereas TSK and FABQ-act both seemed to make an additional contribution to the model. Daily stress did add to the model but only when its intensity was included. With regard to problem solving, negative problem orientation (NPO) was the only independent variable significantly contributing to the model of functional disability. Days of sick leave did not make a significant contribution to any of the analyses. In those cases where sex and/or age made a significant contribution to the model, Beta values were small.

DOES PROBLEM SOLVING HAVE ADDITIONAL VALUE IN A MODEL EXPLAINING FUNCTIONAL DISABILITY? (HYPOTHESIS 2)

In table 3, 5 blocks of independent variables were added to the predictive model of functional disability in a stepwise procedure (steps 1 to 5). Table 3 shows that pain intensity, pain catastrophizing, and sex explained 34% of the variance. Pain catastrophizing explained an extra 11% of the variance when added to the model that already included pain intensity. In step 3, pain-rela-

TABLE 2. Summary of Multiple Regression Analysis of Functional Disability measured by the QBPDS.

<i>Analysis nr.</i>	<i>Adj R²</i>	<i>Independent variables</i>	<i>Standardized Beta</i>
1) Pain intensity	.23	Sex	.19*
		Age	.19*
		MPQ-DV	.46***
2) Pain catastrophizing	.28	Sex	.19*
		PCS	.51***
3) Pain-related fear	.19	Sex	.30**
		TSK	.28**
		FABQ-act	.22*
4) Daily stress	.10	Sex	.18
		EPCL-int	.29**
5) Problem solving	.11	SPSI-NPO	.28**
		SPSI-RPS	.15

Note. Covariates: Sex (1=m, 2=f), age, and days of sick leave. Five blocks of independent variables: Analysis 1: MPQ-DLV; Analysis 2: PCS; Analysis 3: TSK, FABQ-work, FABQ-act; Analysis 4: EPCL-freq, EPCL-int; Analysis 5: SPSI-PPO, NPO, RPS, IMP, AS. Variables with $p>0.1$ were eliminated from the analyses. QBPDS: Quebec Back Pain Disability Scale; MPQ-DV: McGill Pain Questionnaire-Dutch Version; PCS: Pain Catastrophizing Scale; TSK: Tampa Scale for Kinesiophobia; FABQ: Fear Avoidance Beliefs Questionnaire, work (work) and physical activity (act); EPCL: Everyday Problem Checklist, frequency (freq) and intensity (int); SPSI: Social Problem Solving Inventory; PPO: Positive Problem Orientation; NPO: Negative Problem Orientation; RPS: Rational Problem Solving Skills; IMP: Impulsivity; AS: Avoidance Style. * $p<.05$, ** $p<.01$, *** $p<.001$.

ted fear measures were added. TSK did not add to the model. Only FABQ-act turned out to be predictive, with an R^2 -increase of 2%. In step 4, the intensity of daily stress was added, but this did not substantially change the predictive value of the model (2%). Finally, problem solving was added, but none of the problem solving scales turned out to be significant, while the predictive value did not increase either. In short, it might be concluded that from step 2 onwards, the percentage of explained variance did not show any marked increase. When pain intensity and pain catastrophizing were included in the model, pain-related fear, daily stress, and problem solving did not add to the model of functional disability. Of the covariates that were added to the model, only sex turned out to be significant.

TABLE 3. Multiple Regression Analysis (STEPWISE) of Functional Disability measured by the QBPDS.

<i>Step nr.</i>	<i>Adj R2</i>	<i>Independent variables</i>	<i>Standardized Beta</i>
1	.23	Sex	.19*
		Age	.19*
		MPQ-DV	.46***
2	.34	Sex	.19*
		MPQ-DLV	.30**
		PCS	.38***
3	.36	Sex	.21**
		MPQ-DLV	.29**
		PCS	.33**
		TSK	-.02
4	.38	FABQ-act	.19*
		Sex	.20*
		MPQ-DLV	.25**
		PCS	.31***
		FABQ-act	.18*
5	.37	EPCL-int	.14
		Sex	.20*
		MPQ-DLV	.24**
		PCS	.30**
		FABQ-act	.17*
		EPCL-int	.13
		SPSI-NPO	.03

Note. Covariates: Sex (1=m, 2=f), age, and days of sick leave. In each step, a block of independent variables is added in the following order. Step 1: pain intensity; Step 2: pain catastrophizing; Step 3: pain-related fear; Step 4: daily stress; Step 5: problem solving. Variables with $p>0.1$ were eliminated from the analyses in the next step. QBPDS: Quebec Back Pain Disability Scale; MPQ-DV: McGill Pain Questionnaire-Dutch Version; PCS: Pain Catastrophizing Scale; TSK: Tampa Scale for Kinesiophobia; FABQ: Fear Avoidance Beliefs Questionnaire, physical activity (act); EPCL: Everyday Problem Checklist, intensity (int); SPSI: Social Problem Solving Inventory; NPO: Negative Problem Orientation; RPS: Rational Problem Solving Skills; * $p<.05$, ** $p<.01$, *** $p<.001$.

DOES PROBLEM SOLVING MODERATE THE EFFECTS OF DAILY STRESS ON FUNCTIONAL DISABILITY? (HYPOTHESIS 3)

Six separate analyses were executed to test if there was a significant interaction between problem solving (NPO, RPS, IMP) and daily stress (FREQ, INT). None of the interaction variables were significant ($p > 0.05$), indicating that problem solving did not moderate the effects of daily stress on functional disability.

DISCUSSION

It can be concluded that pain intensity, pain catastrophizing, pain-related fear, daily stress, and problem solving are individually related to functional disability, but that this is most convincingly the case for pain intensity and pain catastrophizing. With regard to problem solving, only negative problem orientation (NPO) had a significant relationship with functional disability. There was no evidence that problem solving added to the predictive model of functional disability that included pain intensity and pain catastrophizing. Similarly, the additional value of pain-related fear and daily stress was trivial when pain and pain catastrophizing were controlled for. Finally, there was no evidence to support the hypothesis that problem solving moderates the effect of daily stress on functional disability.

In contrast to the study by Crombez et al. (1999), which concluded that pain-related fear is more disabling than pain itself, the current study found pain intensity to be a more important predictor, in addition to and regardless of pain-related fear. There might be some explanations for these findings. First, Crombez et al. (1999) studied chronic low back pain patients of whom the majority was not in regular employment and who were receiving a disability pension. Patients in the current study were all employed and had been sick-listed for no longer than 20 weeks. For employees, pain intensity may be an important criterion, on a daily basis, to decide whether or not to resume functions and activities. For those patients who have stopped working a long time ago and are suffering chronic pain, pain intensity may not be a very reliable criterion to decide whether daily activities can be increased or not. A second difference between the two studies is the way pain intensity was operationalized. The study by Crombez et al. (1999) used a Visual Analogue Scale (VAS), whereas the current study used the MPQ-DLV. The latter, measuring both pain intensity and pain quality, may be more sensitive to small differences than the VAS. The fact that pain-related fear added to the model of functional disability is in line with the findings by Crombez et al. (1999). However, it must be noted that the model presented in table 3 (step 3) explained 36% of the variance, which is considerably lower than the percentage reported in earlier studies which included pain-related fear and pain intensity to construct a model of functional disability (Crombez et al., 1999). Again it may be questioned what influence the nature of the population in the current study had. As described

earlier, patients in the current study all managed to continue in their job despite pain, and none of them suffered from severe psychopathology. Therefore, compared to the study by Crombez et al. (1999), the current sample might be relatively well adjusted and the influence of pain-related fear might be less prominent.

Problem solving did not convincingly contribute to a predictive model of functional disability. This is not what we had expected to find on the basis of the studies described in the introduction (Aldrich et al., 2000; Nezu et al., 1998). There are some plausible explanations for the fact that rational problem solving skills (RPS) were not associated with disability. First, the SPSI-R describes skills to solve problems in general, not to solve more specific functional problems that result from pain, like getting out of bed or lifting and carrying a heavy suitcase (as measured with QBPDS; (Kopeck et al., 1995)). Secondly, as Aldrich et al. (2000) suggested, ineffective and misdirected coping leads to chronification of pain. They hypothesized that those most disabled by chronic pain might possess problem-solving skills, but apply them repeatedly to an insoluble problem. According to this hypothesis, problem definition and formulation in particular may be crucial, not rational problem solving skills in general.

With regard to negative problem orientation (NPO), high correlations with pain catastrophizing (PCS) and pain-related fear (TSK), as well as high partial correlations, might leave less room for NPO to express its unique relationship with functional disability. Moreover, there is some content overlap between NPO, PCS and TSK, in that they share a negative appraisal of events. NPO describes a negative orientation towards problems in general, whereas PCS and TSK describe a negative orientation towards pain in particular, which may be more easily recognized by a pain patient as applying to himself. There was no evidence to support the hypothesis that problem solving moderates the effect of daily stress on disability. As our analyses were analogous to those used in the studies by Nezu and colleagues (1988), it may be surprising that none of our hypotheses were supported. An important difference with the studies mentioned above is the dependent variable: we studied the effect on functional disability (the performance of specific behaviors), as representing quality of life, instead of depression (psychological and cognitive processes), which could have hampered the generalization of the foregoing to a model of pain.

There are a number of caveats to be considered. Because the study was cross-sectional in nature, it does not permit conclusions to be drawn about the time-sequence of the relationships found. With regard to problem solving, the sequence of its relationship with functional disability remains unclear. It may be true that ineffective problem solving leads to stress-related avoidance and disability in the long term. On the other hand, disability may result in negative thinking about one's capacity to solve a problem since physical functions are invalidated. Prospective or experimental designs would be needed to ans-

wer questions about the role of problem solving in pain and disability. What are the consequences of the current findings for the management of patients with sub-acute low back pain? It was found that, besides pain intensity, pain catastrophizing is a variable that has a solid relationship with disability. This finding may be important for treatment and prevention. Challenging catastrophic and irrational thoughts about pain and its consequences may likewise be a promising method and may prevent disability in the longer term (Vlaeyen, De Jong, Geilen, Heuts, & Van Breukelen, 2001). Finally, catastrophic thoughts may be prevented, by controlling the sources by which patients derive information about pain. Non-ambiguous and consistent information about back pain may help to give the patient a more clear and rational image of the pain problem he or she is confronted with.

As for problem solving, research needs to be done to examine its possibilities and consequences with regard to pain problems in the longer term. The idea that a positive problem orientation and proper problem solving skills can prevent chronicity in acute and sub-acute pain patients is still appealing and plausible. However, the current explorative study provides a warning against premature conclusions. The role of problem solving in disability may be far more complicated than has been assumed so far, which raises new questions to be dealt with in future research.

CHAPTER 6

Has problem solving therapy supplemental
value when added to behavioral graded activity
in non-specific low back pain patients?
A randomized clinical trial.

A shortened version of this chapter is under review as:

Van den Hout, J.H.C., Vlaeyen, J.W.S., Heuts, P.H.T.G., De Vet, H.C.W., Sillen, W.J.T., Willen, A.J.E.H.L., Wijnen, J.A.G., Passchier, J.. Has problem solving therapy supplemental value when added to behavioral graded activity in non-specific low back pain patients? A randomized clinical trial. *Clinical Journal of Pain*.

ABSTRACT

The aim of this study was to investigate the supplemental value of problem solving therapy (PST) when added to behavioral graded activity in non-specific low back pain patients in the improvement of their functional status. One hundred and fifteen patients were randomly assigned to either graded activity + problem solving therapy (GAPS) or graded activity + group education (GAGE). Both GAPS and GAGE conditions resulted in significant improvements with regard to functional disability, pain intensity, pain-related fear, pain behavior and depression at post treatment. Improvements were maintained at 6- and 12-month follow-up. Although modest, patients in the GAPS condition showed significant less disability than patients in the GAGE condition. The GAPS condition also included the highest proportion of patients yielding a clinically relevant decrease of functional disability at 6- and 12-month follow-up. Intention-to-treat and per-protocol analyses generally confirmed previous findings. The results suggest that cognitive PST might be a valuable adjunct to behavioral graded activity programs.

INTRODUCTION

Chronic low back pain (CLBP) has wide consequences on a personal as well as a socio-economic level. Pain threatens personal health, affecting physical function, independence, emotional wellbeing, and family life. Furthermore, in employed patients, pain might threaten patient's work-status and consequently, economic welfare of the society in general. The economical consequences of back pain disability have shown an increasing impact for society in the last two decades (van Tulder, Koes, & Bouter, 1995). Back pain is one of the most expensive illnesses of the western world (OECD, 1996). This is particularly caused by indirect costs resulting from sick leave and disability pensions (Meerding, Bonneux, Polder, Koopmanschap, & van der Maas, 1998). Spitzer, LeBlanc and Dupuis (1987) found that, although most workers with compensated back injury (74%) return to work within 4 weeks, only half of the remaining patients (13%) returns to work within 12 weeks. Furthermore, half of the patients absent from work for longer than 7 weeks, will not return to work within 6 months (Spitzer, LeBlanc, & Dupuis, 1987). When a pain problem develops into a chronic state, potential to reverse of the process decreases, and therapy rarely reports large improvements anymore. Prevention of chronicity in an early stage has therefore aroused the interest of health care providers, researchers and policymakers.

To give new directions to the management of pain and pain disability, it is suggested to consider a multidimensional process in which psychosocial factors and factors related to work (i.e. job task enjoyment) play an important role (Bigos et al., 1991; Burton & Tillotson, 1991; Gallagher et al., 1989; Gatchel & Turk, 1996). There are strong indications that fear-avoidance, referring to the avoidance of movements and activities based on fear, is important with regard to the prevention of a long-term back pain problem (Crombez, Vlaeyen, Heuts, & Lysens, 1999; Klenerman et al., 1995; Vlaeyen & Linton, 2000). Kendall, Linton and Main (1997) therefore included pain-related fear as an essential aspect of a broader early assessment of the so-called psychosocial 'yellow flags', a screening instrument to select back pain patients at risk for chronicity in primary care. In line with this, pain treatments have adopted the fear-avoidance model and implemented cognitive-behavioral techniques. The operant approach, i.e. graded activity or graded exposure, as originally introduced by Fordyce (1976), has been successfully applied in both chronic pain patients (Kole Snijders et al., 1999; Nicholas, Wilson, & Goyen, 1991; Turner & Clancy, 1988; Turner, Clancy, McQuade, & Cardenas, 1990; Turner & Jensen, 1993) and sub-acute and acute low back pain patients (Fordyce, Brockway, Bergman, & Spengler, 1986; Lindström et al., 1992; Linton & Bradley, 1992; Linton, Bradley, Jensen, Spangfort, & et al., 1989). Lindström and colleagues (1992) found that in patients with subacute low back pain, a graded activity program is more effective than usual care regarding return to work. Moreover, in their systematic review, Van Tulder and colleagues (2000) concluded that there is strong

evidence that behavioral treatment has small to moderate positive effects in chronic back pain patients when compared to no treatment, placebo or waiting list control conditions. However, there was also moderate evidence that addition of behavioral treatment to another treatment did not have an extra value, neither in the short, nor in the longer term (van Tulder et al., 2000). Two critical remarks may be made with respect to the efficacy of cognitive behavioral treatments: First, effects of behavioral treatment that are found immediately after treatment may not always be maintained in the longer term. Secondly, interventions applied at an early stage, e.g. between 2 and 6 months duration of sick leave, show better results in the short and the longer term as compared to interventions in pain patients with long-term work disability (Marhold, Linton, & Melin, 2001).

Operant-behavioral treatments, like graded activity, intervene in the process of activity-related avoidance and pain-related fear (Fordyce, 1989; Lethem, Slade, Troup, & Bentley, 1983; Philips, 1987; Vlaeyen & Linton, 2000; Waddell, Newton, Henderson, Somerville, & Main, 1993). However, pain behavior not only allows for the avoidance of physical and painful activities, but also, pain behavior might legitimate the avoidance of situations that are stressful and for which the patient does not find acceptable solutions. In this context, sick leave because of back pain complaints can be considered a medical solution of a work-related problem (Knepper & Feenstra, 1991). In order to improve treatment effects in the longer term and to go into the role of stress-related avoidance as well, problem-solving strategies (D'Zurilla & Goldfried, 1971) were suggested to add to the overall effectiveness of pain treatments. There are several reasons that make the application of PST in pain patients appealing. First, problem-solving skills would offer a coping style that could be of use when pain patients are confronted with daily hassles and stress. Being able to come up with some effective solutions, for instance when pain relapses, might enable a patient to stay at work despite problems resulting from pain. Secondly, as stated by Alldrich, Eccleston and Crombez (2000), in most cases of chronic pain, interventions should focus at the consequences of pain in everyday life rather than the pain itself. As stated before, consequences of chronic pain cover a broad field of problems and more generic skills than those that focus only on pain might be needed to cope with them. Problem solving therapy (Nezu, Nezu, Friedman, Faddis, & Houts, 1998; Nezu & Perri, 1989) is designed to be applied in all kinds of problems, whether these are trivial or complex, and therefore might be useful in those diseases where everyday functioning and wellbeing are at stake. Third, PST is known to be a particularly appropriate treatment for stress-related disorders such as generalized anxiety, psycho-physiological disorders (e.g. tension headaches) and depression because of the possible critical role of perceived uncontrollability in the etiology and/or maintenance of these disorders (D' Zurilla, 1988). Evidence confirming the effectiveness of PST, was especially found with regard to the treatment of depression (Arean et al., 1993; Hussian & Lawrence, 1981; Mynors

Wallis, Gath, Lloyd Thomas, & Tomlinson, 1995; Nezu, 1986; Nezu & Perri, 1989). Finally, because skills are trained to cope with "the normal ebb and flow of living life" (D'Zurilla & Goldfried, 1971), PST is expected to contribute to the generalization of skills to everyday life, and the maintenance of behavior change over a longer period of time.

Next to problem-solving skills, the therapy might increase a positive orientation towards daily problems as well. Problem orientation is a motivational process, that can be described as a set of orienting responses that represent the immediate cognitive-affective reactions of a person when first confronted with a problem (Nezu et al., 1998). Positive problem orientation might result in reduced negative affect, avoidance behavior and disability (D'Zurilla & Nezu, 1982). Conversely, negative problem orientation has shown to increase negative affect, impulsive behavior, and avoidance motivation, which in turn might inhibit or disrupt subsequent problem-solving attempts (Nezu et al., 1998). Furthermore, patients with a negative orientation tend to worry and complain about their health (Dugas, Letarte, Rheame, Freeston, & et al., 1995; Godshall & Elliott, 1997). In back pain, negative problem orientation is positively related to functional disability (Van den Hout, Vlaeyen, Heuts, Sillen, & Willen, 2001b). Corroborating these findings, Shaw, Feuerstein, Haufler, Berkowitz & Lopez (2001) found low back pain symptoms producing more functional loss when positive problem orientation was low and impulsiveness and avoidant style were high. The authors suggest that the prolonged impact of low back pain on daily functioning may be reduced by assisting workers to conceptualize LBP and associated problems as hindrances that can be overcome, and to use active strategies in reducing risks for low back pain disability.

Although the effects of PST have been studied with regard to several populations within the field of behavioral medicine, application in pain populations is scarce. Nevertheless, there are some studies that suggest that the treatment might be very well applicable in this area. First, Linton and colleagues (Linton & Bradley, 1992; Linton et al., 1989) evaluated the effects of a multi-modal program including the training of problem solving strategies in nurses at risk for developing a chronic pain problem. The authors found a significant improvement with regard to pain intensity, mood, sleep, pain behaviors, activity level and days of sick leave at 6-month follow-up. The multi-modality of the program prevented to detect which treatment ingredients were responsible for the improvement. Philips, Grant and Berkowitz (1991) added a problem solving-like counseling session to graded exposure in acute low back pain patients. At 6-month follow-up, the counseling approach showed a trend towards an earlier return to normal activities and reduced the likelihood of persisting pain difficulties as compared to the same approach without counseling. It was suggested that it would be worthwhile to elaborate counseling sessions in future research. Finally, Wilkinson and Mynors-Wallis (1994) found PST to be feasible and acceptable in a pilot study regarding 11 patients with unexplained physical symptoms. In conclusion, the studies above all indicate that PST could

be a valuable addition in the treatment of pain patients in an early phase, however, none of the previous studies investigated PST as an isolated ingredient.

So far, little is known about the necessary ingredients for effective cognitive-behavioral treatments. Component analyses, such as the studies by Turner et al. (1990), Vlaeyen et al. (1996) and Kole-Snijders et al (1999) are scarce, but badly needed in order to isolate the active components of pain rehabilitation programs, which inevitably must lead to more efficient use of health care resources. The present study is designed as an attention-controlled randomized clinical trial that is aimed at examining the supplemental value of PST when added to graded activity in employees who have a recent new episode of sick leave as a result of low back pain, in terms of functional disability, medical consumption, depression and self-reported days of sick leave. Because PST may require a certain level of abstract thinking, it is further hypothesized that treatment effect is moderated by education level, in that high-educated patients gain most from supplemental problem solving therapy, whereas low-educated patients scarcely gain from this addition. The study also includes a long-term follow up regarding days of sick leave based on data provided by occupational health services, that is described elsewhere (Van den Hout, Vlaeyen, Heuts, Zijlema, & Wijnen, accepted pending vision).

METHODS

PARTICIPANTS

Patients were all employees referred to the study by General Practitioners, Occupational Physicians or Rehabilitation Physicians. Eligibility criteria were as follows: age between 18 and 65 years, low back pain (LBP) for more than 6 weeks, on sick leave with LBP but no longer than 20 weeks, no more than 120 days of sick leave during the last year. Non-eligible were those patients who had specific back complaints (vertebral fracture, infectious disease, rheumatoid arthritis, ankylosing spondylitis, or herniated disc; Waddell & Turk, 1992), were pregnant, were not proficient in Dutch, or were in consultation and/or treated by a medical specialist for diagnosis or treatment of their LBP at the time of referral. Furthermore, patients had to agree that they would stop other therapies they were receiving for their back complaints at the start of the intervention.

During the selection procedure the Dutch version of the Symptom Check List 90 (SCL-90; Derogatis, Lipman, & Covi, 1973; Arrindell & Ettema, 1986) and the Dutch Personality Questionnaire (NPV; Luteijn, Starren, & Van Dijk, 1985) were used. Criteria from a previous study in CLBP/fibromyalgia patients were adopted (Vlaeyen et al., 1996) to check for: i) psychopathology that would hamper the rehabilitation process, and ii) disturb group processes¹. Patients with medical comorbidity were excluded when the disorder interfered with one of both treatment programs or rendered them unable to participate in

Note 1 Criteria are described in chapter 5 of this thesis, pg.101.

every part of the program, as decided by the rehabilitation physician. Patients, who currently were involved in litigation concerning their ability to work, were excluded as well. During the screening procedure, demographic variables and the Distress and Risk Assessment Method (DRAM; Main, Wood, Hollis, Spanswick, & Waddell, 1992) were assessed.

In the period between September 1996 and December 1998, 138 patients were assessed for eligibility for the trial. Twenty-three patients did not meet the inclusion criteria and were excluded from the study. The remaining 115 patients all gave informed consent to participate in the study and were randomly assigned to one of two treatment conditions that are compared in this study.

STUDY DESIGN

The study design was a randomized clinical study. The randomization scheme was computer-generated and only known to the logistic planner of the rehabilitation center. The selection procedure was carried out by a rehabilitation physician (P.H.) and a mental health scientist (J.H.), who were both blind for the allocated condition. After patients had been included and signed an informed consent, they were assigned to one of two treatment conditions in blocks of at maximum 5 patients.

Two treatment conditions were compared, the first condition being Graded Activity plus Problem Solving Therapy (GAPS), and the second condition being Graded Activity plus Group Education (GAGE). Group Education was chosen as an attention-control in that non-specific factors were equally provided in GAPS and GAGE, but for keeping contrast, explicitly no skills were trained in the latter. Measurements took place pre-treatment, post-treatment, and at 6- and 12-month follow-up. When patients were not able to attend the measurements, a package with self-report measures and instructions was sent to the patient's address. Consequently, only self-report measures could be obtained when patients filled out the measures at home. The research assistants responsible for data-collection, behavioral observations, and data-entry, were blind for the allocated condition. The same therapists in both treatment conditions guided the graded activity program. Therapists were not blind for the condition, because multidisciplinary consultation was part of the treatment program, so that all therapists were aware of the therapy contents. To avoid contamination, GAPS- and GAGE-groups had their program planned separately such that they did not encounter one another.

TREATMENTS

The treatment program consisted of 19 half-day-sessions that were spread over 8 weeks, with decreasing intensity from 3 to 2 sessions a week. The treatment was given in small groups of at most 5 patients. The program started with an introduction of the treatment rationale. In the course of the program the therapists' team had three meetings with the individual patients. During these

meetings, progresses and hindrances for goal achievement and return to work were discussed. Two months after the final treatment session, a booster session was planned in which the treatment contents was summarized, and individual developments were discussed in the group. In both treatment conditions, patients received a workbook in which specific treatment information, handouts, graded activity graphs, and homework assignments could be collected. There were no external differences between workbooks in each condition.

Graded Activity (GA)

Graded activity is an operant behavioral treatment based on the work of Fordyce (1976) that aims at increasing activity levels and restoring occupational function using quota systems. Included in the training were registration of baseline levels during the first two weeks, treatment contract, positive reinforcement for activity increments, and a workplace visit (Fordyce, 1976). The physiotherapist trained a circuit of 5 basic exercises (bicycling, lateral pulley, steps, shoulder and abdominal exercises) in 15 one-hour sessions. Three additional sessions were dedicated to back education and lifting instructions. The occupational therapist trained patients for 30 minutes per week on an individual basis, in which principles of graded activity were applied in personal relevant activities like work, housekeeping, or leisure activities. Furthermore, the occupational therapist contacted the occupational physician and the patient's direct supervisor at the workplace, to plan a workplace-visit when necessary with all persons involved. Patients were encouraged to return to their previous work place.

Problem Solving Therapy (PST)

Problem solving therapy is a cognitive behavioral therapy in which problem-solving skills are trained according to the theory as designed by D'Zurilla & Goldfried (1971). The theory describes 5 steps in which problems are typically solved: problem orientation, problem definition and formulation, generation of alternatives, decision making, implementation and evaluation. In the current trial a protocol as originally designed by Nezu and colleagues (Nezu, 1986; Nezu & Perri, 1989) was translated and applied as a group intervention. Skills were trained in 10 ninety-minute sessions and provided by 2 experienced behavior therapists (4 behavior therapists were available during the study). A detailed session-to-session therapist's manual was used in order to make sure that treatment protocols were strictly followed. Moreover, there were regular consultations with the therapists and the project's supervisor, to check therapist's adherence to the treatment protocol. Skills training and application were the focus of the therapy, rather than one specific problem area. Patients were free to select their own problem areas, which did not need to be pain-related. Between sessions, homework assignments were given to practice skills in everyday life. Every session homework assignments were discussed in the group,

and patients were encouraged to keep on practicing and applying problem solving skills in everyday life.

Group Education (EDU)

An attention-control intervention was designed in order to make both conditions comparable with regard to non-specific factors. Since educational approaches turned out to have no effects on worry, symptoms, functional status, health care utilization or other primary outcome measures, at least not in the long term (Berwick, Budman, & Feldstein, 1989; Cherkin, Deyo, Street, Hunt, & Barlow, 1996), this intervention was chosen as the attention-control. Group education consisted of 10 lessons of 90 minutes, in which issues related to the back and to back pain were discussed. Physiotherapist, occupational therapist, and/or behavior therapist, using a protocolized manual, served as lecturers. The following themes were discussed: physiology of the healthy back; development of low back pain, posture and lifting instructions, pain, healthy activity, stress and relaxation, fatigue, and pain and the environment. Explicitly, no skills were trained and each theme was discussed during no more than one protocolized session.

OUTCOME MEASURES

In this study self-report measures, a cost diary, and observational measures are included to assess the effectiveness of both interventions. To check if problem-solving skills were exclusively manipulated in the GAPS condition, a social problem solving skills inventory was administered.

Primary outcome measures

Functional disability. The Dutch Version of the Roland Disability Questionnaire (RDQ; Roland & Morris, 1983) and the Quebec Back Pain Disability Scale (QBPDS; Kopec et al., 1995) were included to measure functional disability. The RDQ is a 24-item 2-point scale assessing to what extent the performance of daily activities is hampered by back pain. Reliability, responsiveness and validity of the RDQ are satisfactory (Beurskens, de Vet, & Koke, 1996). The QBPDS is a 20-item 6-point self-report scale describing activities commonly affected by back pain. Subjects rate the level of difficulty executing these activities. The QBPDS is a reliable, valid, and responsive outcome measure (Gommans, Koes, & Van Tulder, 1996; Kopec et al., 1995; Schoppink, van Tulder, Koes, Beurskens, & de Bie, 1996). An advantage of the QBPDS as compared to the RDQ may be that the QBPDS has been specifically designed for patients with low back pain by selection and reduction of items, whereas the RDQ has been derived from the more generic Sickness Impact Profile (Bergner, Bobbitt, Carter, & Gilson, 1981).

Physical (health) activities, health care utilization, and days of sickleave. A cost-diary as designed by Goossens, Rutten- van Mólken, Vlaeyen, and Van der Linden (2000), was included to obtain data on physical (health) activities,

health care utilization, and days of sickleave. Patients were instructed to record on a weekly basis, activities, direct health care consumption, direct non-health care consumption, and indirect costs due to back pain. The cost-diary turned out to be a feasible and valid instrument to obtain data on medical consumption and absenteeism in chronic pain patients (Goossens et al., 2000). Patients kept diaries during 52 weeks post-treatment, starting one week after the intervention had finished. In this study, the following variables will be used for comparison between the conditions: physical (health) activities; general practitioner contacts; physiotherapy contacts; specialist care contacts; alternative medicine contacts; days of sickleave.

Secondary outcome measures

Pain intensity. Pain intensity was measured by the total Pain Rating Index (PRI) of the McGill Pain Questionnaire (MPQ; Melzack, 1975; van der Kloot, Oostendorp, van der Meij, & van den Heuvel, 1995), a frequently used and reliable measure of pain intensity. The MPQ-DLV consists of 20 groups of 3 or 4 words describing pain qualities. The subject can choose those words that fit in best with his/her pain. The PRI is calculated by adding the individual values of the words marked.

Pain Catastrophizing. The Pain Catastrophizing Scale (PCS; Sullivan, Bishop, & Pivik, 1995; Crombez & Vlaeyen, 1996) was applied to measure catastrophizing thoughts about pain. The PCS is a 13-item 5-point scale that has shown satisfactory internal consistency for the total scale (Cronbach $\alpha = .93$; Osman et al., 1997), as well as good temporal stability (Pearsons $r = .92$; Crombez, Vervaeke, Lysens, Baeyens, & Eelen, 1998). The total score was used in the present study.

Fear of movement/(re)injury. Pain-related fear was assessed by the Dutch version of the Tampa Scale for Kinesiophobia (TSK; Kori, Miller, & Todd, 1990). The TSK is a 17-item 4-point scale that measures the fear of injury or re-injury due to movement. The Dutch version has been shown to possess good reliability and validity (Goubert et al., 2000; Vlaeyen, Kole Snijders, Boeren, & van Eek, 1995b).

Fear avoidance beliefs. Fear avoidance beliefs were measured by the Dutch version of the Fear Avoidance Beliefs Questionnaire (FABQ; Waddell et al., 1993). The FABQ consists of two factors and focuses specifically on patients' beliefs about how physical activity (FABQ-act) and work (FABQ-work) affect their low back pain. Test-retest reproducibility and internal consistency of the FABQ have turned out to be satisfactory (Waddell et al., 1993).

Depression. The Beck Depression Inventory (BDI; Beck, Ward, & Mendelson, 1961), a 21-item self-report measure, was conducted to measure depressive symptoms. The BDI is widely used in psychotherapy research and has shown to be reliable and valid (Beck, Steer, & Garbin, 1988).

Pain behavior. A research assistant who was blind for the condition observed pain behaviors during a Behavioral Approach Test (BAT). During the BAT

(Kole-Snijders, Vlaeyen, van Eek, Schuerman, & Groenman, 1990) patients were asked to walk and stop if pain prevented them from continuing, up to a preset maximum of 7 minutes. Walking distance was recorded. Furthermore, pain behaviors were observed by means of the Pain Behavior Scale (PBS; Richards, Nepomuceno, Riles, & Suer, 1982; Vlaeyen et al., 1990). Reliability of the PBS turned out to be satisfactory (Kole Snijders et al., 1999).

Risk assessment. For descriptive purposes, the Distress and Risk Assessment Method (DRAM; Main et al., 1992) was administered. The DRAM classifies patients in 4 subclasses: normal, at risk, distressed-depressive, and distressed-somatic, in order to identify distress and evaluate the risk of poor outcome.

Manipulation check

Problem solving skills. Problem solving was measured by the Dutch translation of the Social Problem Solving Inventory-Revised (SPSI-R; D'Zurilla, Nezu, & Maydeu-Olivares, 1994). The SPSI-R is a 52-item self-report measure of social problem solving processes. The SPSI-R consists of five scales measuring two different problem-orientation dimensions (positive problem orientation (PPO; range: 0 - 15) and negative problem orientation (NPO; range: 0 - 30)), and three problem-solving specific dimensions (rational problem solving (RPS; range: 0 - 60), impulsive/carelessness style (IMP; range: 0 - 30) and avoidance style (AS; range: 0 - 21)).

Compliance

To check whether patients equally attended the treatment conditions, attendance rate of separate treatment components (GA, PST, EDU, and booster session) was registered. With regard to workplace visit, it was registered whether or not the visit actually took place at the workplace. If the workplace visit had been canceled, the work situation was discussed with the patient at the rehabilitation center instead. In the GAPS condition, compliance to homework assignments was checked per session. Participants received written homework assignments, 21 in total, which were checked by the therapists. The number of completed assignments was used as an additional compliance measure.

Treatment credibility

In order to check whether GAPS and GAGE conditions were equally credible to participants, treatment credibility was measured according to a procedure similar to the one described by Borkovec and Nau (1972; Devilly & Borkovec, 2000) and applied in Vlaeyen et al. (1996). After the introduction of the complete treatment program, and after the first session of either PST or EDU, patients rated on a Visual Analogue Scale (VAS), to what extent they expected that treatment would help them with regard to return to work, the resumption of daily activities, and to cope better with their pain. The VAS ranged from "not at all" (0) to "very much" (100). The VAS on treatment credibility was rated once more at the end of treatment to check whether treatment credibility changed.

STATISTICAL ANALYSES.

Initial analyses

Dropout-rate will be described per condition, at each separate measurement point in time. Furthermore, the influence of condition on dropout will be analyzed by means of logistic regression. To that purpose, attrition will be coded as a dichotomous outcome measure.

To check whether randomization succeeded, demographic variables and outcome measures at pretest are compared between conditions. Variables that are unequally distributed will be added to the regression equation as a covariate.

To test whether problem-solving skills were actually manipulated in the GAPS condition, pre-post differences on the SPSSI-r subscales are tested by means of paired t-tests within each condition separately. Furthermore, between group differences are compared by means of multiple regression analyses.

Effect evaluation

Selfreport and observational measures. In order to test within-group changes, paired t-tests are executed contrasting post-treatment, 6-month follow-up and 12-month follow-up measurement with pre-treatment, post-treatment and 6-month follow-up, respectively.

Multiple linear regression analyses are executed in order to test the hypothesis that PST has supplemental value when added to GA. Between-group differences are tested with regard to problem solving skills (manipulation check), functional disability, and depression, at post-test and 6- and 12-month follow-up. A stepwise hierarchical backward elimination method is used to eliminate non-significant predictors of outcome. The initial regression model includes the following independent variables: pretreatment measurement of the independent variable, treatment condition (GAPS=1; GAGE=2), age, gender (male=1; female=2), pain duration of the current pain episode (in weeks), sickleave duration prior to intake (in days), education level (low/medium=1; high=2), variables that turned out to be unequally divided between treatment conditions (covariates), and the interaction variable (treatment condition x education level). Pretreatment measurement, treatment condition, age, gender, and covariates always remain in the regression equation. In case the interaction variable is significant ($p < 0.05$), analyses with regard to the concerning outcome measure, are repeated within each education level separately. Non-significant interaction terms are eliminated from the model. Next, non-significant predictors ($p \geq 0.1$), concerning education level, current pain duration and sickleave duration, are eliminated one by one. At each step of the analysis, tests are done to test for high collinearity (Variance Inflation Factor > 10) and/or outliers (Cook's distance > 1; Studentized Residual < -3, or > 3). When a case is identified as an outlier, it is removed from the analysis. Finally, linearity of the dependent-independent relationship, and normality of the residuals is controlled for by means of scatter- and normal probability plots.

Cost-diary. Results of the cost-diary are split in two 26-weeks time periods corresponding with the 6- and 12-month follow-up measurements. In case diaries are missing, the mean score of the valid weeks in that particular period is imputed. A priori, it is decided that at least 18 out of 26 weeks have to be available in order to include the data into the analysis. Comparisons between the conditions are made with regard to the proportion of patients that reported activities, medical consumption or sickleave during the 26-weeks periods, and with regard to the absolute number of activities, health care visits and days of sickleave patients registered. Where appropriate, either Chi-square or Mann-Whitney *U*-tests will be used, rather than parametric *t*-tests, when the assumption of normality is violated.

Intention-to-treat analysis

First, in order to prevent bias from participants not completing the trial, an intention-to treat analysis is conducted. Intention-to-treat means that all the study participants are included in the analyses as part of the groups to which they were randomized regardless of whether they completed the study or not (Jadad, 1998). Second, missing data are imputed according to preset rules: 1) In case other data are available regarding the particular measurement (posttest, 6- or 12-month follow-up) the SPSS missing value analysis module is used. 2) In the absence of information from simultaneously conducted outcome measures, sickleave registration available from the occupational health services is used. It is assumed that patients, who drop out of the study but have fully resumed their jobs at the time of the measurement, have most favorable outcomes. In that case, the mean score of the most favorable quartile of all patients at the respective measurement is imputed. Those patients who drop out and have not or only partly resumed their jobs, are assumed to have least favorable outcomes, and the mean score of the least favorable quartile of all patients at the respective measurement is imputed. 3) Finally, when no information is available at all, neither from simultaneously conducted measures nor from sickleave registration, missing patients are classified on basis of their pretest scores and by ratio, and in concordance with their classification, given most favorable or least favorable outcomes. The ratio is computed within each condition on basis of the available sickleave data. The data-file resulting from this procedure is analyzed with the same statistical procedure as described above. In a sensitivity analysis we compare the results of the intention-to-treat analysis with those of the initial effect analysis.

Per-protocol analysis

In this procedure only those patients are analyzed who sufficiently participated in the protocolized treatment sessions. In the current study we decided that patients had to attend at least 66% of GA-sessions, and 8 sessions of PST or EDU. Analyses are repeated with the same statistical procedure as described above, and results of the per-protocol analysis are compared to those of the initial effect analysis.

Clinical relevance

In order to check whether the improvements with regard to functional disability (RDQ) are of clinical significance, criteria set by Stratford, Binkley, Riddle, and Guyatt (1998) are applied. In the absence of comparable studies with regard to the QBPDS, clinical relevance can only be calculated with respect to the RDQ. Change-scores are calculated by distracting post-treatment, 6- and 12-month follow-up RDQ-scores, respectively, from pre-treatment RDQ-scores. Because a clinically relevant change depends on initial RDQ-scores (Stratford et al., 1998), criteria are related to RDQ-scores at pre-treatment: higher initial RDQ-scores demand higher change-scores. The RDQ-range (0-24), was divided into 6 equal classes (0-4; 5-8; 9-12; 13-16; 17-20; 21-24), and minimum change-scores were defined per class (2; 2; 4; 5; 8; 8; Stratford et al., 1998). The proportion of improved patients is compared between the 2 conditions by means of a Chi-square test.

Process evaluation

After the 9th session of either PST or EDU, a process evaluation form was handed out to the participants to fill out and bring with them to the last therapy session. The evaluation consisted of 4 open questions on therapy content (e.g. "What was the most important you learned in this training?"), one question in which the patient could value the intervention on a 10-point rating scale; and one question to indicate the level of improvement on a 4-point Likert scale. Secondly, positively or negatively formulated statements concerning: treatment expectations and the achievement of treatment goals (e.g. "I gained more control over my life as a result of the training"), treatment process (e.g. "the time invested in the training was adequate"), working in a group (e.g. "I would have preferred to do this training individually"), therapists (e.g. "therapists showed adequate commitment to their patients"), and organization and facilities (e.g. "the training was well organized"), were assessed on a 5-point Likert scale (1="completely disagree"; 5="completely agree"). In the GAPS condition, 2 clusters were added on homework assignments (e.g. "homework assignments helped me to apply problem solving skills in practice"), and problem solving skills (e.g. "the training helped me to see daily hassles from another angle"). Answers to the open questions were clustered and the number of answers within each cluster counted. Next, answers were compared between conditions. Ratings of the intervention on the 10-point rating scale and 4-point Likert scale were compared by means of *t*-test and Chi-square test respectively. Finally, the percentage of statements agreed with, disagreed with, or neither (neutral), within each cluster of statements (after reflecting the negatively formulated statements) was compared between GAPS and GAGE conditions. The higher the percentage of statements agreed with, the more positive the particular part of the therapy process was thought of.

RESULTS

STUDY POPULATION

The final sample included in the study, consisted of 115 patients, 84 males and 31 females, with a mean age of 39.8 years (*SD* = 8.9, range = 21-55). On the average pain duration since the first pain episode was 7.8 years (*SD* = 8.4, range = 5.6 weeks – 40.3 years). The current pain episode had a mean duration of 1.8 years (*SD* = 5.0, range = 3 days – 40.3 years), 68% of the current pain episodes had a duration of more than 12 weeks. In eleven patients (9.6%) the current episode was shorter than 6 weeks, but in none of these patients this was the first episode of LBP. Overall 69% of all patients had recurrent pain. On the average, patients were on the sicklist for more than 8 weeks (57 calendar-days; *SD* = 53.2, median=42) before they took part in the screening procedure. The mean score on functional disability, as measured by the Roland Disability Questionnaire (RDQ; Roland & Morris, 1983), was 13.0 (*SD* = 5.0), and 34.8% of the patients had a score higher than 15, suggesting that they were moderately to highly disabled (Patrick et al., 1995; Roland & Morris, 1983). Forty-eight percent of the patients were classified as normal, 38% at risk, 7% distressed- depressive, and another 7% distressed-somatic as measured by the DRAM (Main et al., 1992).

ATTRITION

The number of patients within each treatment condition at pre-, and post-treatment, 6-, and 12-month follow-up is described in table 1. During treatment, 2 patients dropped out, both in the GAPS condition. One patient never showed up after treatment introduction, and a second patient terminated the

TABLE 1. Number of patients within each Treatment Condition at Post-Treatment, 6- and 12-Month Follow-Up

Treatment condition	Pre-treatment	Post-treatment	6-month Follow-up	12-month Follow-up
GAPS	58	56	48	44
GAGE	57	57	52	50
Total	115	113	100	94

program halfway because he had found a new job. Both refused further participation in the study and were missing at the post-treatment measurement. At 6-month follow-up, another 13 patients dropped out of the research, 8 in the GAPS condition and 5 in the GAGE condition, consequently, data from 100 patients (87%) were available for analyses at this point. At 12-month follow-up, 2 patients who dropped out at 6-month follow-up attended measurements again at 12-month follow-up (both GAPS condition). Another 8 patients dropped out at 12-month follow-up, 6 in the GAPS condition and 2 in the GAGE condition. Consequently, 94 patients (81.7%) were available for data analysis one year after the intervention. The reasons for dropout were comparable between conditions. In most patients who dropped out the reason remained unclear (43%) and in 3 patients, all 3 allocated to GAPS, new addresses could not be traced. Two patients passed away before one-year follow-up, one in each condition. In one case it can not be ruled out that pain and disability reported during the research, (partly) stemmed from the disease that eventually caused death.

With regard to observational measures, an additional 10, 27, and 34 patients, dropped out at respectively post-test, 6-, and 12-month follow-up, because they could not attend the measurement at the rehabilitation center. In most cases, patients could not attend the measurement because they had returned to work. The number of additional dropouts in each condition was 6 (GAPS) versus 4 (GAGE) at post-test, 16 versus 11 at 6-month follow-up, and 17 versus 17 at 12-month follow-up.

In order to detect whether attrition could be predicted by treatment condition, a logistic regression analysis was carried out with dropout at 12-month follow-up as dependent variable (0=no; 1=dropout). Condition did not predict dropout rate ($p>0.1$).

RANDOMIZATION CHECK

In table 2, baseline characteristics within each condition are displayed. There were no significant differences between GAPS and GAGE with regard to baseline characteristics. Although not significant ($p<0.1$), patients in the GAPS condition had their first pain episode about 10 years ago whereas patients in the GAGE condition had their first pain episode about seven years ago. To prevent this inequality from biasing the results, pain duration was added to all analyses as a covariate. There were no significant differences between treatment conditions at pretreatment measurements.

COMPLIANCE

The percentage of GA-sessions missed was low in both conditions. Patients in the GAPS condition on the average missed 4.5% of GA-sessions, whereas patients in the GAGE condition missed 6.5% of GA-sessions. Presence during either PST or EDU-sessions was equally high: on the average 9.6 and 9.3 sessions out of 10 in respectively GAPS and GAGE conditions. With regard to

TABLE 2. Baseline characteristics of patients in GAPS (*n*=58) and GAGE (*n*=57).

Variable	%GAPS		%GAGE		<i>p</i> ^a
Gender (%male)	74.1		71.9		.79
Education					
High	39.7		40.4		.94
Low/Medium	60.3		59.6		
Current medication use					
Yes	37.9		31.6		.48
Disability compensation					
Yes	6.9		5.3		.71

Variable	GAPS		GAGE		<i>p</i> ^b
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Age	40.3	9.1	39.3	8.8	.53
Pain duration since first pain episode (years)	9.0	9.7	6.6	6.7	.13
Duration current pain episode (years)	2.3	6.5	1.3	2.5	.27
Sickleave duration (weeks)	7.7	6.1	8.6	8.9	.56

Note. GAPS = Graded Activity + Problem Solving Therapy; GAGE = Graded Activity + Group Education. ^a Based on Chi-square test. ^b Based on independent samples *t*-test

workplace visits, administration was missing with regard to 14 patients. Relatively more patients in the GAGE condition received a workplace visit as part of the therapy, 53% versus 38% in the GAPS condition. The difference, however, was not significant ($\chi^2=2.39$; $p>0.1$). Data on booster-attendance were available with regard to 83 patients. Relatively more patients in the GAGE than in the GAPS condition attended the booster session, 76% versus 58%. However, this difference was not significant ($\chi^2=3.17$; $p<0.1$).

Finally, compliance regarding homework assignments was registered in the GAPS condition. Unfortunately, homework registration of 3 patients was missing. On the average patients completed 15.7 of the 21 assignments ($SD=5.1$), with half of the patients completing more than 16 assignments and seven patients (13%) completing less than 10 assignments.

TREATMENT CREDIBILITY

Treatment credibility with regard to the complete treatment program (including graded activity) did not differ between conditions. However, credibility of problem solving therapy versus group education, differed significantly between conditions (68 vs. 79; $t(94)=-3.49$; $p<0.01$), and was therefore added to the effect analyses as a covariate. At post-treatment there were no significant dif-

ferences between GAPS and GAGE condition anymore (62 versus 69, respectively). Apparently, patients in both conditions lowered their expectations somewhat at the end of treatment, mostly in the GAGE condition.

MANIPULATION CHECK

In table 3, pre- and post-treatment scores on the SPSI-R subscales are shown. At pre-treatment, scores were comparable to a group of 475 soldiers with back pain, as well as to a group of 495 soldiers without back pain (Shaw et al., 2001). Paired *t*-tests show that the GAPS condition changed significantly with regard to 4 of 5 subscales: patients in the GAPS condition increased positive problem orientation (PPO) and rational problem solving skills (RPS) and decreased negative problem orientation (NPO) and impulsiveness/carelessness style (IMP). There were no significant changes in the GAGE condition. Between-group differences were significant with regard to PPO ($p < 0.01$) and RPS ($p < 0.05$), and showed borderline significance with regard to NPO ($p < 0.1$). In general, it might be concluded that self-reported problem-solving skills changed in the GAPS condition only.

POST-TREATMENT ANALYSES

Both conditions demonstrated changes in the hypothesized direction on primary and secondary outcome measures (table 3). The pre-post changes within each treatment condition were significant ($p \leq 0.01$). Between-group differences were analyzed by means of multiple regression analyses. One subject in the GAPS condition was eliminated from the analyses (manipulation check, post treatment, 6- and 12-month follow-up, per-protocol analyses), because he turned out to be an influential case with regard to several dependent variables and measurements (Cook's distance > 1). Outcomes with regard to primary outcome measures and depression are displayed in table 4. The table shows the final step of multiple regression analyses after application of hierarchical backward elimination. There were no significant differences between the conditions on any of the outcome measures at post-treatment (GAPS vs. GAGE). A higher education level and a longer duration of the current pain episode were positively related to depression at post-treatment.

SIX-MONTH FOLLOW-UP

A half-year after the intervention, most of the outcome measures still change in the hypothesized direction as compared to post-treatment (table 3), although a statistically significant decrease is only found with regard to functional disability (QBPDS, RDQ) in both conditions and pain behavior (PBS) in the GAPS condition. Multiple regression analyses showed that functional disability as measured by the QBPDS was predicted by treatment condition ($p < 0.06$), in that the GAPS condition had less functional disability than the GAGE condition at 6-month follow-up (table 4). Other predictors in the models were pretreatment measurement and duration of the current pain epi-

sode. Furthermore, a high education level predicted higher depression (BDI) at 6-month follow-up (table 4).

TWELVE-MONTH FOLLOW-UP

One year after the intervention, both conditions maintain the outcome-levels of 6-month follow-up (table 3). Treatment condition turned out to be predictive of functional disability as measured by the RDQ at 12-month follow-up ($p < 0.05$; table 4). Patients from the GAGE condition report significantly more disability than those in the GAPS condition. Pretreatment measurement was the most powerful predictor in each of the models. Furthermore, duration of pain since the first pain episode was predictive of functional disability as measured by the QBPDS ($\beta = .213$, $p < 0.05$; table 4).

THE COST DIARY

In the first 26 weeks after the intervention, 28 patients (17 GAPS and 11 GAGE) filled out less than 18 diary-weeks and these patients were therefore eliminated from this analysis. Nine of these patients, 3 GAPS and 6 GAGE did not fill out any diary at all. In the second half-year after the intervention, 32 patients (21 GAPS and 11 GAGE) filled out less than 18 diary-weeks, of which 25 patients (15 GAPS and 10 GAGE) did not fill out any of the diaries. Consequently, data of 87 and 83 patients respectively were available for the first and second half-year follow-up data analysis. Table 5 displays cost-diary data. Proportion, mean and standard deviation (*SD*) of physical (health) activities (i.e. fitness, aerobics, physical exercises, swimming, relaxation exercises, walking, cycling), health care utilization and days of sickleave during 2 half-year follow-up periods, by treatment condition, are displayed.

It is striking, that health care utilization in the current sample is about 2 to 3 times smaller as compared to a group of chronic low back pain patients with longstanding work disability (Goossens et al., 1998). An exception to this observation is the number of alternative medicine contacts. Comparing the second with the first half-year period, fewer patients in both conditions report contacts with a General Practitioner and a medical specialist, whereas more patients report contacts with a physiotherapist, during the second half-year follow-up period. The proportion of patients within each condition that reports sickleave due to back pain reduces in both conditions from the first to the second half-year follow-up. With regard to physical health activities, most of the patients (GAPS: 95%; GAGE: 94%) indicate that they carried out such activities during the first half-year follow-up period. In the GAPS condition, the same proportion is observed during the second half-year. In the GAGE-condition, however, the proportion of patients carrying out physical health activities drops to 78%. The difference between treatment conditions is statistically significant ($\chi^2(1) = 4.42$; $p < 0.05$). None of the other comparisons between conditions were significant.

Secondly, it was analyzed whether there were between-group differences

Table 3. Means (*M*) and Standard Deviations (*SD*) for all dependent measures by condition at pre-treatment, post-treatment, 6- and 12-month follow-up assessments.

Measure	Pre-treatment <i>m (sd)</i>	post-treatment <i>m (sd)</i>	6-month Follow-up <i>m (sd)</i>	12-month Follow-up <i>m (sd)</i>
Manipulation Check:				
Social Problem Solving Inventory (SPSI-r)				
Positive Problem Orientation (PPO)				
GAPS	11.9 (3.7)	12.9 (3.3)		
GAGE	12.6 (2.9)	12.0 (3.1)		
Negative Problem Orientation (NPO)				
GAPS	8.5 (5.8)	6.5 (5.6)		
GAGE	8.6 (5.2)	7.8 (5.0)		
Rational Problem Solving (RPS)				
GAPS	41.4 (13.1)	47.1 (12.7)		
GAGE	42.3 (12.8)	43.4 (14.2)		
Impulsive/carelessness style (IMP)				
GAPS	12.2 (4.9)	10.8 (5.0)		
GAGE	12.8 (4.8)	12.1 (5.0)		
Avoidance Style (AS)				
GAPS	6.1 (4.0)	6.3 (3.4)		
GAGE	6.3 (3.9)	5.9 (4.0)		
Primary outcomes:				
Quebec Back Pain Disability Scale (QBPDS)				
GAPS	41.5 (16.9)	30.0 (19.9)	25.0 (19.0)	25.8 (22.5)
GAGE	39.6 (13.1)	31.3 (18.1)	27.6 (19.3)	27.3 (21.4)
Roland Disability Questionnaire (RDQ)				
GAPS	13.7 (5.2)	9.1 (6.3)	6.2 (6.0)	6.5 (6.7)
GAGE	12.4 (4.8)	8.5 (5.6)	6.8 (6.3)	7.7 (6.5)

Secondary outcomes:

McGill Pain Questionnaire (MPQ)

GAPS	17.6	(10.1)	13.4	(9.6)	12.6	(10.4)	11.3	(11.0)
GAGE	18.1	(8.4)	15.2	(10.5)	13.6	(9.8)	14.0	(11.5)

Pain Catastrophizing Scale (PCS)

GAPS	16.6	(11.1)	12.2	(10.8)	11.3	(11.0)	12.9	(11.8)
GAGE	16.1	(8.3)	11.5	(9.7)	11.7	(8.0)	12.0	(10.1)

TAMPA Scale for Kinesiophobia (TSK)

GAPS	36.7	(8.7)	33.5	(7.9)	32.3	(8.2)	33.1	(9.2)
GAGE	37.1	(6.8)	33.3	(8.3)	33.3	(8.3)	32.8	(8.7)

Fear Avoidance Beliefs Questionnaire – Physical Activity (FABQ-act)

GAPS	10.8	(5.0)	7.3	(4.6)	6.4	(5.5)	8.5	(7.2)
GAGE	10.3	(5.0)	6.7	(5.2)	6.2	(5.6)	7.8	(5.9)

Fear Avoidance Beliefs Questionnaire – Work (FABQ-work)

GAPS	17.8	(9.8)	13.0	(9.3)	12.6	(11.4)	12.0	(11.7)
GAGE	17.3	(9.6)	14.2	(10.3)	13.3	(10.8)	13.6	(11.2)

Beck Depression Inventory (BDI)

GAPS	7.5	(5.8)	5.4	(5.6)	6.2	(6.8)	6.0	(8.0)
GAGE	6.8	(5.1)	5.4	(5.7)	5.5	(5.4)	6.2	(6.6)

Observational measures:

Behavior Approach Test (BAT, distance in *m*)

GAPS	429.1	(160.9)	530.2	(134.8)	546.1	(155.4)	531.8	(150.7)
GAGE	473.1	(140.3)	543.3	(139.8)	512.9	(152.1)	538.4	(138.7)

Pain Behavior Scale

GAPS	2.2	(1.6)	0.9	(1.3)	0.8	(1.2)	1.0	(1.8)
GAGE	1.9	(1.5)	0.8	(1.4)	0.8	(1.2)	1.0	(1.5)

Note. GAPS = Graded Activity + Problem Solving Therapy; GAGE = Graded Activity + Group Education.

Table 4. Results of Multiple Linear Regression Analyses at Post-treatment, 6- and 12-Month Follow-Up for Primary Outcome Measures and Depression

Dependent Variable	<i>n</i>	Adj. <i>R</i> ²	Independent Variables	Standardize β
<i>Post-treatment:</i>				
Functional Disability – QBPDS	112	.446***	Treatment condition	.083
			Pre-treatment QBPDS	.693***
Functional Disability – RDQ	111 ^a	.444***	Treatment condition	.038
			Pre-treatment RDQ	.691***
Depression – BDI	112	.533***	Treatment condition	.068
			Pre-treatment BDI	.719***
			Duration current pain episode	.149*
			Education level	.149*
<i>6-Month Follow-Up:</i>				
Functional Disability – QBPDS	99	.282***	Treatment condition	.173†
			Pre-treatment – QBPDS	.553***
Functional Disability – RDQ	97 ^b	.139**	Duration current pain episode	.153†
			Treatment condition	.152
			Pre-treatment – RDQ	.363***
Depression – BDI	99	.390***	Duration current pain episode	.238*
			Treatment condition	.075
			Pre-treatment – BDI	.571***
			Duration current pain episode	.213*
			Education level	.172*
<i>12-Month Follow-Up:</i>				
Functional Disability – QBPDS	93	.192***	Treatment condition	.135
			Pre-treatment – QBPDS	.446***
Functional Disability – RDQ	92 ^b	.161**	Duration since first pain episode	.213*
			Treatment condition	.233*
Depression – BDI	89 ^a	.161**	Pre-treatment – RDQ	.358***
			Treatment condition	.101
			Pre-treatment – BDI	.408***

Note. Treatment condition: GAPS = 0; GAGE = 1; Adj. = adjusted; QBPDS = Quebec Back Pain Disability Scale; RDQ = Roland Disability Questionnaire; BDI = Beck Depression Inventory; ^a Data of one or more patients were eliminated from the regression analyses because Studentized Residuals were out of range (Stud. res. < -3 or > 3); ^b Data of one or two patients were missing on this particular measure because of administratively failure. Other predictors were only displayed when they had a (marginally) significant contribution to the model ($p < 0.1$). † $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 5. Proportion and mean (*SD*) of physical (health) activities, health care utilization and days of sickleave per half-year follow-up period by treatment condition.

Measure	0-26 weeks follow-up (N=87)			27-52 weeks follow-up (N=83)		
	% Checked 'yes'	<i>M</i>	(<i>SD</i>)	% Checked 'yes'	<i>M</i>	(<i>SD</i>)
Physical (health) Activities						
GAPS	95	80.53	(67.17)	95	76.72	(71.50)
GAGE	94	88.48	(83.07)	78	74.36	(82.67)
General Practitioner contacts						
GAPS	51	1.45	(1.87)	38	2.91	(10.08)
GAGE	46	1.63	(3.13)	39	1.33	(3.00)
Physiotherapy contacts						
GAPS	17	0.97	(2.54)	27	4.98	(15.86)
GAGE	24	3.25	(7.72)	37	4.94	(9.22)
Specialist care contacts						
GAPS	32	1.11	(2.19)	14	0.5	(1.62)
GAGE	22	1.01	(2.48)	17	1.13	(4.19)
Alternative Medicine contacts						
GAPS	7	0.32	(1.47)	14	1.60	(7.46)
GAGE	15	0.66	(1.92)	4	0.11	(0.56)
Days of sickleave						
GAPS	63	25.46	(37.07)	35	20.63	(39.92)
GAGE	61	21.25	(37.10)	39	24.64	(43.69)

Note. GAPS = Graded Activity + Problem Solving Therapy ($N_{0-26\text{wks}}=41$ and $N_{27-52\text{wks}}=37$); GAGE = Graded Activity + Group Education ($N=46$).

with regard to the absolute numbers reported in the diaries. Because tests for normality (Kolmogorov-Smirnov test) were all significant ($p < 0.001$), it was decided to use non-parametric tests for this comparison. There were no significant differences between the conditions during the first or second half-year period.

Regarding sickleave, changes in absolute days of sickleave within each condition are only small. However, when means are calculated regarding only those patients reporting sickleave, the mean score increases in both conditions from the first to the second half-year period (GAPS: from 40 to 59 days; GAGE: from 35 to 63 days). This might indicate that the smaller group of patients reporting sickleave during the second half-year period, are those patients most at risk to develop chronic work-disability, and some of them might already receive disability pensions.

TABLE 6. Dropouts and work status within each condition.

Work status	6-month Follow-up		12-month Follow-up	
	GAPS	GAGE	GAPS	GAGE
Total	10	5	14	7
• 100% return to work ¹	3	1	7	1
• < 100% or no return to work ²	4	1	2	3
• not known ³	3	3	5	3
(% of patients with full return to work*)	(77.1)	(70.7)	(85.4)	(68.3)

Note. GAPS = Graded Activity + Problem Solving Therapy; GAGE = Graded Activity + Group Education; ¹ Imputed the best outcomes (mean score of the best quartile); ² Imputed the worst outcomes (mean score of the worst quartile); ³ Imputed the best or the worst outcomes depending on the within-group ratio; * Based on available data from the Occupational Health Services.

INTENTION-TO-TREAT ANALYSIS

Table 6, shows the number of dropouts with regard to self-report measures, imputed the best outcomes (mean score of the best quartile), and the worst outcomes (mean score of the worst quartile). In case the work status of the dropout was unknown, ratios (% of patients with full return to work) were applied. Two patients who dropped out at post-treatment were given best and worst outcomes, according to their work status at that time. At 6-month follow-up, 77.1% of the patients in GAPS and 70.7% in GAGE had full (100%) return to work. At 12-month follow-up, this was 85.1% in GAPS and 68.3% in GAGE. Multiple regression analyses were executed to check whether results with regard to disability and depression changed as a result of data-imputation according to the preset rules mentioned above. With regard to post-treatment and 6-month follow-up, there were no outcomes that differed from those ini-

tially reported when dropouts were not imputed. With regard to 12-month follow-up, however, treatment condition was not only predictive of RDQ (standardized $\beta = 0.27$; $p < 0.01$), but also of QBPDS (standardized $\beta = 0.21$; $p < 0.05$) and BDI (standardized $\beta = 0.17$; $p < 0.05$). In other words, patients in the GAPS condition were less disabled and less depressed in comparison to patients in the GAGE condition, one year after the intervention. These results are in concordance with those found before and strengthen the impression that the GAPS condition has better outcomes in the longer term. In some cases, one or more significant predictors were added to the final model. At 12-month follow-up, treatment credibility significantly predicts RDQ and BDI. Patients with higher credibility ratings report less disability and less depression. With regard to the cost diary, there were no differences in comparison to the initial effect analyses.

PER-PROTOCOL ANALYSIS

In order to execute the per-protocol analysis, patients who attended less than 66% of the GA-sessions, or less than 8 sessions of PST were left out from the analyses. As a result, 4 patients were omitted in the GAPS condition and 5 in the GAGE condition. One subject in the GAPS condition turned out to be an influential case in several analyses (Cook's distance >1), and was therefore dropped from the analyses. Consequently, data of 105 patients were left for the per-protocol analysis.

Regarding the manipulation check, conditions still differed significantly with regard to RPS and PPO, favoring the GAPS condition. The trend found earlier with regard to NPO, however, disappeared now that therapy attendance was taken into account. With regard to evaluation of primary measures, no differences were found between per-protocol analysis and initial effect evaluation. Therapy conditions did not differ with regard to post-treatment, and showed borderline effects with regard to QBPDS at 6-month follow-up, and with regard to RDQ ($p < .06$) at 12-month follow-up. These results endorse those initially found. However, the per-protocol analysis failed to show the stronger effects one would expect when only data are analyzed of those patients who sufficiently attended the therapy program.

Other predictors in the model were comparable to those in the initial effect evaluation. There was one exception: in comparison to the initial analysis, duration of pain since first pain episode becomes an additional predictor of the QBPDS at post-treatment in the per-protocol analysis. Results regarding the cost-diary were comparable to those initially found.

CLINICAL RELEVANCE

Next to statistical significance, clinical relevance of the results with regard to RDQ-changes was considered. Changes were related to the initial scores at pre-treatment. Comparison of the number of patients in 6 equally divided

Figure 1. Percentage of patients with a clinically relevant improvement by condition with regard to functional disability (RDQ) at post-treatment (POST), 6- (FU 1) and 12-month (FU 2) follow-ups. GAPS = Graded Activity + Problem Solving Therapy; GAGE = Graded Activity + Group Education; RDQ = Roland Disability Questionnaire

RDQ-classes at pre-treatment, showed a significant difference ($\chi^2 = 14.89$; $p < 0.05$) between the 2 conditions in that GAPS had a higher percentage of patients in the RDQ-classes representing higher functional disability (RDQ > 16; 36.2% versus 17.5%). According to the preset criteria, to assess clinical relevance, a change of at least 8 RDQ-points was expected when the initial RDQ-score was higher than 16.

In figure 1, the percentage of patients that had an improvement that was clinically relevant is displayed at each measurement point. At 12-month follow-up, 3 patients (7%) in the GAPS condition and 7 patients (15%) in the GAGE condition increased their RDQ-scores in comparison to initial scores. Because criteria for a clinically relevant 'worsening' were not defined, these patients were regarded as unchanged in the analyses. At post-treatment, conditions were equally improved. During the first half-year following the intervention, the proportions of patients with a clinically relevant improvement were still increasing in both conditions. Finally, at both 6- and 12-month follow-up, the GAPS condition had an extra 15% of patients that showed a clinically relevant improvement as compared to the GAGE condition. However, this difference was not statistically significant.

PROCESS EVALUATION

Fifty-three patients in the GAPS condition and 50 patients in the GAGE condition filled out the process evaluation form. Although it was stressed that the

TABLE 7. Process evaluation. Percentage (%) of statements agreed with, disagreed with, or neither (neutral) by condition.

Statements cluster:	GAPS			GAGE		
	agree (+)	neutral (0)	disagree (-)	agree (+)	neutral (0)	disagree (-)
treatment expectations and achievement of treatment goals	50.8	33.0	16.2	82.2	11.6	6.2
treatment process	75.0	14.4	10.6	75.8	14.6	9.6
working in a group	92.3	4.5	3.2	93.4	4.8	1.8
therapists	95.6	4.4	0.0	95.5	3.6	0.9
organization and facilities	84.1	12.1	3.8	82.9	14.6	2.5
homework assignments	67.1	20.9	12.0	-	-	-
problem solving skills	64.7	26.6	8.7	-	-	-

Note. GAPS = Graded Activity + Problem Solving Therapy; GAGE = Graded Activity + Group Education.

evaluation only concerned the PST or EDU part of the program, still some patients used the form to evaluate the entire program. As a consequence, the report of answers in this section will cover themes from PST, EDU and graded activity. Remarkably, the content of the answers to the open questions given in the GAPS condition differed from the answers given in the GAGE condition. The most prominent difference was that GAPS patients mentioned problem solving skills as being the most important learning dimension (54% of the answers) whereas GAGE patients mainly mentioned ergonomics and healthy movement as important themes (58% of the answers). Other themes that were mentioned by both treatment groups were personal growth (e.g. saying "no", accept personal limits), and general goals (e.g. how to cope with pain). To the question which part of the treatment was indispensable to the patient, both GAPS and GAGE patients chiefly noted physiotherapy, ergonomics and occupational therapy (respectively 56% and 65% of the answers). Secondly, they mentioned specific themes that were discussed in either PST or EDU. The fact that problem solving skills were exclusively mentioned in the GAPS condition, might indicate that there had been no contamination between the two treatment conditions. The open question referring to what patients had missed in the training was commented on by only a small part of the patients. In the GAPS condition, themes concerning ergonomics or physiological topics were most frequently mentioned, whereas remarks concerning practical aspects of the treatment were most frequently mentioned in the GAGE condition. Next to the open questions patients were asked to rate the training on a 10-point scale. On the average patients in the GAPS condition rated 7.6 ($SD=1.0$) and patients in the GAGE condition rated 7.7 ($SD=0.7$). These ratings were not significantly different ($p>0.05$). Moreover, patients evaluated their personal improvement on a 4 point rating scale. Five patients did not fill out this scale because they

wanted to rate "neutral", but this, however, was not an option on the scale. Eighty-six percent of the patients in the GAPS condition and 92 percent of the patients in the GAGE condition rated their personal improvement as "satisfactory" or "very satisfactory". The remainder rated their personal improvement as "dissatisfactory" ($n=10$) or "very dissatisfactory" ($n=1$, GAPS). The Chi-square test was not significant ($p>0.05$).

Finally, the percentages of positive, negative and neutral answers to each cluster of statements were calculated per condition (table 7). A higher percentage of patients in the GAGE condition (82.2%) answered positive to statements on "treatment expectations and the achievement of treatment goals" in comparison with patients in the GAPS condition (50.8%). With regard to the clusters "treatment process", "working in a group", "therapists", and "organization and facilities", conditions were equally positive. The statements concerning homework assignments and problem solving skills were only rated in the GAPS condition. As shown in table 7, the majority of patients rated statements on these themes positively.

DISCUSSION

An intervention that included Graded Activity (GA) and Problem Solving Therapy (PST) was designed against the background of a bio-psycho-social model of pain and a growing interest in cognitive-behavioral techniques to prevent the development of chronic disability. The current study aimed to examine whether the PST had supplemental value when added to GA in a group of employees with non-specific low back pain. Data analyses showed that both GAPS and GAGE conditions resulted in significant improvement as compared to pre-treatment with regard to primary and secondary outcome measures at post-treatment, and as compared to post-treatment with regard to functional disability at 6-month follow-up. Moreover, the results at 6-month follow-up sustained at 12-month follow-up. Between-group differences were found with regard to functional disability at 6- and 12-month follow-ups, favoring the GAPS condition. With the exception of physical health activities that were performed by relatively more patients in the GAPS condition during the second half-year follow-up, data from the cost diary revealed no significant differences between the conditions. The intention-to-treat analysis confirmed previous findings and moreover, underscored the supplemental effectiveness of the GAPS-condition at 12-month follow-up with regard to functional disability and depression. The per-protocol analysis was in agreement with the initial analysis. Both intention-to-treat and per-protocol analysis revealed that although the supplemental value of the GAPS condition is modest, it appears to be robust. Regarding the clinical relevance of the findings, the GAPS condition excelled the GAGE condition by having the highest proportion of patients yielding a clinically relevant change with regard to functional disability, though not statistically significant. Finally, the process evaluation showed that patients in GAPS and GAGE condition were

equally satisfied about therapy and their personal improvement. With regard to treatment expectations and achievement of treatment goals, more patients in the GAGE condition agreed that the therapy came up to their expectations and that general therapy goals were attained.

Several tests were executed to check the internal validity of the study. First, a manipulation check demonstrated that problem-solving skills were exclusively changed in the condition including PST. This suggests that PST apparently effectively changes self-reported problem-solving skills, and that these changes were restricted to the condition in which PST was delivered. Secondly, the randomization check revealed that conditions were comparable with regard to baseline characteristics, pre-treatment measurements and compliance rate, expressed as the percentage of sessions attended during treatment. Dropout rate, however, was higher in GAPS as compared to GAGE condition. Finally, treatment credibility was measured in each condition and turned out to differ between the PST and EDU part of the treatment, EDU yielding higher credibility than PST. As differential treatment credibility can influence outcome (Turk, Rudy, & Sorkin, 1993), it was controlled for in the subsequent effect analyses. In conclusion, several checks ruled out potential threats to the internal validity of the study.

One of the reasons why the supplemental value of the PST turned out to be modest, might be that the Group Education (EDU) intervention could have been more effective than anticipated. Not only this intervention turned out to be somewhat more credible than PST, it also seemed to fulfill patients expectations better. This was endorsed by the process evaluation, showing that patients in the GAGE condition more often found the treatment to meet their expectations and therapy goals to be achieved. This might also indicate that EDU is more preferred and accepted by patients than PST. Finally, the EDU-sessions were for the greater part attended by the same therapists that attended graded activity, which might have resulted in carry-over effects, in that issues that came about during the GA-sessions were more easily discussed in detail during EDU and *visa versa*.

The difference found with regard to dropout rate, GAPS having twice as much dropouts at 12-month follow-up, deserves further exploration. Reasons for dropout were unknown in most cases, but could have been various. A reasonable explanation in this population of employees may be the fact that return to work made it too much an effort to appear at the rehabilitation center during a working day. As we had at our disposal objective data from the occupational health service on work-status (Van den Hout et al., accepted pending revision), we found out that most of the dropouts (table 6; % of patients with full return to work) had a 100% return to work. The return to work-rate further supports the suggestion that dropouts had other obligations at the time of the measurement at the rehabilitation center. Comparison between patients who had a 100% return to work versus patients who only partly or not resumed their work, showed that the former had significantly less disability, depression, health care uti-

lization and days of sickleave than the latter, at both 6- and 12-month follow-up. The same evidence was found with regard to most of the secondary outcome measures. These data support the assumption that patients dropped out because they were occupied with their jobs, rather than that they dropped out because their health status deteriorated. Finally, dropout might be due to motivational reasons. Indeed, patients who dropped out found the treatment significantly less credible at pre-treatment than those who did not ($p < 0.05$).

How likely is it that the supplemental efficacy of GAPS, one year after the intervention, is due to the increase of problem solving skills? The manipulation check indicated that problem-solving skills were exclusively changed in the GAPS condition, and between-group differences at post-treatment were only found with regard to problem-solving skills, supporting the hypothesis that it was improvement of problem-solving skills that yielded the supplemental value of the PST. On the other hand, there are some aspects to consider. First, analyses failed to find pre-post differences for 'Avoidance Style', whereas this was one of the subscales that was expected to decrease and consequently yield its effect on primary outcome measures. Second, because of the lack of normative data on the SPSI-r, it is impossible to rule out ceiling and/or bottom effects interfering with the magnitude of change scores. Third, because the SPSI-r is a self-report measure, it might not necessarily indicate that actual behaviors changed. Therefore, validation of the SPSI-r against an observational measure might be of great value. However, there is some evidence that problem orientation is important in the process of stress-related avoidance, in that a positive problem-orientation leads to less negative affect and avoidance behavior (Chang & D'Zurilla, 1996; D'Zurilla & Nezu, 1982). Moreover, low scores on positive problem orientation, and high scores on impulsivity and carelessness style, are indirectly related to functional loss in employees with low back pain (Shaw et al., 2001).

A number of interesting questions needs to be addressed in future research. First, the comparison from the cognitive-behavioral interventions described in this study to treatment-as-usual might reveal the effectiveness of the complete program and not just the supplemental value of PST. Second, since consequences of back pain involve large amounts of money spent to sickleave and disability pensions, cost-effectiveness of the rehabilitation program will be of interest to convince policymakers and employers of the use and gain of such programs. Next, one of the questions that could not be answered in this study was whether problem-solving therapy would be acceptable and effective when given as an individual therapy, as is the case in among others unexplained physical symptoms (Wilkinson & Mynors Wallis, 1994). Finally it would be interesting to find strong predictors of treatment effect. These might be found in variables other than included in the current study. Especially patient characteristics as pain-related fear (McCracken & Gross, 1998; Vlaeyen & Linton, 2000), personality traits (Bigos et al., 1991) and comorbidity (Main et al., 1992), might be of interest.

This study suffers from a number of limitations. First, the design of the

study does not allow drawing conclusions about the effects of the complete program, as no no-treatment condition was included. However, graded activity was shown to be more effective than treatment-as-usual (Lindström et al., 1992; Linton & Bradley, 1992; Linton et al., 1989). Because graded activity, as presented in the current study, is comparable to the protocols of Lindström and Linton (Lindström et al., 1992; Linton et al., 1989), the superiority of the GAPS condition may be attributed to the PST ingredient and adds supplemental value to the effect of graded activity.

Second, for practical reasons, the study was not double blind in that all therapists knew in which of two conditions patients took part. In case therapists were favoring one of two conditions this preference might have biased their efforts during therapy. Moreover, the study lacks a treatment integrity check, what might have cast doubt on the performance of the therapies according to treatment protocols. Nevertheless, potential bias regarding therapists' preferences and how they performed therapy was restricted to a minimum. First, therapies were strictly protocolized such that there was no room for divergent interpretations. Furthermore, the research team supervised therapists on a regular basis.

Third, with the exception of the observational measures, outcome measures are all based on self-report. It would be interesting to add measures that were collected from an objective source. With regard to days of sickleave, data were also obtained from the occupational health care services (OHS) involved in the study, and are described elsewhere (Van den Hout et al., accepted pending revision). In the current study, OHS-data were used as objective source to impute missing data for the intention-to-treat analysis. Remarkably, and in line with the results from this study, OHS-data revealed that patients in the GAPS condition had a higher rate of 100% return to work than patients in the GAGE condition, at one-year follow-up.

Despite these limitations, some recommendations for the implementation of PST in pain patients might be called for. First, patients must be aware of the fact that therapy will focus on the consequences of pain in daily life rather than pain reduction (Aldrich et al., 2000). Moreover, in order to improve therapy success, therapists must agree on this point and convey concordant messages to the patient. In line with this, it might be important to check whether patients belief in therapy rationale on forehand. Kole-Snijders et al. (1999) also found treatment credibility to be one of the most important predictors of therapy success. To introduce patients into the rationale at the start of the therapy or even before might be helpful to prevent false expectations. As already discussed, education might not necessarily be an effective part of therapy, but certainly may reassure patients for whom worries about pain are a continuous source of distress. Secondly, it might be useful to check whether the patient is prepared to change behaviors, as only then therapy can render fruitful outcomes (Dijkstra, Vlaeyen, Rijnen, & Nielson, 2001). As for treatment protocol, two adjustments might further improve the treatment protocol. First, in the current

study graded activity was initially taught by training of a set of standard physical exercises. In case pain-related fear is high, it might be useful to identify those specific movements and activities the patient is afraid of. In a single case experimental design, Vlaeyen, de Jong, Geilen, Heuts, and Van Breukelen (2001) found promising results with the application of graded exposure in vivo with behavioral experiments, i.e. challenging expectations of harmfulness of physical activities, in patients reporting elevated levels of pain-related fear. One of the characteristics of the therapy was that it focused on the specific movements and activities the patient fears and the idiosyncratic thoughts that go along with it. Especially when patients show high levels of pain-related fear, it might be useful to introduce exposure treatment as a part of graded activity, and thereby increase the possibility that therapy principles will be generalized to daily life and work. Secondly, according to the therapy protocol problem-solving skills were exclusively discussed during the PST-sessions. However, the generalization of skills to daily living might be even more fruitful when other disciplines within the rehabilitation team appeal to these skills whenever the patient brings about a problem. In that way, patients might be encouraged to discuss problems and use problem-solving skills to solve them, not only in therapy but also in daily lives and environments (Loscalzo & Bucher, 1999).

CHAPTER 7

Secondary prevention of work-related disability in non-specific low back pain: Does problem-solving therapy help? A randomized clinical trial.

This chapter was accepted pending revision:

Van den Hout, J.H.C., Vlaeyen, J.W.S., Heuts, P.H.T.G, Zijlema, J.H.L., Wijnen, J.A.G.:
Secondary prevention of work-related disability in non-specific low back pain: Does
problem-solving therapy help? A randomized clinical trial. *Clinical Journal of Pain*.

ABSTRACT

Given the individual and economic burden of chronic work disability in low back pain patients, there is a call for effective preventive interventions. In a randomized clinical trial, it was investigated whether problem-solving therapy had a supplemental value when added to behavioral graded activity, regarding days of sick leave and work status. Data from 84 employees, who were recently out off work as a result of non-specific low back pain, were retrieved from occupational health services. Forty-five employees had been randomly assigned to the experimental condition including problem-solving therapy and behavioral graded activity (GAPS), and thirty-nine employees had been randomly assigned to behavioral graded activity and group education (GAGE). Randomization and manipulation were checked. Moreover, a sensitivity analysis was executed regarding the calculation of days of sick leave. Data analyses show that employees in the GAPS condition had significantly less days of sick leave in the second half-year after the intervention. Moreover, work status was more favorable for employees in this condition, in that more employees had a 100% return-to-work and fewer patients ended up receiving disability pensions one year after the intervention. Sensitivity analyses confirmed these results. It was concluded that the addition of problem solving therapy to behavioral graded activity had supplemental value in employees with non-specific low back pain.

INTRODUCTION

Indirect costs, e.g. absenteeism and disability pensions, make up about 90% of the economic burden of low back pain (Goossens et al., 1998; OECD & Development, 1996; Van Tulder, Koes, & Bouter, 1995). In the Netherlands, indirect costs due to back pain are estimated at US \$4.6 billion, equivalent to 1.6% of the gross national product (Van Tulder et al., 1995). Moreover, total costs of back pain are not normally distributed. A relatively small group of patients with chronic low back pain, about 10 to 25%, is responsible for about 75% of the economic burden of back pain (Cats-Baril & Frymoyer, 1991). Yet, no satisfactory answer has been given to the rising trend of low back pain disability and as Waddell (1998) concludes: "...trends of low back pain disability show that current management and health care have failed to solve the problem" (pp. 404). Obviously, the development of effective interventions to prevent chronic work disability is of highest priority.

Secondary prevention of low back pain is defined in terms of initiatives taken from the point of the first reported pain incident until a full chronic problem has developed, that is pain that continues after 6 months of health care or sick leave (Linton & Bradley, 1996). In a recent review by Linton & Van Tulder (2001b), the effects of subsequent educational efforts, lumbar supports, exercises, ergonomics and risk factor modification in non-healthcare settings, e.g. in the general population or in the workplace, were evaluated. The effects reported from the 27 investigations were very sobering in that only 'exercises' provided sufficient evidence to conclude that this form of prevention was effective. According to Linton and Van Tulder (2001b), the reasons for this failure may be found in the methodology of the studies involved, as well as in the way interventions were designed.

In the last decades, a shift took place towards more cognitive-behaviorally oriented interventions, with accents on gradual exposure to physical activities and the application of operant conditioning principles (Fordyce, Brockway, Bergman, & Spengler, 1986; Lindström et al., 1992; Linton & Bradley, 1992; Linton, Bradley, Jensen, Spangfort, & et al., 1989; Philips, Grant, & Berkowitz, 1991). In addition, cognitive therapy has often been successfully applied whether or not in combination with other forms of therapy (Kole Snijders et al., 1999; Turner & Jensen, 1993). Moreover, there are several indications that a multidisciplinary approach is effective in the prevention of long-term disability and work loss (Feuerstein, Menz, Zastowny, & Barron, 1994; Loisel et al., 1997). Recently, a cognitive-behavioral return-to-work program was found to be more effective than treatment as usual in reducing days of sick leave and increasing general activity levels (Marhold, Linton, & Melin, 2001). However, these effects were only found in patients with short-term sick leave (2-6 months) and not in those with long-term sick leave (>12 months), underscoring the importance for an early return-to-work intervention.

Work-related problems, such as not being able to meet job demands or conflicts with co-workers cause additional stress, on top of the burden already caused by back pain. In this context, sick leave because of back pain complaints can sometimes be considered a medical solution to a work-related problem (Knepper & Feenstra, 1991). Recently, a negative orientation towards problems was found to be associated with higher levels of functional disability in employees with low back pain (Van den Hout, Vlaeyen, Heuts, Sillen, & Willen, 2001b). In line with this, Shaw, Feuerstein, Haufler, Berkowitz & Lopez (2001) found low scores on positive problem orientation and high scores on impulsiveness and avoidant style associated with functional loss in low back pain patients. The authors suggest that the prolonged impact of low back pain on daily functioning may be reduced by assisting workers to conceptualize low back pain (LBP) and its consequences as problems that can be overcome and to use active strategies in reducing risks for low back pain disability. Problem-solving therapy, based on the model by D'Zurilla and Goldfried (1971), might successfully be used in the prevention of work disability. It teaches strategies which might help to feel confident and in control of stressful situations, for example to solve work-related problems when pain relapses. Problem-solving strategies have been successfully applied in pain patients previously (Linton & Bradley, 1992; Linton et al., 1989; Philips et al., 1991), but integrated in multimodal treatment packages. To our knowledge the unique contribution of a protocolized problem-solving therapy has never been evaluated.

The current study describes a secondary preventive intervention developed against the background of a biopsychosocial model of pain. The experimental intervention consists of behavioral graded activity and problem-solving therapy. In the complementary part of this study (Van den Hout et al., under review), both treatment conditions, i.e. graded activity with and without problem solving therapy, resulted in significant improvements with regard to functional disability, pain intensity, pain-related fear, pain behavior and depressive symptoms at post treatment. Improvements were maintained at 6- and 12-month follow-up. Although modest, patients receiving problem-solving therapy in addition to graded activity reported less functional disability than patients in the GAGE condition did. In addition, the problem-solving 'group' included the highest proportion of patients yielding a clinically relevant decrease of functional disability at 6- and 12-month follow-up.

The aim of the current study is to determine whether problem-solving therapy has a supplemental value in addition to behavioral graded activity in employees who have a new episode of sick leave as a result of low back pain, with regard to days of sick leave and work status.

MATERIALS AND METHODS

STUDY DESIGN

A two-armed randomized clinical trial was executed in an occupational rehabilitation setting, in order to specifically test the supplemental value of problem-solving therapy when added to behavioral graded activity (Figure 1). A rehabilitation physician (PH) and a mental health scientist (JvdH), both of whom were blind for the allocated condition, carried out the selection procedure. Patients had to sign an informed consent. After patients had been included, they were assigned to one of two treatment conditions in blocks of at maximum 5 patients. The randomization scheme was computer-generated and known only to the logistics planner of the rehabilitation center.

Two treatment conditions were compared, the first being Graded Activity plus Problem Solving Therapy (GAPS), and the second being Graded Activity plus Group Education (GAGE). Group Education was chosen as an attention-control such that non-specific factors were equally provided in GAPS and GAGE conditions. The same therapists in both treatment conditions guided the graded activity program. Therapists were not blind for the condition because multidisciplinary consultation was part of the treatment program. To avoid contamination, GAPS- and GAGE-groups had their program planned separately such that they did not encounter one another.

Figure 1. The study design

R: Randomization; Pre: Pre-treatment assessment; Post: Post-treatment assessment; FU1: 6-month follow-up; FU2: 12-month follow-up; Period 1: one to half-year before intervention; Period 2: half-year before intervention; period 3: first half-year after intervention; period 4: second half-year after intervention.

SUBJECTS

All patients were employees who were recently sicklisted with low back pain (LBP), and referred to the study by general practitioners, occupational physicians or rehabilitation physicians. Eligibility criteria for patients in the study were as follows: age between 18 and 65 years, LBP for more than 6 weeks, on sick leave with LBP but no longer than 20 weeks, no more than 120 days of sick leave during the last year. Non-eligible were those patients who had specific back complaints (vertebral fracture, infectious disease, rheumatoid arthritis, ankylosing spondylitis, or herniated disc; Waddell & Turk, 1992); who had predominant psychopathology; were pregnant; were not proficient in Dutch; or were under medical treatment for their LBP at the time of referral. Furthermore, patients had to agree to stop any other ongoing treatments they were receiving for their back complaints at the start of the intervention.

Patients with medical comorbidity were excluded when the disorder interfered with the treatment program or rendered them unable to participate in every part of the program, as decided by the rehabilitation physician. Patients who were involved in any litigation regarding work conflicts were excluded as well.

TREATMENTS

The treatment program consisted of 19 half-day sessions that were spread over 8 weeks and given in small groups of at most 5 patients. In the course of the program the team of therapists had three meetings with individual patients. During these meetings, progresses and hindrances for goal achievement and return to work were discussed. Two months after the final treatment session, a booster session was planned in which treatment components were summarized and individual developments were discussed in the group.

Graded Activity (GA)

Graded activity is an operant behavioral treatment that aims to increase activity levels using quota systems. Included in the training were registration of baseline levels during the first two weeks, a treatment contract, positive reinforcement for activity increments, and a workplace visit (Fordyce, 1976). The training consisted of 15 one-hour sessions and 3 additional sessions dedicated to back education and lifting instructions. The occupational therapist saw patients for 30 minutes per week on an individual basis in which principles of graded activity were applied in personally relevant activities like work, house-keeping, and leisure-time activities. Furthermore, the occupational therapist contacted the occupational physician and the patient's supervisor at the workplace to discuss a plan of return-to-work. A workplace-visit was planned when considered necessary by the occupational therapist.

Problem Solving Therapy (PST)

Problem solving therapy is a cognitive-behavioral therapy in which problem-

solving skills are taught according D'Zurilla & Goldfried (1971). The PST-model describes 5 steps in which problems are typically solved: 1) problem orientation, 2) problem definition and formulation, 3) generation of alternatives, 4) decision making, and 5) implementation and evaluation. In the current trial, a protocol, as originally designed by Nezu and colleagues (Nezu, 1986; Nezu & Perri, 1989) was translated and applied as a group intervention. Skills were trained in 10 ninety-minute sessions, and were guided by two experienced behavior therapists. The foci of the therapy were skills training and application of skills in daily life, rather than one specific problem area. Patients were free to select their own problem areas, which did not need to be pain-related. Between sessions, homework assignments were given to practice skills in everyday life. Homework assignments were discussed in the group in all sessions.

Group Education (EDU)

An attention-control intervention was designed in order to make both conditions comparable with regard to non-specific factors. Since educational approaches turned out to have limited effects on worrying, functional status, health care utilization or other primary outcome measures (Berwick, Budman, & Feldstein, 1989; Cherkin, Deyo, Street, Hunt, & Barlow, 1996), this intervention was chosen as the attention-control. EDU consisted of 10 ninety-minute lessons in which issues related to the back and to back pain were discussed. A physiotherapist, an occupational therapist, and a psychologist, using a protocolized manual, served as lecturers. Explicitly, no skills were taught and each theme was discussed during no more than one protocolized session.

OUTCOME MEASURES

Days of sick leave

Data on sick leave were obtained from the occupational health service (OHS) associated with the workplace of the employee. Employees were requested to give written consent, which would allow the research team to retrieve data from the OHS. In total, 23 OHS were involved in the data collection process. There is no national administration of sick leave data in the Netherlands, therefore, data were recorded quite differently. One research assistant who was blind to the condition calculated the number of days of sick leave taking into account: 1) a work-ratio of 5 days a week; 2) the percentage of full-time equivalents (fte); 3) disability pensions, which were defined in terms of days of sick leave. Data were calculated with respect to 4 time periods: 2 half-year periods preceding the intervention (period 1 and 2); and 2 half-year periods following the intervention (period 3 and 4). Also a variable regarding general sick leave was calculated in which all days of sick leave were cumulated, regardless of the primary reason.

Because return-to-work, whether full-time, part-time or on a therapeutic basis (i.e. reduced workload), was considered to have a preventive effect on

the patient's recovery, the calculation of days of sick leave accounted for part-time work resumption and work on a therapeutic basis. Because OHS irregularly reported whether work was fully or partly lost, and whether work was done on a therapeutic basis, a second calculation was made in order to conduct a post-hoc sensitivity analysis. According to the second calculation, sick leave was considered to be 100% work loss, regardless of part-time work-resumption or work on a therapeutic basis.

Work-status

Next to days of sick leave, the patient's work status was gathered from the OHS data-report. Work-status was defined one week before the intervention, a half-year after the intervention, and one year after the intervention. The classification was as follows: 1) 100% return-to-work; 2) part-time return-to-work; 3) no return-to-work; 4) 100% disability pension as a result of back pain; 5) 100% disability pension not as a result of back pain. In cases where sick leave was reported not to be associated with back pain, work-status was defined as 100% return-to-work. Return to the original job or return on a therapeutic basis were both defined as full return-to-work, irrespective of whether or not a part-time factor was reported. Cases in which data on period 2 and post-treatment data on period 3 and/or 4 were available were included in the study.

Data Quality

Data quality was considered at three levels: good, moderate and poor quality. Data were qualified as good when the report provided information on work on a therapeutic basis and on the percentage of work resumption. Quality was judged moderate when data were provided on the percentage of work resumption, but not with respect to work on a therapeutic basis. Finally, the quality of the sick leave registration was considered as poor when no data were available on either part-time resumption of work or work on a therapeutic basis. In order to optimize data quality, one researcher (JZ) contacted the OHS to obtain more detailed information that had not been included in their reports.

Problem solving skills

Problem solving skills were measured by the Social Problem Solving Inventory (D'Zurilla, Nezu, & Maydeu-Olivares, 1994) in order to conduct a manipulation check. The social problem-solving inventory consists of 5 sub-scales: positive problem orientation (PPO), negative problem orientation (NPO), rational problem-solving skills (RPS), avoidant style (AS), and impulsive / carelessness style (IMP).

Treatment credibility

In order to check whether GAPS and GAGE conditions were equally credible to participants, treatment credibility was measured according to a procedure

similar to that described by Borkovec and Nau (1972) and applied in Vlaeyen et al. (1996). At the start of the intervention, patients rated on a Visual Analogue Scale (VAS), the extent to which they expected the treatment would help them return to work, resume daily activities, and cope better with their pain. The VAS ranged from "not at all" (0) to "very much" (100). Psychometric properties of this credibility questionnaire are good (Deville & Borkovec, 2000).

Baseline characteristics

Additional data were collected pre-treatment, regarding demographics, Distress and Risk Assessment Method (DRAM; Main, Wood, Hollis, Spanswick, & Waddell, 1992), functional disability (Roland Disability Questionnaire (RDQ); Roland & Morris, 1983), RDQ-level at baseline (Stratford, Binkley, Riddle, & Guyatt, 1998), pain intensity (McGill Pain Questionnaire-Dutch Version (MPQ); Van der Kloot, Oostendorp, van der Meij, & van den Heuvel, 1995), and job satisfaction in terms of work (VAG_work) and work conditions (VAG_cond); Van Veldhoven & Meijman, 1994).

STATISTICAL ANALYSES

Randomization check

Demographic variables, treatment credibility, data quality and data at baseline are compared between the two treatment conditions (GAPS and GAGE) by means of *t*-tests or Mann-Whitney *U*-tests (in case of non-normality) and by means of *Chi*-squared tests when ordinal variables are concerned.

Manipulation check

Paired *t*-tests within each condition are executed with regard to the SPSI-R (per subscale; D'Zurilla et al., 1994) in order to check whether the manipulation succeeded. The manipulation is considered successful when problem-solving skills are exclusively changed in the GAPS condition.

Outcome evaluation

Multiple linear regression analyses are conducted in order to test whether days of sick leave in period 3 and 4 independently, differ by treatment condition (using definitions: GAPS=1, GAGE=2). Next to condition, days of sick leave in the half-year before treatment (period 2) is added to the model as a covariate in order to control for pre-treatment differences. Finally, variables that are unequally divided between the conditions, despite randomization, are included in the analyses as covariates. Analyses are executed with respect to days of sick leave (a) as a result of back pain, and (b) in general. In each analysis, collinearity (Variance Inflation Factor > 10) and outliers (Cook's distance > 1; Studentized Residuals < -3 or > 3) are tested. Normal distribution of the residuals will be checked by means of scatter- and normal-probability plots.

Differences in work-status are assessed by means of *Chi*-squared tests

regarding work status one week before the intervention and 6 and 12 months after the intervention.

Sensitivity analysis

In order to control for the possible confounding influence of incomplete data, a multiple linear regression is executed on days of sick leave, calculated according to the second definition (part-time resumption and therapeutic work are 100% sick leave) as the dependent variable. It is checked whether outcomes are comparable to those initially found.

RESULTS

BASELINE CHARACTERISTICS

One-hundred-and-thirty-eight patients completed the selection procedure and 115 patients met the inclusion (Van den Hout et al., under review). Another 7 patients refused consent to retrieve data regarding sick leave and work status from the OHS. Furthermore, OHS retrospectively provided data on 84 of the remaining 108 patients included in the study, regarding period 2 (pre-treatment) and one or both of periods 3 and 4 (post-treatment). The group of 31 patients who dropped out for this part of the study (n : GAPS = 13; GAGE = 18) were comparable to the current sample regarding baseline characteristics ($p > .05$). Eventually, 45 patients had been randomly assigned to the GAPS condition and 39 to the GAGE condition. In table 1, baseline characteristics of the sample included in this part of the study are reported by condition. Overall, the average age is 40 years (range: 23 – 54 years of age), and 76% of all patients is male. Mean pain duration from the onset of the current pain episode is 1.5 years (range: 0.5 – 25 years), and from the first pain episode ever, 7.6 years (range: 6 weeks – 38 years). Four patients reported their current pain episode to be shorter than 6 weeks, however, all four of them had recurrent pain for more than 6 weeks and were therefore included in the study. The current pain episode was sub-acute (6 – 12 weeks) in 28% of the cases and chronic (> 12 weeks) in 67% of the cases. On average, patients were sicklisted for more than 8 weeks (60 calendar days; $SD = 57$ days; median = 44 days) when they entered the study. Mean pain intensity was 17.7 ($SD = 9.9$), which is higher than that of a group of Dutch physiotherapy patients who reported a mean pain intensity of 11.6 ($SD = 7.0$) just before the start of treatment (Van der Kloot et al., 1995). The mean score on functional disability, as measured by the RDQ (Roland & Morris, 1983), was 13.2 ($SD = 5.3$) and 31% of the patients had a score higher than 15 suggesting that they were moderately to highly disabled (Patrick et al., 1995; Roland & Morris, 1983). Overall, 50.6 % of the patients were classified as normal, 34.6% at risk, 6.2% distressed-depressive, and another 8.6% distressed-somatic as measured by the Distress and Risk Assessment Method (DRAM; Main et al., 1992)

Table 1. Baseline characteristics of patients in GAPS ($n=45$) and GAGE ($n=39$).

Variable	% GAPS	% GAGE	<i>p</i> ^a
Gender (%male)	73.3	79.5	.51
Education			
High	48.9	35.9	.23
Low/Medium	51.1	64.1	
Current medication use			
Yes	44.4	33.3	.30
DRAM-classification			
N	47.6	53.8	.60
R	33.3	35.9	
DD	9.5	2.6	
DS	9.5	7.7	
RDQ-level at baseline			
0-8	20.0	12.8	.05
9-16	40.0	66.7	
17-24	40.0	20.5	
Data quality			
poor	11.1	5.1	.47
moderate	57.8	53.8	
high	31.1	41.0	
	GAPS	GAGE	
Variable	<i>M (SD)</i>	<i>M (SD)</i>	<i>p</i> ^b
Age	40.3 (9.3)	40.8 (8.4)	.80
Pain duration since first pain episode (years)	8.7 (9.4)	6.4 (6.5)	.21
Duration current pain episode (years)	1.7 (4.5)	1.4 (2.6)	.69
Sick leave duration (weeks)	7.4 (6.1)	10.0 (9.9)	.14
Treatment credibility, PST / EDU (0-10)	6.9 (2.0)	8.0 (1.1)	.00
Functional disability (RDQ)	13.8 (5.4)	12.5 (5.2)	.28
Pain intensity (MPQ)	17.8 (10.6)	17.6 (9.0)	.93
Job satisfaction			
VAG_Work	19.2 (3.2)	18.4 (2.7)	.24
VAG_Cond	36.4 (4.6)	37.8 (4.1)	.16

Note. GAPS = Graded Activity + Problem Solving Therapy; GAGE = Graded Activity + Group Education. ^a Based on *Cbi*-square test. ^b Based on independent samples *t*-test. DRAM: Distress and Risk Assessment Method; N: normal; R: at risk; DD: distressed-depressive; DS: distressed-somatic.

RANDOMIZATION CHECK

With the exception of two variables, there were no significant differences between treatment conditions regarding baseline characteristics (Table 1; $p > .05$). The distribution of RDQ-scores at baseline (low / moderate / high) showed differences between the conditions ($p = 0.05$), in that, more patients in the GAPS condition had high RDQ-scores (17 – 24) compared to the GAGE condition where more patients had moderate RDQ-scores (9 – 16). Because high disability was regarded as a potential predictor of later work-related disability, this variable was added to the regression analyses as a covariate. The second variable that differed between treatment conditions was treatment credibility, with patients in the GAGE condition reporting higher expectations of Group Education than patients in the GAPS condition of Problem Solving Therapy (8.0 vs. 6.9 respectively). Treatment credibility was found an important predictor of cognitive-behavior therapy in pain patients (Kole Snijders et al., 1999) and was therefore added to the analyses as a covariate.

MANIPULATION CHECK

Table 2 shows the pre- and post-treatment scores on the SPSI-r. As expected, patients in the GAPS condition increased Positive Problem Orientation and Rational Problem Solving and decreased Negative Problem Orientation

Table 2. Means (*M*) and Standard Deviations (*SD*) for the Social Problem Solving Inventory-Revised at pre- and post-treatment.

	Pre-treatment	Post-treatment	
	$N_{\text{GAPS}}=45; N_{\text{GAGE}}=39$	$N_{\text{GAPS}}=43; N_{\text{GAGE}}=39$	
Measure	$M (SD)$	$M (SD)$	P^a
Social Problem Solving Inventory (SPSI-r)			
Positive Problem Orientation (PPO)			
GAPS	12.0 (3.5)	13.1 (3.0)	.01
GAGE	12.7 (2.9)	12.3 (3.1)	.34
Negative Problem Orientation (NPO)			
GAPS	8.0 (5.4)	6.1 (5.1)	.01
GAGE	8.3 (4.5)	7.5 (5.1)	.15
Rational Problem Solving (RPS)			
GAPS	41.8 (12.1)	48.8 (11.1)	.00
GAGE	42.7 (14.1)	44.6 (15.5)	.21
Impulsive/carelessness style (IMP)			
GAPS	11.4 (4.4)	10.1 (4.6)	.08
GAGE	12.7 (4.8)	12.2 (5.4)	.44
Avoidant Style (AS)			
GAPS	5.6 (3.6)	4.8 (3.0)	.17
GAGE	6.6 (3.8)	5.9 (4.5)	.14

Note. GAPS = Graded Activity + Problem Solving Therapy; GAGE = Graded Activity + Group Education. ^a Based on paired pre-post treatment *t*-tests.

(all p s < 0.05) and Impulsive Style (p < 0.1). Avoidance Style however, had not significantly changed. In the GAGE condition, problem-solving skills had not changed. The conclusion seems justified that, although patients in the GAPS condition did not significantly change avoidance style, problem-solving skills were exclusively changed in the GAPS condition in which the PST-intervention took place.

DAYS OF SICK LEAVE

Days of sick leave as a result of back pain are reported in table 3a. Although not statistically significant (p > 0.05), the number of days of sick leave in the half year before the intervention (period 2) are lower in the GAPS than in the GAGE condition (difference = 10.5 days). Figure 2 shows the developments regarding days of sick leave within each condition in the year before the intervention (periods 1 and 2) and in the year after the intervention (periods 3 and 4). In the first half-year after the intervention, days of sick leave decreased in both conditions. However, during the second half-year after the intervention, conditions developed in remarkably different manner: in the GAPS condition, days of sick leave continued to decrease, whereas in the GAGE condition, days of sick leave slightly increased again.

The results of multiple regression analyses are reported in table 4. In the first half-year after the intervention, treatment condition did not have a significant contribution to the regression models of days of sick leave. The RDQ-level at baseline was the only variable that had a significant contribution to the model at 6-month follow-up (p < 0.05). The regression analyses regarding days of sick leave during the second half-year after the intervention, on the other hand, show that treatment condition was the only variable that had a significant contribution to the model, in that, patients in the GAPS condition reported significantly less days of sick leave than patients in the GAGE condition (p < .05). Remarkably, none of the covariates had predictive value at this point in time. The results were comparable regardless of whether days of sick leave in general or as a result of back pain were regarded.

Sensitivity analyses were conducted in order to control whether data quality (incomplete data) influenced outcomes of the effect evaluation. Regression analyses were therefore repeated, but this time sick leave was defined as 100% work loss, regardless whether work was partly resumed or took place on a therapeutic basis. Analyses did not show any marked differences and endorsed the previous finding that treatment condition was the only variable successfully predicting days of sick leave as a result of back pain during the second half-year after the intervention. When sick leave in general was regarded, RDQ-level at baseline had a significant contribution to the prediction of days of sick leave during the second half-year after treatment (standardized β = .218; p < .05) next to treatment condition (standardized β = .287; p < .05).

Table 3. a) Means (*M*) and Standard Deviations (*SD*) for days of sick leave as a result of back pain by condition in period 1, 2, 3, and 4.
b) Work status at pre-treatment and at 6- and 12-months follow-up, by condition.

Condition							
a) Days of sick leave, <i>M</i> (<i>SD</i>) *	GAPS			GAGE			<i>p</i> ^a
			<i>n</i>			<i>n</i>	
Period 1: 1 – 0.5 year pre-treatment	8.3	(16.9)	45	10.4	(22.8)	39	.99
Period 2: 0.5 – 0 year pre-treatment	30.8	(24.7)	45	41.3	(27.8)	39	.09
Period 3: 0 – 0.5 year post-treatment	24.5	(31.3)	44	34.2	(44.3)	39	-
Period 4: 0.5 – 1 year post-treatment	18.5	(36.4)	44	37.9	(50.1)	39	-

	GAPS			GAGE		
b) Work status, % (absolute number)	Pre	FU1	FU2	Pre	FU1	FU2
1) 100% return to work	8.9 (4)	75.6 (34)	86.4 (38)	20.5 (8)	71.8 (28)	66.7 (26)
2) part-time return-to-work	28.9 (13)	13.3 (6)	2.3 (1)	33.3 (13)	7.7 (3)	2.6 (1)
3) no return-to-work	62.2 (28)	6.7 (3)	2.3 (1)	46.2 (18)	17.9 (7)	10.3 (4)
4) 100% disability compensation as a result of back pain	0	2.2 (1)	4.5 (2)	0	2.6 (1)	17.9 (7)
5) 100% disability compensation not as a result of back pain	0	2.2 (1)	4.5 (2)	0	0	2.6 (1)

Note. GAPS = Graded Activity + Problem Solving Therapy; GAGE = Graded Activity + Group Education; *Means are based on all available data in the period concerned; Pre = one week before treatment, FU1 = 6 months after treatment, FU2 = 12 months after treatment; Pre-treatment: $N_{\text{GAPS}} = 45$, $N_{\text{GAGE}} = 39$; ^a Based on Mann-Whitney *U* test.

Figure 2. Days of sick leave as a result of back pain, by condition.

Note. ga: Graded Activity; pst: Problem-solving Therapy; edu: Group Education; period 1: one to half year pre-treatment; period 2: half-year pre-treatment; period 3: half-year post-treatment; period 4: half to one year post treatment; I: Intervention.

WORK STATUS

In table 3b, work status at pre-treatment, and at 6- and 12-month follow-up is reported per treatment condition. In the week prior to the start of the intervention, most patients were reported to be on the sick list. Twelve patients had a full return to work between the selection procedure and the start of the intervention. In general, work status of patients in the GAPS condition is less favorable, in that, relatively more patients did not return to work in this condition. The Pearson *Chi-Square* test, however, did not reveal a significant difference between treatment conditions at pre-treatment. Six months after the intervention, most patients (74%) had a full return to work in both conditions. In 4 cases (1 GAPS and 3 GAGE) work status was defined as 100% return to work because sick leave was not caused by back pain. At this time, relatively more patients in the GAGE condition did not return to work as compared to the GAPS condition (18% vs. 7% respectively). Differences between conditions were not significant ($p > 0.1$). Finally, work status one year after the inter-

Table 4. Results of Multiple Linear Regression Analyses in the first and second half-year after the intervention (periods 3 and 4 respectively), for days of sick leave: a) as a result of back pain, and b) in general.

Dependent Variable	<i>n</i>	Adj. R^2	Independents	Standardized β
Days of sick leave period 3:				
a) as a result of back pain	83	.069*	Treatment condition	.117
			Days of sick leave period 2	.195†
			RDQ-level at baseline	.229*
			Treatment credibility	.006
b) in general	81	.066†	Treatment condition	.117
			Days of sick leave period 2	.151
			RDQ-level at baseline	.260*
			Treatment credibility	-.011
Days of sick leave period 4:				
a) as a result of back pain	83	.038	Treatment condition	.247*
			Days of sick leave period 2	.080
			RDQ-level at baseline	.156
			Treatment credibility	-.082
b) in general	81	.077*	Treatment condition	.284*
			Days of sick leave period 2	.113
			RDQ-level at baseline	.179
			Treatment credibility	-.106

Note. GAPS = Graded Activity + Problem Solving Therapy; GAGE = Graded Activity + Group Education; period 2: half-year before intervention; period 3: first half-year after intervention; period 4: second half-year after intervention. † $p < 0.1$; * $p < 0.05$.

vention was evaluated. Data of one patient in the GAPS condition were not available at 12-month follow-up because he changed his job. Furthermore, work status was defined as 100% return-to-work in 8 cases (4 GAPS and 4 GAGE) because sick leave had causes other than back pain. Figure 3 illustrates the differences between conditions. The percentage of patients with 100% return to work increased in the GAPS condition from 6- to 12-month follow-up (from 76% to 86%), whereas the GAGE condition shows a slight decrease regarding this variable (from 72% to 67%). At one-year follow-up more patients in the GAGE condition received disability pensions, as compared to patients in the GAPS condition (21% vs. 9% respectively). This may indicate that patients in the GAPS condition were more successful in their job resumption as compared to patients in the GAGE condition. However, differences between conditions, although of clinical importance, were not statistically significant ($Cbi^2 = 6.88$; $p = .142$). Nevertheless, developments regarding work status were more favorable in the GAPS condition regarding the data at one-year follow-up.

Figure 3. Work status at 12-month follow-up, by condition.

Note. ga: Graded Activity; pst: Problem-solving Therapy; edu: Group Education; DP: Disability pension.

DISCUSSION

The aim of the current study was to investigate the supplemental value of problem-solving therapy when added to behavioral graded activity, with regard to days of sick leave and work status in employees with low back pain. It was concluded that problem-solving therapy had supplemental effects, which became apparent 12 months after termination of the program. First, patients who received problem-solving therapy in addition to graded activity were reported to have 50% less days of sick leave during the second half-year after the intervention, than patients who did not receive additional problem-solving therapy. Moreover, work status in these patients was more favorable in that more patients had a full return-to-work and fewer patients received disability pensions compared to patients who only received graded activity. Several checks to control for possible sources of confounding revealed that the manipulation was successful, as well as the randomization, with the exception of the level of functional disability and treatment credibility. Furthermore, sensitivity analyses did not show any marked differences when days of sick leave were computed according to an alternate definition in which sick leave was defined as 100% work loss, regardless of partial or therapeutic work resumption.

The sample described in the current study was a smaller part of the origi-

nal sample of 115 patients, as described by Van den Hout et al. (under review). Seven patients did not consent to data-retrieval from the OHS (3 GAPS and 4 GAGE condition). In another group of 24 patients, registration of sick leave by the OHS was missing or incomplete, in some cases because employees changed their jobs and could not be contacted. In case of incomplete registration of sick leave, when available, medical files of OHS were used to substitute missing data. Although unlikely, slight selective dropout of data cannot be ruled out completely. On the other hand, baseline characteristics were comparable between the current sample and the group of patients whose data on sick leave were not available. Neither can we rule out that the dropout was more favorable for one condition than the other, although the equal distribution of baseline characteristics between conditions may indicate that the original randomization was not violated.

Outcomes regarding sick leave and work status are in line with the results of the first part of this study in which self-report measures are reported (Van den Hout et al., under review). In accordance with outcomes of the current study, the supplemental value of the GAPS condition became clear in the long-term regarding functional disability. Not only the patients in the GAPS condition had significantly less functional disability at 12-month follow-up. Also, relatively more patients made a clinically relevant change regarding RDQ-level.

Current results on days of sick leave are different from data obtained by means of a cost diary. In the latter no marked differences between conditions were found regarding self-reported days of sick leave (Van den Hout et al., under review). Although moderate to high correlations between both data sources (cost diary and OHS) were found at 6- and 12-month follow-up (Pearson $R = .78$ and $.84$, respectively; $p < .001$), self-report measures such as the cost diary, are more vulnerable to a number of biases and misinterpretations on the part of the subject. For example, subjects may have reported sick leave due to causes other than back pain, whereas the cost-diary was explicitly meant to report only those costs resulting from back pain. Regarding OHS-data, it was possible to control for possible shortcomings in the report to a reasonable degree. First, analyses were executed regarding sick leave in general, and second, to control for inequalities in data quality, sensitivity analysis was performed in which data were calculated according to definition 2. The consistent findings of these additional analyses may further contribute to the reliability of OHS-data and outcomes reported in the current study.

Remarkably, in the half-year following the intervention, the level of perceived functional disability at baseline is predictive of days of sick leave, next to and regardless of days of sick leave in the half-year preceding the intervention. From this, it seems to be important to screen employees regarding the level of self-reported disability (Hurley, Dusoir, McDonough, Moore, & Baxter, 2001; Linton & Hallden, 1998). At 12 months follow-up, treatment condition was a significant predictor of the number of days of sick leave. Teaching problem-solving skills in addition to gradually increasing activity levels through

operant principles successfully contributes to the prevention of sick leave and unfavorable work status in the long term. This would suggest that problem-solving skills indeed play a role in disability as a result of low back pain, as was suggested in two recent cross-sectional studies (Shaw et al., 2001; Van den Hout et al., 2001b). For example, Shaw et al. (2001) suggest that a secondary prevention program could use problem solving skills to help patients to overcome psycho-social, vocational and ergonomic risk factors. The current prospective study is the first to demonstrate the preventive effects of a problem-solving intervention regarding sick leave in the long term. The results indicate that behavioral graded activity may have an important contribution to the effects in the first half-year after the intervention. However, only in employees who also received the problem solving intervention the positive developments continued during the second half-year after the intervention. These long-term effects result in a more favorable work status than if the problem solving intervention is not received. As suggested by Shaw et al. (2001), effects in the long term may indeed be related to a more positive orientation, not only to the pain problem, but even more so to problems in general, either in the workplace or at home.

The sample selected in this study consisted in the greater part of employees with chronic and/or recurrent low back pain, but also had been able to continue to work despite pain for quite some time. It would be more appropriate, therefore, to consider duration of sick leave rather than duration of pain when indicating the target population for the intervention described in this study. It would be interesting to select employees with recurrent short episodes of back pain to prevent them from work loss when back pain becomes less manageable in future situations. These suggestions are supported by significant preventive effects of a cognitive-behavioral group intervention with regard to sick leave in a non-patient population who experienced 4 or more episodes of intense spinal pain during the last year, but had not been out of work more than 30 days (Linton & Ryberg, 2001a).

The introduction of problem solving techniques, especially in a population of employees at risk for developing chronic disability, may be a promising innovation with respect to early intervention at the workplace in primary care settings. Indeed, Mynors Wallis, Davies, Gray, Barbour, and Gath (1997) found problem solving therapy to be effective regarding disability days and days off work in patients with emotional disorders in primary care. However, before introducing these problem-solving techniques some conditions need to be fulfilled. In the current study, cooperation between the occupational health physician, the general practitioner and the multidisciplinary rehabilitation team was initiated regarding individual cases. To achieve a consistent plan agreed by all parties, one party (e.g. the occupational therapist) initiates contacts and communicates about therapy to parties outside the rehabilitation center. First, a plan for return-to-work must be agreed upon between the occupational health physician, the employer, the employee and the rehabilitation team.

Second, it is imaginable that problem-solving techniques yield the best effects when broadly implemented, not only within the rehabilitation setting, but also at the workplace and at home. When employees find results in their new way of dealing with work-related problems, it may facilitate the use of problem solving in future situations.

A limitation of the current study may be that the study design did not include a treatment-as-usual condition, and therefore does not allow drawing conclusions about efficacy of the entire program. Nevertheless, graded activity was previously shown to be more effective than treatment-as-usual (Lindström et al., 1992; Linton & Bradley, 1992; Linton et al., 1989). From this, it may be cautiously suggested that the PST-ingredient added supplemental value to the effect of graded activity.

Finally, the supplemental value of problem solving only occurred at long term. This may indicate that the generalization of problem-solving skills to daily life needs time and eventual effects would not have been detected when measurements were restricted to the short-term. In line with this, Linton and Van Tulder (2001b) found that almost all studies included in their review had relatively short follow-up periods, and suggested that the full preventive effect may not yet be visible. From the current findings it is recommended that long-term follow-up measurements be included, especially when behavioral change and maintenance are important aims of the intervention.

In conclusion, problem-solving therapy turned out to be an effective treatment in low back pain. It showed favorable effects in the course of sick leave in the year following the intervention. The intervention may alter the course of work disability and even protect employees against new episodes of sick leave. A logical continuation of the current study would be to examine cost-effectiveness and to explore possibilities for implementation of problem-solving techniques in occupational health care.

CHAPTER 8

General Discussion

*Employ the best idea found,
And watch what happens most precisely,
Making sure by looking 'round,
That things are working out quite nicely.*

*From: Winnie-the-Pooh on Problem Solving,
Allen & Allen, 1995*

GENERAL DISCUSSION

Does problem-solving play a role in the development of chronic low back pain and can problem-solving therapy add to the efficacy of the existing behavioral rehabilitation programs of non-specific low back pain? That is the main question of this thesis. In the introductory section, five research questions are formulated that are directly or indirectly related to this main question. First, the question is posed whether problem-solving therapy is generally an effective treatment in the field of mental health care and behavioral medicine. In chapter 2, a systematic review is described that addresses this question. Second, the role of non pain-related failure on pain and pain behavior is studied in two experimental studies (chapter 3 and 4), as it is hypothesized that failure experiences influence the experience of pain, pain tolerance and pain avoidance. Furthermore, the role of negative affectivity, either as a mediator or as a moderator, is examined. The next question refers to the role of several pain-related variables, daily stress and problem-solving skills in functional disability of low back pain patients. Answers to these questions are examined by means of a cross-sectional study (chapter 5). Finally a randomized clinical trial is conducted to answer the question whether problem-solving therapy has supplemental value when added to a more physically oriented graded activity approach. This question is addressed in two parts: First, regarding self-report and observational measures (chapter 6), and second, regarding data from the occupational health services (OHS) on sick leave and work status (chapter 7). In the current chapter, conclusions from the foregoing chapters will be brought together and discussed in the light of methodological strengths and flaws, outcomes, and external validity. Moreover, new research questions following from the current thesis will be summarized. Finally, a general consideration will speculate on the role of problem solving, relative to pain-related fear, in low back pain.

THE EFFECTS OF PROBLEM SOLVING THERAPY (PST) WITHIN THE FIELD OF MENTAL HEALTH CARE AND BEHAVIORAL MEDICINE. (A SYSTEMATIC REVIEW)

Problem-solving therapy is an effective therapy within the field mental health care and behavioral medicine, immediately after treatment and in the long term. About 75% of all comparisons included in the systematic review concluded that PST was more effective than the reference treatment (other treatment or no-treatment control). Moreover, in 80% of the studies that included a follow-up next to a post-treatment measurement, PST turned out to be more effective than the reference treatment in the long term. PST was especially effective regarding problem-solving skills, quality of life, domain specific outcomes, and outcomes referring to general improvement. Surprisingly, success rates were lowest with regard to outcomes of negative affect, especially in the long term (25% of the comparisons). It was suggested that failure to find effects with regard to negative affect was due to floor-effects in non-depressed samples.

This was confirmed by the fact that success rates increased when clinically depressed samples were concerned.

In our review we were restrictive regarding which trials to include. First of all we included only the randomized controlled trials. We realize that by doing so, we might have excluded studies in which the PST protocol was properly used and outcomes rightly measured, but the design did not fit our inclusion criteria. Another restriction took place by including only those studies describing that problem-solving skills were actually trained. Therefore, the conclusion drawn above refers to PST interventions in which problem-solving skills are explicitly trained. In cases the problem-solving theory was merely used as a framework (e.g. Toseland, Blanchard, & McCallion, 1995; Johnson & Greenberg, 1985; Halford, Goodall, & Nicholson, 1997), the study was not included. For example, in a study by Toseland et al. (1995), a problem-solving intervention for caregivers of cancer patients was evaluated. The oncology social worker utilized the problem-solving model within the context of a supportive helping relationship, whereas patients were helped to reflect on how they coped with three pressing problems and what they wanted to change. A special concern in this 'framework approach' might be whether coping skills applied in a few particular problems, would generalize to other problems.

At least two variables turned out to be influential regarding positive effects of the PST interventions. First the methodological quality of the studies was positively related to favorable outcomes of the PST intervention. Apparently, the chance of finding positive effects depends on internal and external validity of the study, and consequently, the chance of making a type II error (inaccurately accepting the hypothesis that H_0 is true) increases as the quality is low. A second variable that was shown to be influential regarding the efficacy of the PST intervention was the number of sessions. More favorable outcomes were found when the number of sessions was more than five, and a success rate of almost 100% was reported when the intervention consists of 10 or more sessions. A very plausible explanation might be that skills training needs time and rehearsal, which can not be provided in only a few sessions, at least not when behavioral and mental health problems are regarded.

Finally, publication bias may not be excluded as a possible treat to the internal validity of the review. Publication bias is defined as the tendency, which appears to favor trials with positive results, in that these are published preferentially over those, which failed to find significant ("positive") results (Jadad, 1998). Keeping this in mind, we must reckon with the risk of overestimating the effects of PST.

Generally, PST seemed to be an effective therapy, which made it worthwhile to investigate whether PST could be of importance within low back pain patients. Next, the question was addressed what mechanisms could possibly explain the working of problem solving. In line with this, it would be interesting to find out more about the influence of failure on pain, and in particular failure to solve non-pain-related problems. Levine, Krass and Padawer

(1993) found that non-pain-related failure increased pain. At the same time, however, their study had some methodological problems. Replication of the Levine-study therefore seemed advisable.

DOES NON-PAIN-RELATED FAILURE HURT? (2 LABORATORY EXPERIMENTS)

In healthy volunteers, non-pain-related failure hurts and leads to a decrease of pain tolerance. In chronic low back pain patients, however, non-pain-related failure resulted in lowered pain avoidance, but only when Negative Affectivity was low.

Outcomes of the first experiment (chapter 3) were in line with those from Levine et al. (1993). In a healthy female population, non-pain-related failure (negative feedback on an intelligence test) resulted in increased pain report and lowered pain tolerance. The hypothesized mediating role of negative affect (state) and the moderating role of negative affectivity (trait) were not confirmed. Although the direction of findings regarding the influence of failure was consistent with those from the Levine-study, outcomes were less convincing. This discrepancy could possibly be explained by the methodological modifications we made to the Levine-design. Still another explanation may be found in the impact of the failure manipulation, which may have been more powerful in the Levine-study. But apart from this, the effects of failure may indeed be very unpredictable, depending on personality traits or particular mood states. Subsequently, negative affect (being either depression, fatigue, anxiety, or anger) may have influenced outcomes very differently.

Based on the results of the first experimental study in healthy volunteers, it was hypothesized that in chronic low back pain patients, helplessness resulting from feelings of failure and frustration about not having control over pain consequences and life in general, might have started to influence pain and pain behavior. In that way, failure experiences may help to explain how chronic pain patients got stuck in a vicious circle of chronic pain and disability. Surprisingly, although results were only statistically significant with regard to pain avoidance, the overall pattern in CLBP patients was opposite to the one expected and not comparable to what we found in healthy controls. It seemed as if CLBP patients were triggered to give a better performance after failure feedback in that they avoided less, reported less pain and had increased pain tolerance. How can the different findings of the 2 experiments be explained? There may be several explanations.

a) The study sample

In the first experiment healthy subjects were involved, who did not have pain at that moment, and in whom acute pain was induced. In the second experiment, on the other hand, chronic pain patients were involved in whom pain was elicited in addition to the already existing pain. Besides the fact that the current pain influenced the performance on the lifting task, the increase of pain during the lifting task might not have come as a complete surprise to the

CLBP patients. Moreover, in chronic pain, the experience of failure may not have had an overwhelming impact, as other variables, such as personality traits, coping style and attributions, have become more important in this phase. Indeed, we found Negative Affectivity to moderate the results regarding pain avoidance. It turns out that not all patients had a better performance after failure feedback, but specifically those not affected by Negative Affectivity. In those cases patients reported Negative Affectivity, feedback did not seem to matter, which might be an even more worrying scenario, as patients were neither encouraged by success nor by failure, as if they had surrendered themselves already to the fact that "things just are as bad as they are".

b) The experimental pain induction

The pain induced during the lifting task might be far more familiar to the CLBP patients, than pain induced by cold water was for a non-pain sample. Explanations may be found in the fact that healthy controls did not know what they could expect, this might have aroused anxiety or curiosity, and attention would be allocated to pain. Attention was found to mediate the influence of anxiety on pain, in that pain increases when allocated attention to (Arntz & de Jong, 1993; Arntz, Dreessen, & De Jong, 1994; Arntz, Dreessen, & Merckelbach, 1991). The lifting task, on the other hand, might have been familiar and CLBP-patients might have known what they could expect, such that attention was also allocated to stimuli other than pain. Different pain induction may have hampered comparability of the studies, nevertheless, for 2 reasons we chose the lifting task in stead of the cold pressor in the second experimental study. First, the introduction of another source of pain (in the hand) in CLBP patients might have complicated our findings, as mechanisms like attention and attribution would involve two different sources. Second, the lifting task seemed to have more ecological validity. Especially because we also wanted to investigate the influence of pain-related fear, the lifting task was expected to be more relevant to the CLBP patients than the cold pressor task.

c) The effects of the manipulation

Although the manipulation check suggested that the manipulations succeeded in both experiments, emotions were differently affected. The manipulation in healthy controls, as hypothesized, resulted in increased negative affect. In the second experiment, however, negative affect decreased in both failure and success conditions, which might have been crucial as negative affect was thought to mediate between failure and pain. Negative affect decreasing after failure may imply that the manipulation did not impress patients enough to result in compatible mood. Another explanation may be that the composite score from the POMS, a common scale to measure negative state affect, was not sensitive enough to more specific negative emotions as anger and frustration. Next the role of negative state affect will be further explored.

In neither of the experiments, we found negative state affect to mediate

between failure and pain. The hypothesis was that failure feedback would result in depressed feelings, what would result in increased pain (Brown, 1990; Wade, Price, Hamer, Schwartz, & Hart, 1990). Suppose failure feedback did bring about negative mood, but the character of this mood was not depressed but rather angry and frustrated. In a recent study by Janssen, Spinhoven and Brosschot (2001), an angry state was induced by computer task and harassing comments in healthy subjects. Pain and pain tolerance were operationalized by a cold pressor task. Results showed some unexpected findings, in that subjects in the "harassment-condition" had increased pain tolerance compared to subjects in the neutral condition. These results are in line with our findings in CLBP patients following failure feedback. Increased tolerance and decreased pain avoidance, in line with the results of Janssen et al. (2001), may have resulted from unintentionally induced anger and frustration, that in turn resulted from the negative feedback regarding the social empathy test. Why anger and frustration cause increased pain tolerance is still subject to debate. One possibility may be increased cardiovascular reactivity, although findings are contradictory to this respect (Caceres & Burns, 1997; France & Stewart, 1995; Janssen et al., 2001). Another explanation may be motivational, in that dissatisfaction and frustration about the earlier task, motivates patients to improve on the second cold pressor task and to endure despite pain.

Some patients in the second experimental study, namely those reporting low Negative Affectivity, showed 'perseverance' after failure. Although perseverance may be a very healthy coping style and normally very useful to solve complex and difficult problems, in the long term perseverance may also have some negative consequences. By ignoring emotional and physical boundaries, pain patients may start to wear themselves out, and instead of improving their functional status, they will increase disability and the probability that new failure experiences will occur. The vicious circle of perseverance, failure and increased disability may feed the feeling that the pain problem is out of control (Aldrich, Eccleston, & Crombez, 2000), which may hypothetically be the moment that perseverance changes into passivity and helplessness.

In line with the foregoing, and as an extension to the fear-avoidance model (e.g. Vlaeyen, Kole Snijders, Boeren, & van Eek, 1995b; Vlaeyen & Linton, 2000), Hasenbring, Marienfeld, Kuhlendahl and Soyka (1994) introduced the avoidance-endurance model. Contrary to 'avoiders', 'endurers' ignore their pain, use cognitions which help them to endure (e.g. "get yourself together", "don't be such a baby"), as well as endurance strategies (e.g. take no breaks despite pain, go on with sports at the same level). In stead of inactivity, they show hyperactivity, which may in the long-term lead to chronification of pain. Although still quite hypothetical, the distinction between patients who avoid and patients who endure may be recognized in clinical practice and both groups may ask for divergent approaches. Questions for future research might be directed towards these hypothesized distinct groups, e.g. do they exist? And if so, how are they characterized? Moreover, the role of endurance

(cognitions and strategies) in pain, may be addressed in prospective designs.

The foregoing hypotheses, although plausible, may question the external validity of the experimental studies on failure and pain, described in chapter 3 and 4 of this thesis. There are two main features in both experiments that mainly determine the external validity of the studies. These are the failure manipulation and the pain induction. The manipulation was successful in that respect that failure feedback leads to lowered predictions of future achievement, whereas success feedback leads to elevated predictions of future achievement. Regarding mood, however, CLBP patients did not show mood compatible with the outcomes of the social empathy test. This may implicate that the subjects did not find failure in social empathy important, or, again another explanation, the failure experience did not resemble real-life failures. It is imaginable that 'real-life failure' is far more implicit, regarding the broad spectrum of daily problems, but in particular elicited by the feeling that one is disabled and invalidated as a result of pain. Moreover, real-life failure may be experienced repeatedly and continuously, whereas in the experimental study failure feedback was repeated no more than 3 times. With regard to pain induction, both the cold pressor and the lifting task were found to be valid methods to induce laboratory pain. Nevertheless, in both experiments, pain avoidance was operationalized in ways never validated before. Therefore, outcomes must be cautiously interpreted.

One of the most important differences between the laboratory pain and chronic low back pain is the fact that pain was acute in the former and could be escaped from whenever the subject wanted to. Chronic low back pain can not be escaped from by simply quitting a task but usually appears to be persistent and unpredictable. To that respect, the lifting task might have had the highest ecological validity as pain was induced by a daily activity next to the already existing pain.

THE ROLE OF PAIN-RELATED FEAR. (LABORATORY EXPERIMENT AND CROSS-SECTIONAL STUDY)

Pain-related fear increases pain report and decreases pain tolerance in CLBP patients, and is related to functional disability in non-specific low back pain patients with a recent onset of work disability.

The fear-avoidance model of pain hypothesizes that pain-related fear is an important predictor of pain avoidance and disability. In Crombez, Vlaeyen, Heuts and Lysens (1999), it was found that pain-related fear was more disabling than pain itself. One of the research questions was whether the findings of Crombez et al. (1999) could be replicated. Two studies within this thesis investigated the role of pain-related fear in more detail. First, in the experimental study on failure experiences in CLBP patients, pain-related fear was added to the hypothetical model of pain report, pain tolerance and pain avoidance. It was found that CLBP patients with high levels of pain-related fear, measured by means of the TSK (Kori, Miller, & Todd, 1990), reported more

pain and tolerated the lifting task shorter than patients with low pain-related fear. Most influential in the model was the pain before lifting (pre-lifting pain), in that higher pre-lifting pain resulted in increased pain report during the lifting task and decreased pain tolerance. When pre-lifting pain was controlled for, pain-related fear did not predict pain avoidance, probably because pre-lifting pain explained most of the variance.

On the other hand, the cross-sectional study in this thesis (chapter 5) found pain catastrophizing, here measured by the Pain Catastrophizing Scale (Sullivan, Bishop, & Pivik, 1995), to be the most important predictor of functional disability. Although pain intensity, was running up as second predictor in the model, the functional disability of workers with a new episode of sick leave as a result of back pain, seems to be most importantly influenced by the level of pain catastrophizing. In a comparable design, Gheldof et al. (2001) found pain intensity to have a stronger relation to functional disability than pain-related fear (TSK), in a sample of 1118 employees with low back pain. Although both studies (Gheldof et al., 2001; Van den Hout, Vlaeyen, Heuts, Sillen, & Willen, 2001b) did not find pain intensity equally important, its role may not be underrated in these kind of pain populations, who are still at work despite pain. It might be hypothesized that pain catastrophizing, pain-related fear, and hypervigilance (i.e. increased attention towards the source of the threat, in *casu* bodily sensations) mediate the influence of subjective pain to disability (Peters, Vlaeyen, & van Drunen, 2000; Vlaeyen et al., 1995b; Vlaeyen & Linton, 2000). This hypothesis, however, may better be tested in prospective or experimental designs.

For clinical practice, it may be important that valid predictors are detected, which are changeable without invasive therapy, and can mainly be controlled by the patient himself. To that respect, pain-related fear offers a number of possibilities to intervene, such as operant, respondent, and cognitive techniques. One new development may be considered in particular, that is graded exposure in vivo with behavioral experiments (Vlaeyen, de Jong, Geilen, Heuts, & van Breukelen, 2001). Remarkably, this technique aims especially at those patients who report a high level of pain-related fear. Not only the avoidance of personally relevant activities is addressed, but also the irrational idiosyncratic thoughts that go along with it. Behavioral experiments and thought challenging are important techniques applied in this treatment. With regard to the current clinical trial, it would be interesting to investigate whether the level of pain-related fear predicted treatment outcome. The more because graded activity not explicitly included the specific cognitive behavioral techniques included in graded exposure in vivo, the former might have been less successful in changing very high levels of pain-related fear.

THE ROLE OF PROBLEM SOLVING SKILLS. (CROSS-SECTIONAL STUDY)

A negative problem orientation is positively related to functional disability in non-specific low back pain patients. Problem solving skills do, however, not add

predictive value to the model of functional disability, when pain intensity and pain catastrophizing are included as predictors. Furthermore, problem-solving skills do not moderate the effects of daily stress on functional disability.

The cross-sectional study described in chapter 5 (Van den Hout et al., 2001b) could not convincingly subscribe the idea that problem solving and pain disability, are directly related to one another. Some methodological features may have interfered with the conclusions. First, the study is cross-sectional, and must therefore be restrictive regarding the directionality of the results. Second, problem-solving skills were measured by means of the SPSI-R (D'Zurilla, Nezu, & Maydeu-Olivares, 1997). The SPSI-R is a self-report measure and may particularly assess how people think that problems should be solved, but may not necessarily assess how they actually behave when confronted with a problem. Generally, the lack of an observational measure that assesses real-life problem solving, is one of the needs of research regarding problem-solving theory and therapy, and development of a measure as such would be of great value. Third, as we suggested in the discussion of chapter 5, functional disability, as measured by the QBPDS, may be too specific as a dependent variable, as it is rather quality of life in general that is assumed to be affected. Recently, Shaw, Feuerstein, Haufler, Berkowitz and Lopez (2001) published a comparable cross-sectional study, studying the influence of problem solving skills on functional health status, as measured by the SF-12 (Ware, Kosinski, & Keller, 1995). With respect to the main effects of problem solving on functional status, both studies (Shaw et al., 2001; Van den Hout et al., 2001b) concluded that rational problem solving skills, apart from problem orientation, did not play a role of significance. Although modest, Van den Hout et al. (2001b) found that a negative problem orientation was related to increased functional disability. Van den Hout et al. (2001b) did not find an interaction between daily stress and problem solving. Shaw et al. (2001), however, found two- and three-way-interactions between several problem-solving scales, symptom severity, and pain duration. It was concluded that problem solving moderated the influence of symptom severity on functional health status.

These outcomes (Shaw et al., 2001; Van den Hout et al., 2001b) may indicate that the role of problem solving is not straightforward but may especially be sought in indirect pathways, protecting a person against negative influences on functional status and quality of life. Moreover, it may be noticed that not the skills, but the orientation towards problems mattered. Although it may be obvious that problem orientation and problem-solving skills are intertwined, remarkably, we did not find studies explicitly unraveling this relationship. It may be quite possible that the experience of success, while using rational problem-solving skills, increases a positive orientation towards future problems. On the other hand, starting with an optimistic point of view and having confidence in one's capacity to solve problems may improve subsequent actions to that respect. In line with this, attributional biases ("effort won't help me to solve this problem") were found to lead at ineffective attempts of problem

solving coping (Baumgardner, Heppner, & Arkin, 1986; Heppner, Reeder, & Larson, 1983). Probably, problem orientation and rational problem solving skills depend on one another and therefore therapy protocols should preferably incorporate both to yield best effects. Indeed, in a dismantling investigation by Nezu and Perri (1989), the inclusion of training of problem orientation added significantly to the overall efficacy of PST. More research is needed to unravel the direct or indirect role of problem solving in pain and disability.

In the foregoing, we examined the hypothesized role of problem solving with regard to the maintenance of pain and disability. Although rational problem solving skills were not related to disability, problem orientation was. As problem orientation is a cognitive construct, this may motivate the use of cognitive behavioral techniques to change problem orientation in a positive direction. For instance, when a negative thought prevails that one is not able to bring problems to a good ending, an effective way to change this irrational thought may be to expose the patient to the actual problem. While experiencing that he or she is able to solve the problem satisfactory, alternate and more positive thoughts may be formulated. Problem solving therapy is one of the methods exposing patients to the problems of everyday life, and training them in effective problem solving. Conclusions of a randomized clinical trial (Chapter 6 and 7) will be discussed next.

PROBLEM SOLVING THERAPY AND GRADED ACTIVITY IN NON-SPECIFIC LOW BACK PAIN. (A RANDOMIZED CLINICAL TRIAL)

The problem solving therapy has supplemental value when added to graded activity: One year after the intervention, patients who received both graded activity and problem solving therapy had less functional disability than patients who received graded activity and group education. Moreover, in the experimental group, more patients had a clinically relevant improvement regarding functional disability. There was no supplemental value with regard to medical consumption and negative affect. High-educated patients did not gain more from the supplemental PST than low-educated patients did.

Finally, the problem solving therapy had a supplemental value with regard to days of sick leave and work status. After one year, patients receiving PST had less days of sick leave and relatively more patients had a 100% return-to-work. The conclusions above are supported by the fact that several sensitivity analyses, such as intention-to-treat and per-protocol analyses, had comparable outcomes. The results also fit into the findings of the systematic review of problem-solving therapy, in that the supplemental value was apparent in the long term and the effects were referring to increased quality of life (functional disability), problem solving skills and general improvements (days of sick leave and work status). These domains were identified as yielding highest success rates in the systematic review.

Next, comments regarding methodological issues, external validity of the study, and limitations will be discussed.

Table 1. Methodological quality of the RCT

Methodological Criteria List		Comment
a. Specification inclusion/exclusion criteria.	+	Criteria are specified in methods-section
b1. Randomization treatment allocation *	+	Randomization scheme is computer-generated
b2. Concealment of treatment allocation *	+	Randomization scheme only known to logistic planner, who is not involved in inclusion procedure
c. Similarity of baseline characteristics *	+	Groups are comparable regarding most important prognostic factors: age, sex, main outcome measures
d1. Description experimental intervention(s)	+	Experimental intervention described in methods-section
d2. Description control intervention(s)	+	Control intervention described in methods-section
e. Co-interventions comparable *	?	Not described regarding intervention period
f. Compliance rate *	+	Compliance to the interventions is described and acceptable
g. Blinding of patients *	+	Patients could not know which condition was experimental
h. Blinding of outcome assessor *	+	Yes, described in methods-section
i. Sort of outcome measures	+	Self-report, behavioral observations and sick-leave registration
j. Withdrawal/drop-out rate *	+	2% at post-treatment, 17% at 12-month follow-up
k1. Description short-term follow-up measurement	+	Yes, post-treatment results
k2. Length of long-term follow-up measurement *	+	Adequate length of follow-up: 12 months
l. Identical timing of the outcome assessment	+	Yes
m. Adequate sample size for each group	+	N(GAPS) = 58; N(GAGE) = 57 (>25 subjects per group)
n. Intention-to-treat analysis *	+	Yes, dropout imputation on basis of work status
o. Presentation point estimates and distribution measures	+	Yes, means and standard deviations are presented

Note. *internal validity criteria.

A) METHODOLOGICAL ISSUES

In chapter 2 we reviewed several RCT's on PST with regard to their methodological quality. What are the methodological characteristics of our study? Therefore we used the criteria list applied in the methodological quality assessment of our systematic review, which was based on Van Tulder et al. (2000). The methodological quality of the RCT turned out to be high, in that 17 of 18 criteria were scored positive (table 1). In the review in chapter 2, the methodological quality of the studies ranged from 6 to 16. The only criterion, on which information was not explicitly described in the current RCT, was criterion e on co-interventions. Although medical consumption was registered in the cost-diary, this concerned the year after the intervention, and not the period during the intervention. Nevertheless, most methodological quality criteria were addressed in the clinical trial and described in this thesis (chapter 6 and 7), what may further support the reliability of the findings.

One critical comment may be that high methodological quality still does not guarantee the quality of the intervention provided. As discussed in response to the systematic review, a good balance has to be found between the development of an effective treatment and strong methodology to validate conclusions. In the current RCT, some things can be said to assure the quality of the treatment. These are the extensive treatment protocol, the experience of the therapists regarding protocolized cognitive-behavioral therapies, number and duration of treatment sessions, and supervision of the therapists by the project manager. Nonetheless, the study lacks a treatment integrity check, which may be one of its most important shortcomings. There is no direct assessment that the treatment was performed according to the protocol. Nevertheless, there may be 2 reasons to assume that the PST-protocol was performed as intended. First, homework assignments were checked as a compliance measure. An important part of the therapy was spent on explaining new homework assignments and discussing the ones completed between treatment sessions. Homework assignments were linked to the themes discussed in the accompanying sessions, and could hardly be accomplished without the explicit explanation during the session. As we learned from the compliance assessment, homework assignments were kept up very reasonably by most of the patients, i.e. 50% of the patients in the GAPS condition completed more than 16 out of 21 homework assignments. Secondly, the manipulation check showed that problem-solving skills were explicitly changed in the experimental condition. Although this may not reflect what exactly happened during the therapy, and how treatment was performed, it increases the plausibility that the treatment actually trained problem-solving skills.

Another methodological issue that needs further discussion, is the choice of group education as the attention-control condition. The research question was very specific on this part, as we wanted to know whether the specific therapy method, i.e. training problem-solving skills, had a supplemental value. In that sense, we were especially interested in the theoretical meaning of problem

solving in low back pain patients, which can best be addressed by means of a dismantling design (Strayhorn, 1987). In such a design, the control condition had to be chosen in a way identical to the experimental condition regarding graded activity and non-specific features. Moreover, the attention-control treatment (here group education) should be as similar as possible to the active treatment, to produce equal credibility and to create similar expectations (Turk, Rudy, & Sorkin, 1993). Patients receiving the control treatment should have visit frequency, contact, and support equivalent to that in the active treatment. The choice of the attention-control condition is a dilemma for intervention studies, especially when psychotherapy is concerned (Parloff, 1986; Strayhorn, 1987). First, it is difficult to distinguish specific from non-specific effects. Second we can not be sure that both treatments create the same amount of non-specific effects. Third, an attention control condition may include some specific elements as well. And fourth, the therapist cannot be blinded like in drug efficacy research (Parloff, 1986). Parloff (1986) advises to choose a known alternate treatment or modality as the attention control, but preferably this treatment should differ markedly regarding underlying theory, assumptions, and procedures. The latter is what we tried to achieve when choosing group education as a control therapy, as it is an alternate therapy that has a very different rationale and complete different procedure (providing information versus skills training).

The dismantling design also has some disadvantages. Because both treatment conditions have to be comparable regarding other treatment components in the package (here graded activity) and non-specific elements, the contrast between conditions decreases. Consequently, the power becomes smaller and a greater sample size is needed to detect differences at a statistically significant level (Strayhorn, 1987). In the clinical trial discussed here, it is not unlikely that modest effects were due to reduced contrast between contents of both conditions. Nevertheless, outcomes were all in the hypothesized direction. Regarding measures that were hypothesized to be specifically addressed by graded activity (pain-related fear, pain catastrophizing, fear-avoidance, behavioral measures), we did not find supplemental effects of the problem solving therapy, as expected. However, with regard to problem solving skills and more general indices of outcome, such as functional status and sick leave, expected supplemental effects were found. With regard to state negative affect, as measured by the BDI, we did not find the hypothesized effect. One of the reasons may be that BDI-levels were low at pre-treatment and did not allow great changes anymore (floor effect). Indeed the mean scores on the BDI ($M = 7.2$; $S.D. = 5.4$) did not indicate that this group was clinically depressed (Hoevenaars & Van Son, 1984). Another remark in this context is that the systematic review concluded that success rates regarding measures referring to negative affect were low, especially when the population was not a clinically depressed one (chapter 2 of this thesis).

As it was important that the treatments were equally credible (Turk et al.,

1993), treatment credibility was assessed after the first session when the rationale had been explained to the patients. Remarkably, group education was rated as somewhat more credible than the problem-solving therapy condition. At least two explanations might be given for this discrepancy. First, group education was developed such that it would be very acceptable and informative to the patients. To that respect, we did more than our best to create a therapy that would be as appealing as the problem solving therapy. Secondly, group education has a longer history in the treatment of low back pain, and has generally become an accepted way of treatment in Dutch health care. Furthermore, patients in the experimental condition were frequently indicating that they expected that therapy would include education about the back, posture and physiology. The introduction of the problem-solving therapy might not have been what patients expected, and might even have been considered as a rather peculiar treatment. Finally, and not in line with Kolen-Snijders et al. (1999), we did not find treatment credibility to be predictive of outcomes. It was our experience, however, that it took most patients 3 to 4 sessions to get used to the new approach. Credibility might have developed during the first sessions, and as a consequence, the measurement during the first session was not related to what patients eventually thought of the therapy. Unfortunately, credibility was not reassessed halfway the therapy. Post-treatment, credibility was reassessed with regard to the helpfulness of therapy to resume activities and work. At this time, credibility was comparable between conditions, mainly because group education had lower credibility than in the beginning of the therapy.

Still another threat to internal validity might have been selection bias. For instance, referring physicians could have referred only the smart and highly educated patients, as they thought the PST demanded a certain level of intelligence. This could have resulted in failure to find interaction effects (education x condition). However, education level turned out to be normally distributed, and both low- and high-educated patients were presented in the final sample. Selection bias regarding education therefore did not seem to be an issue. On the other hand, referrals could have been biased because only the most alarming cases were referred to the project. Although we had the impression that this was the case in the first period of the project, the final sample was normally distributed regarding duration of pain, duration of sick leave, and disability. The latter may be an indication that referrals were not biased to that respect.

Finally, we would like to discuss the intention-to-treat analyses and the choices we made regarding the imputation of missing data. We considered 2 contrasting hypotheses: 1) patients dropped out because they were too disabled to come to the measurement, and 2) patients dropped out because they had a full return to work, and were too occupied to participate in the research any longer, or did not want to be remembered of their episode of pain anymore. Because we were able to link data from the OHS to most of the cases

that dropped out, we could identify those patients who made a 100% return-to-work, and thus, for whom the second hypothesis was most plausible. In all other cases where the OHS reported less than 100% return-to-work, the first hypothesis was followed, and worst outcomes were imputed. The general rule for imputation of missing data was therefore linked to the work status of the patient. We could have chosen to apply a 'worst case scenario' (hypothesis 1) to all cases that dropped out, which is a more common way of dealing with dropouts. However, by doing so, we would probably not have done justice to the real situation in the group of patients reported 100% return-to-work at the time of measurement. Moreover, post hoc analyses revealed that the groups, "100% return-to-work" versus "less than 100% return-to-work", were significantly different with regard to disability, negative affect, health care utilization and days of sick leave at 6- and 12-month follow-up. The group with 100% return-to-work, as expected, reported lowest levels. The latter may be an indication that work status was a valid measure to judge which of the two dropout hypotheses was most probable.

B) EXTERNAL VALIDITY AND GENERALIZATION TO CLINICAL PRACTICE

To what extent the results of the trial may be generalized to the treatment of back pain patients in general? Two aspects will be considered which might have influenced the external validity of the study. 1) Whether the study sample represented the target population for which the intervention was initially designed? 2) Whether the experimental intervention can be generalized to other treatment settings? Not only the practical feasibility of the project, but also conclusions regarding the value of problem solving therapy must be considered.

1) Generalisation to study population

At the start of the research in 1995, we had a lot of difficulties defining in- and exclusion criteria to select our target population. Part of the problem was due to the fact that definitions of acute, sub-acute and chronic pain were based on either duration of pain or duration of sick leave (Faas et al., 1996; Spitzer, LeBlanc, & Dupuis, 1987). These definitions did not coincide, and as a result there was a lot of misunderstanding regarding the description of "sub-acute", the target population of our intervention. Eventually, the inclusion criteria, which were communicated to the referring physicians, involved both pain and sick leave duration. Pain duration had to be longer than 6 weeks, so that the first phase, in which natural recovery has the most powerful effect, was passed. At the same time, sick leave duration might not have exceeded 20 weeks, so that the employee still would have a reasonable chance to return to his or her former job, which was one of the aims of the secondary prevention program. Taking into consideration the Dutch situation at that moment, the criteria for inclusion were chosen such that the study sample would be a good reflection of the population at stake. Our initial plan to select only those

employees, who were on the sick list for 6 to 12 weeks, soon turned out to be incompatible with the situation in the Netherlands. Occupational health services (OHS), which had just been privatized, were helped by an early return-to-work (< 6 weeks) to satisfy their clients (the firms where our potential subjects were working at). Moreover, we found out that the natural history of sick leave was as whimsically as that of pain. It turned out that there existed a substantial group of employees who had recurrent short periods of sick leave and would fit perfectly within the targets of our intervention, which were among others to reduce sick leave. Therefore, we decided that workers could be included from the first day of sick leave. Patients who were on the sick list for more than 20 weeks were not included in our study. The maximum of 20 weeks was chosen because, in the Netherlands, at 26 weeks the request to assign a disability pension will be discussed for the first time. With regard to the generalization of effects, we therefore must bear in mind that results refer to a group with sick leave duration smaller than 20 weeks, but pain duration longer than 6 weeks.

It was hypothesized that education level would interact with effects, in that high-educated patients would profit more from the problem solving condition. This hypothesis was not confirmed in our study. Can problem-solving therapy thus be implemented regardless of education level? We think it can, but under the conditions that therapy contents is kept very practical, therapists use a lot of examples, and patients are helped to understand how these techniques could fit into their own lives. This is what therapists in our setting tried to do and what does not naturally follow from the treatment protocol.

2) Generalization to other treatment settings

As discussed in chapters 6 and 7 of this thesis, the role of the multidisciplinary team, which was quite experienced in the behavioral rehabilitation of chronic pain patients, may not be underscored. Next, the importance of disciplines working together and conveying the same messages, may have been one of the strengths of the setting in which the current trial was executed. Finally, the possibilities of PST in primary care and occupational health care may be examined. Especially because these settings tend to be more accessible and referrals are more direct as compared to secondary health care settings, they lend themselves more for early preventive interventions.

C) LIMITATIONS OF THE TRIAL

Hereafter, the lack of a treatment-as-usual condition, and the validity of self-reported outcomes will be discussed.

1) The lack of a treatment-as-usual condition

By choosing a dismantling design, instead of a no-treatment control, we were not able to control for natural history and regression to the mean (Turner, Deyo, Loeser, Von Korff, & Fordyce, 1994). In other words, we do not know

how patients would have developed when given no extra treatment during the one-year follow-up. However, there are several reasons to assume that patients in a treatment-as-usual condition would have had less favorable outcomes than patients assigned to the GAPS or GAGE conditions. First, Graded activity was shown to be more effective than treatment-as-usual on several occasions (Lindström et al., 1992; Linton & Bradley, 1992; Linton, Bradley, Jensen, Spangfort, & et al., 1989). Because graded activity, as presented in the current study, is comparable to the protocols of Lindström et al. (1992) and Linton et al. (Linton & Bradley, 1992; Linton et al., 1989), we may cautiously suggest that graded activity would have been more effective than treatment-as-usual. Moreover, the course of sick leave of patients included in our study, is much more favorable as compared to the data of the Quebec taskforce regarding the natural history of sick leave as a result of back pain (Spitzer et al., 1987). The Quebec-curve (Spitzer et al., 1987) predicts that patients with a sick leave duration of about 6 weeks (median score in our sample), have a 50% chance to return to work within 6 months. In comparison, patients in our experimental group had a 76% chance to have a 100% return-to-work within 6 months, and 86% within 12 months. From this comparison we may conclude that we successfully deviated from the natural course of pain, by increasing chances to return-to-work.

2) The validity of self-reported outcomes

The use of self-report measures might have hampered the validity of the results. What is the evidence that patients really changed their problem-solving skills? And are patients really less fearful and avoidant with regard to physical activities and work, or were they only taught not to be fearful, but in fact still were? The observational measures might have offered a useful addition to the self-report measures regarding fear-avoidance. Remarkably, sick leave data from the cost-diary, which were self-reported, were not completely in line with those of the OHS. As a result, the cost diary did not show effects regarding days of sick leave, whereas OHS-registration did. This may give reason to doubt reliability of data sources. First, both sources had to deal with missing values. In case of the cost-diary, missing values were due to dropout of patients, which may not have been randomly. At least 18 of 26 week-diaries had to be returned to the research assistant. In those cases, less than 26 weeks were filled out (but more than 18 weeks), missing weeks were substituted by the average score of the available weeks. This might have induced bias, especially when data imputation was more favorable in one condition than the other. On the other hand, when data were missing from the OHS report, the case was removed from the data analyses. As far as we could observe, OHS-data were randomly missing, i.e. there were no reasons to belief that only the worst or the best cases dropped out. In conclusion, OHS-data appear to be more objective, data were randomly missing, and sensitivity analyses had comparable outcomes, and therefore OHS-data are likely to be more reliable than

data from the cost-diary. In future research, it might be advisable to include objective sources to the data collection, such that outcomes do not completely rely on self-reported outcomes.

FURTHER RESEARCH

Problem solving is an iterative loop: solving a problem almost automatically triggers new questions. This thesis is no exception to this rule. By answering the 5 main research questions, new 'scientific problems' to be addressed in future research were identified. Next, we will summarize the questions that followed from the studies described in this thesis.

1) Within the population of low back pain patients, there is a demonstrable group of patients who are characterized by high levels of pain-related fear. In this group avoidance of activities may be a mechanism responsible for high levels of disability. At the same time, and also recognized as such in clinical practice, Hasenbring et al. (1994) hypothesized that an endurance-style, with identifiable cognitions and strategies, may be another mechanism in chronic pain. First, valid measures are needed to recognize patients with high levels of endurance. Next, the role of endurance in the chronification of pain and disability may be studied in prospective designs as well as experimental studies. Finally, when the endurance-model is supported by empirical evidence, consequences for treatment may be considered and tested.

2) Was pain-related fear a predictor of treatment outcome in our study? I.e. did patients with high pain-related fear at baseline profit less from the interventions? This question may link to the more generally formulated question: What's best for whom? Research addressing this question, may help health care professionals to make better choices with respect to the treatment of pain patients. It may prevent patients shopping from one medical specialist to the other, slowly slipping into the vicious circle of chronic pain.

3) Problem solving skills were measured by means of a self-report measure. This may raise questions about the validity of the results (e.g. were problem-solving skills really changed?) It would be interesting to develop a structured observation-method, in which the individual is asked to solve standardized real-life problems. Such a method could serve as validation of the existing self-report measures.

4) Prospective research is needed to unravel the more complex relations between problem orientation, rational problem solving skills, pain, and disability.

5) The comparison with treatment-as-usual is desirable, the more because it would be most interesting to investigate whether the experimental intervention, as well as graded activity only, would do better than the natural course of pain, and thus prevent chronic pain disability.

6) What are the costs of such an intensive program that is guided by various highly skilled professionals? And do benefits surpass the costs? Interested parties, such as employers, occupational health services, and insurance companies, will find it most important to know whether the investment in a specific

program at an early stage of illness, pays itself back in the long term. Especially where savings of indirect costs (sick leave and disability pensions) are concerned, outcomes of cost-effectiveness may add important information for policymakers. It would therefore be interesting to execute a cost-effectiveness study regarding these data.

WHAT ABOUT PROBLEM-SOLVING?

What did the studies included in this thesis teach us about the merits of problem solving in how we should think about pain and pain treatment? In the introductory chapter, it was hypothesized that problem solving would influence a sense of control that people experience about daily stress and life in general. Moreover, a positive problem orientation would hypothetically prevent stress-related avoidance and hence promote functional restoration. This may be of particular importance in pain patients for the following reasons. First, active coping is required now the patient is confronted with several hindrances resulting from pain. Second, feeling confident in dealing with problems, and problems at the workplace in particular, may precipitate recovery and return-to-work. Nevertheless, in comparison to a current cognitive-behavioral model of pain, stressing the role of pain-related fear, problem solving turned out to play only a minor and indirect role in pain disability (Shaw et al., 2001; Van den Hout et al., 2001b). How can these findings be interpreted and how can problem solving and problem solving therapy be situated within current conceptualizations of pain?

In the current debate on the differentiation of constructs associated with negative affectivity, Lilienfeld (1996) proposed a hierarchical factor model (figure 1), in which lower order traits are nested hierarchically within a higher order dimension, e.g. anxiety sensitivity nests within trait anxiety. The higher order dimension, in turn, may covary sufficiently with other traits to form a still higher order dimension, e.g. negative affectivity (Watson & Clark, 1984). The lower in the hierarchy, the more specific the factor becomes. 'Trait anxiety' would refer to a tendency to react anxiously to potentially anxiety-inducing stimuli in general, whereas 'anxiety sensitivity' would refer to a more specific tendency to react anxious to one's own anxiety. Next in the hierarchy, at a lower order level, anxiety becomes even more specific, e.g. 'concerns regarding physical sensations'. Although variation in the lower order factors may be largely explained by variation in the higher order factor, the lower order factor may possess incremental validity in the prediction of more concrete and delineated outcomes, because it possesses unique variance that is especially relevant to these outcomes (Lilienfeld, 1996).

According to this model, 'concern about pain' (pain catastrophizing) and 'concern about movement (re) injury' would logically nest in the high-order factor 'concern regarding physical harm'. 'Concern regarding physical harm', in turn, may nest in 'injury sensitivity'. The model by Lilienfeld (1996) offers a

Figure 1. Proposed hierarchical factor model, based on Lilienfeld, Turner, and Jacob (1993, p.172).

very useful and clear structure to unravel psychological variables sharing variance in the prediction of more specific constructs within psychopathology. Moreover, it may help explain current findings regarding the role of problem-solving and pain-related fear in a model of pain. Not only psychological variables, but also therapies may be classified according to the level at which they are assumed to intervene.

At which level in the hierarchy would problem-solving be situated? Negative problem orientation may be identified as almost synonymous with negative affectivity, as the latter was defined as "the disposition to experience significant levels of distress and dissatisfaction at all times and in any given situation", and moreover, "to dwell on failures and shortcomings" (Watson and Pennebaker, 1989). The idea that problem solving may specifically relate to a higher-order factor, i.e. 'negative affectivity', is supported by the fact that problem-solving therapy was found to be effective in such a broad variety of disorders including depression, conduct disorders, and anxiety. On the other hand, kinesiophobia which is defined as "an irrational and debilitating fear of physical movement resulting from a feeling of vulnerability to painful injury or reinjury", refers to very specific concerns regarding a clearly identifiable stimulus and considered to be a lower-order factor. This explains why kinesiophobia and other kinds of pain-related fear possess a greater part of unique variance with very specific outcomes such as pain disability, as compared to higher-order factors like 'negative affectivity'.

Notwithstanding the foregoing considerations, problem-solving therapy turned out to have supplemental value in the treatment of low back pain problems. How could these effects be explained, or in other words, at which level of the Lilienfeld-model did problem-solving therapy intervene? We already suggested that problem solving would relate to 'negative affectivity' in terms of a negative orientation towards stressful events and problems in general. It is not completely hypothetical to suggest that problem-solving therapy intervened at this level, because we found that indices like negative and positive problem orientation, significantly changed in our clinical trial. There may be several explanations for these changes. First, the increase of rational problem-solving skills may have increased a sense of control ("I can do something about it"). Second, the catastrophes that were expected to happen, especially in patients who perceive stressful situations as threatening, did not occur when patients challenged the problems during therapy. In a certain way, the confrontation with problems that might have been avoided previously, and the challenge of irrational beliefs that go along with it may have worked as a behavioral experiment, the core ingredient of current cognitive therapy. Finally, changes at a higher level may be explained by changes in more specific problems at a lower level of the hierarchy, e.g. regarding work, family, or social situations. It is not clear, however, how these specific changes would have generalized to higher levels in the hierarchy.

Finally, and in line with the foregoing, we may formulate a proposal about

at which level in the hierarchy interventions can best be designed. When the patient shows elevated levels of pain-related fear, it may be wise to intervene at this specific level first, with the appropriate techniques that serve the aim best (e.g. exposure-in-vivo with behavioral experiments; Vlaeyen et al., 2001). Although changes at a higher level in the hierarchy may have very powerful and broad effects, according to the Lilienfeld-model, they may be very hard to achieve. At the highest level, negative affectivity, as a personality trait, is thought of as very stable and difficult to change. Since problem solving skills were applied in more specific and personally relevant problems, problem-solving therapy may probably have affected lower-order variables as well. Accordingly, the supplemental value of general interventions such as the problem solving therapy will for the greater part depend on the personal relevance to the patient. The better the problem-solving model is applied in specific problems at a lower level in the hierarchy, the more convincing the additional effect of the therapy will be.

IN CONCLUSION,

In spite of the modest relationship between problem solving skills and pain disability, an experimental therapy in which problem-solving skills were trained, turned out to increase the efficacy of the more physically oriented graded activity, and consequently, the implementation of problem-solving techniques in the treatment of low back pain, in combination with graded activity may be recommended from the findings in this thesis.

REFERENCES

- Aldrich, S., Eccleston, C., & Crombez, G. (2000). Worrying about chronic pain: Vigilance to threat and misdirected problem solving. *Behaviour Research and Therapy*, 38, 457-470.
- Andersson, G. B. (1998). Epidemiology of low back pain. *Acta Orthop Scand Suppl*, 28128-28131.
- Andersson, G. B. (1999). Epidemiological features of chronic low-back pain. *Lancet*, 354(9178), 581-585.
- Andersson, G. B. J. (1991). The epidemiology of spinal disorders. In J. W. Frymoyer (Ed.), *The adult spine: principles and practice* (pp. 107-146). New York: Raven Press.
- Arean, P. A., Perri, M. G., Nezu, A. M., Schein, R. L., Christopher, F., & Joseph, T.X. (1993). Comparative effectiveness of social problem-solving therapy and reminiscence therapy as treatments for depression in older adults. *Journal of consulting and clinical psychology*, 61(6), 1003-1010.
- Arntz, A., & de Jong, P. F. (1993). Anxiety, attention and pain. *Journal of Psychosomatic Research*, 37(4), 423-431.
- Arntz, A., Dreessen, L., & De Jong, P. (1994). The influence of anxiety on pain: Attentional and attributional mediators. *Pain*, 56(3), 307-314.
- Arntz, A., Dreessen, L., & Merckelbach, H. (1991). Attention, not anxiety, influences pain. *Behaviour Research and Therapy*, 29(1), 41-50.
- Arntz, A., & Schmidt, A. J. M. (1989). Perceived control and the experience of pain.
- Arntz, A., Van den Hout, M. A., Van den Berg, G., & Meijboom, A. (1991). The effects of incorrect pain expectations on acquired fear and pain responses. *Behaviour Research and Therapy*, 29(6), 547-560.
- Arrindell, W. A., & Ettema, J. H. M. (1986). *SCL-90. Handleiding bij een multidimensionele psychopathologie indicator (SCL-90; Manual of a multidimensional psychopathology-indicator)*. Lisse, the Netherlands: Swets & Zeitlinger.
- Asmundson, G. J., Kuperos, J. L., & Norton, G. R. (1997). Do patients with chronic pain selectively attend to pain-related information?: preliminary evidence for the mediating role of fear. *Pain*, 72(1-2), 27-32.
- Asmundson, G. J., Norton, P. J., & Norton, G. R. (1999). Beyond pain: the role of fear and avoidance in chronicity. *Clin psychol Rev*, 19(1), 97-119.
- Asmundson, G. J., & Taylor, S. (1996). Role of anxiety sensitivity in pain related fear and avoidance. *Journal of behavioral medicine*, 19(6), 577-586.
- Atha, C., Salkovskis, P. M., & Storer, D. (1992). Cognitive-behavioural problem solving in the treatment of patients attending a medical emergency department: a controlled trial. *Journal of psychosomatic research*, 36(4), 299-307.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173-1182.
- Baumgardner, A. H., Heppner, P. P., & Arkin, R. M. (1986). Role of causal attribution in personal problem solving. *Journal of Personality and Social Psychology*, 50(3), 636-643.

- Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, 8(1), 77-100.
- Beck, A. T., Ward, C. H., & Mendelson, M. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 561-571.
- Bergner, M., Bobbitt, R. A., Carter, W. B., & Gilson, B. S. (1981). The Sickness Impact Profile: development and final revision of a health status measure. *Medical Care*, 19(8), 787-805.
- Berwick, D. M., Budman, S., & Feldstein, M. (1989). No clinical effect of back schools in an HMO. A randomized prospective trial. *Spine*, 14(3), 338-344.
- Beurskens, A. J., de Vet, H. C., & Koke, A. J. (1996). Responsiveness of functional status in low back pain: a comparison of different instruments. *Pain*, 65(1), 71-76.
- Bigos, S. J., Battie, M. C., Spengler, D. M., Fisher, L. D., Fordyce, W. E., Hansson, T. H., Nachemson, A. L., & Wortley, M. D. (1991). A prospective study of work perceptions and psychosocial factors affecting the report of back injury. *Spine*, 16(1), 1-6.
- Blanchard, C. G., Toseland, R. W., & McCallion, P. (1996). The effects of a problem-solving intervention with spouses of cancer patients. *Journal of Psychosocial Oncology*, 14(2), 1-21.
- Boden, S. D., Davis, D. O., Dina, T. S., Patronas, N. J., & Wiesel, S. W. (1990). Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. prospective investigation. *Journal of bone and joint surgery, The*, 72(3), 403-408.
- Borenstein, D. (1998). Epidemiology, etiology, diagnostic evaluation, and treatment of low back pain. *Current opinion in rheumatology*, 10(2), 104-109.
- Borkovec, T. D., & Nau, S. D. (1972). Credibility of analogue therapy rationales. *Journal of Behavior Therapy and Experimental Psychiatry*, 3 (4), 257-260.
- Bortz, W. M. d. (1984). The disuse syndrome. *Western Journal of Medicine*, 141(5), 691-694.
- Brown, G. K. (1990). A causal analysis of chronic pain and depression. *J Abnorm Psychol*, 99, 127-137.
- Burton, A. K., & Tillotson, K. M. (1991). Prediction of the clinical course of low-back trouble using multivariable models. *Spine*, 16(1), 7-14.
- Burton, A. K., Tillotson, K. M., Main, C. J., & Hollis, S. (1995). Psychosocial predictors of outcome in acute and subchronic low back trouble. *Spine*, 20(6), 722-728.
- Caceres, C., & Burns, J. W. (1997). Cardiovascular reactivity to psychological stress may enhance subsequent pain sensitivity. *Pain*, 69(3), 237-244.
- Catalan, J., Gath, D. H., Anastasiades, P., Bond, S. A., Day, A., & Hall, L. (1991). Evaluation of a brief psychological treatment for emotional disorders in primary care. *Psychological medicine*, 21(4), 1013-1018.
- Cats-Baril, W. L., & Frymoyer, J. W. (1991). The economics of spinal disorders. In J. W. Frymoyer (Ed.), *The adult spine: principles and practice*. New York: Raven Press.
- Chang, E. C., & D' Zurilla, T. J. (1996). Relations between problem orientation and optimism, pessimism, and trait affectivity: A construct validation study. *Behaviour Research and Therapy*, 34(2), 185-194.

- Chaves, J. F., & Barber, T. X. (1974). Cognitive strategies, experimenter modeling, and expectation in the attenuation of pain. *Journal of abnormal psychology*, 83(4), 356-363.
- Cherkin, D. C., Deyo, R. A., Street, J. H., Hunt, M., & Barlow, W. (1996). Pitfalls of patient education. Limited success of a program for back pain in primary care. *Spine*, 21(3), 345-355.
- Ciotă, M. A., Zeichner, A., & Allen, J. D. (1998). *Effects of self-esteem and performance feedback on subjective pain and arousal*. Paper presented at the The fifth International Congress of Behavioral Medicine, Copenhagen, Denmark.
- Cipher, D. J., & Fernandez, E. (1997). Expectancy variables predicting tolerance and avoidance of pain in chronic pain patients. *Behaviour research and therapy*, 35(5), 437-444.
- Clarke, K. M., & Greenberg, L. S. (1986). Differential effects of the Gestalt two-chair intervention and problem solving in resolving decisional conflict. *Journal of Counseling Psychology*, 33(1), 11-15.
- Coche, E., & Douglas, A. A. (1977). Therapeutic effects of problem-solving training and play-reading groups. *Journal of Clinical Psychology*, 33(3), 820-827.
- Coche, E., & Flick, A. (1975). Problem-solving training groups for hospitalized psychiatric patients. *Journal of Psychology*, 91(1), 19-29.
- Cohen, J. E., Goel, V., Frank, J. W., Bombardier, C., Peloso, P., & Guillemin, F. (1994). Group education interventions for people with low back pain. An overview of the literature. *Spine*, 19(11), 1214-1222.
- Coleman, M., Wheeler, L., & Webber, J. (1993). Research on interpersonal problem-solving training: A review. *RASE: Remedial and Special Education*, 14(2), 25-37.
- Council, J. R., Ahern, D. K., Follick, M. J., & Kline, C. L. (1988). Expectancies and functional impairment in chronic low back pain. *Pain*, 33(3), 323-331.
- Craig, K. D. (1994). Emotional aspects of pain. In P. D. Wall & R. Melzack (Eds.), *Textbook of pain*. London: Churchill Livingstone.
- Crombez, G., Vervaeke, L., Lysens, R., Baeyens, F., & Eelen, P. (1998). Avoidance and confrontation of painful, back-straining movements in chronic back pain patients. *Behavior Modification*, 22(1), 62-77.
- Crombez, G., Vlaeyen, J. W., Heuts, P. H., & Lysens, R. (1999). Pain-related fear is more disabling than pain itself: evidence on the role of pain-related fear in chronic back pain disability. *Pain*, 80(1-2), 329-339.
- Crombez, G., & Vlaeyen, J. W. S. (1996). *The Pain Catastrophizing Scale*. Unpublished authorized Dutch/Flemish translation.
- Derogatis, L. R., Lipman, R. S., & Covi, L. (1973). SCL-90: An outpatient rating scale: preliminary report. *Psychopharmacological Bulletin*, 9, 13-27.
- DeVellis, B. M., Blalock, S. J., Hahn, P. M., DeVellis, R. F., & et al. (1988). Evaluation of a problem-solving intervention for patients with arthritis. *Patient Education and Counseling*, 11(1), 29-42.
- Deville, G. J., & Borkovec, T. D. (2000). Psychometric properties of the credibility/expectancy questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry*, 31(2), 73-86.

- Deyo, R. A., Rainville, J., & Kent, D. L. (1992). What can the history and physical examination tell us about low back pain? *Jama*, 268(6), 760-765.
- Di Fabio, R. P. (1995). Efficacy of comprehensive rehabilitation programs and back school for patients with low back pain: a meta-analysis. *Physical therapy*, 75(10), 865-878.
- Dickersin, K., Scherer, R., & Lefebvre, C. (1994). Identifying relevant studies for systematic reviews. *Bmj*, 309(6964), 1286-1291.
- DiGiuseppe, R., McGowan, L., Simon, K. S., & Gardner, F. (1990). A comparative outcome study of four cognitive therapies in the treatment of social anxiety. *Journal of Rational Emotive and Cognitive Behavior Therapy*, 8(3), 129-146.
- Dijkstra, A., Vlaeyen, J. W. S., Rijnen, H., & Nielson, W. (2001). Readiness to adopt the self-management approach to cope with chronic pain in fibromyalgic patients. *Pain*, 90, 37-45.
- Drenth, P. J. D. (1965). *Test voor niet verbale abstractie (Test for non-verbal abstraction)*. Amsterdam: Swets & Zeitlinger.
- Drenth, P. J. D., & Hoolwerf, G. (1970). *Cijferreeksen: Numerieke Aanleg Test (NAT'70) (Numerical Series: Numerical Ability Test)*. Amsterdam: Swets & Zeitlinger.
- Drenth, P. J. D., & Van Wieringen, P. C. W. (1969). *Verbale analogieën: Verbale Aanleg Testserie (VAT'69) (Verbal Analogies: Verbal Ability Testseries)*. Amsterdam: Swets & Zeitlinger.
- Dugas, M. J., Letarte, H., Rheume, J., Freeston, M. H., & et al. (1995). Worry and problem solving: Evidence of a specific relationship. *Cognitive Therapy and Research*, 19(1), 109-120.
- D'Zurilla. (1988). Problem-solving therapies. In K. S. Dobson (Ed.), *Handbook of cognitive behavioral therapies* (pp. 85-135). New York, NY, US: The Guilford Press.
- D'Zurilla, T. J., & Goldfried, M. R. (1971). Problem solving and behavior modification. *Journal of abnormal psychology*, 78(1), 107-126.
- D'Zurilla, T. J., & Nezu, A. M. (1982). Social problem-solving in adults. In P. C. Kendall (Ed.), *Advances in cognitive-behavioral research and therapy* (Vol. 1, pp. 202-274). New York: Academic Press.
- D'Zurilla, T. J., Nezu, A. M., & Maydeu-Olivares, A. (1994). *Manual for the Social Problem Solving Inventory-Revised*. Unpublished manuscript, State University of New York at Stony Brook.
- D'Zurilla, T. J., Nezu, A. M., & Maydeu-Olivares, A. (1997). *Manual for the Social Problem-Solving Inventory - Revised*. North Tonawanda, NY: Multi-Health Systems.
- D'Zurilla, T. J., & Sheedy, C. F. (1991). Relation between social problemsolving ability and subsequent level of psychological stress in college students. *Journal of Personality and Social Psychology*, 61(5), 841-846.
- Eccleston, C., & Crombez, G. (1999). Pain demands attention: a cognitive affective model of the interruptive function of pain. *Psychological bulletin*, 125(3), 356-366.
- Faas, A., Chavannes, A. W., Koes, B. W., Van den Hoogen, J. M. M., Mens, J. M. A., Smeele, L. J. M., Romeijnders, A. C. M., & Van der Laan, J. R. (1996). NHG-standaard lage-rugpijn (Dutch Guidelines low back pain). *Huisarts & Wetenschap*, 39, 18-31.

- Feuerstein, M., Menz, L., Zastowny, T., & Barron, B. A. (1994). Chronic back pain and work disability: Vocational outcomes following multidisciplinary rehabilitation. *Journal of Occupational Rehabilitation, 4*, 229-251.
- Flor, H., & Birbaumer, N. (1993). Comparison of the efficacy of electromyographic biofeedback, cognitive-behavioral therapy, and conservative medical interventions in the treatment of chronic musculoskeletal pain. *Journal of consulting and clinical psychology, 61*(4), 653-658.
- Flor, H., Fydrich, T., & Turk, D. C. (1992). Efficacy of multidisciplinary pain treatment centers: a meta-analytic review. *Pain, 49*(2), 221-230.
- Flor, H., Kerns, R. D., & Turk, D. C. (1987). The role of spouse reinforcement, perceived pain, and activity levels of chronic pain patients. *Journal of psychosomatic research, 31*(2), 251-259.
- Flor, H., Turk, D. C., & Birbaumer, N. (1985). Assessment of stress-related psychophysiological reactions in chronic back pain patients. *Journal of Consulting and Clinical Psychology, 53*(3), 354-364.
- Fordyce, W. E. (1976). *Behavioral methods for chronic pain and illness*. St. Louis: Mosby.
- Fordyce, W. E. (1989). Learning factors in pain. *Scandinavian journal of rheumatology. Supplement, 82*, 13-17.
- Fordyce, W. E., Brockway, J. A., Bergman, J. A., & Spengler, D. (1986). Acute back pain: a control group comparison of behavioral vs traditional management methods. *J Behav Med, 9*, 127-140.
- Foster, S. L., Prinz, R. J., & O'Leary, K. D. (1983). Impact of problem-solving communication training and generalization procedures on family conflict. *Child and Family Behavior Therapy, 5*(1), 1-23.
- France, C. R., & Stewart, K. M. (1995). Parental history of hypertension and enhanced cardiovascular reactivity are associated with decreased pain ratings. *Psychophysiology, 32*(6), 571-578.
- Frymoyer, J. W. (1988). Back pain and sciatica. *The New England journal of medicine, 318*(5), 291-300.
- Gallagher, R. M., Rauh, V., Haugh, L. D., Milhous, R., Callas, P. W., Langelier, R., McClallen, J. M., & Frymoyer, J. (1989). Determinants of return-to-work among low back pain patients. *Pain, 39*(1), 55-67.
- Gamsa, A. (1990). Is emotional disturbance a precipitator or a consequence of chronic pain? *Pain, 42*(2), 183-195.
- Gamsa, A. (1994). The role of psychological factors in chronic pain. I. A half century of study. *Pain, 57*(1), 5-15.
- Gatchel, R. J., & Turk, D. C. (1996). *Psychological approaches to pain management. A practitioner's handbook*. New York, London: Guilford Press.
- Gath, D., & Catalan, J. (1986). The treatment of emotional disorders in general practice: Psychological methods versus medication. *Journal of Psychosomatic Research, 30*(3), 381-386.

- Gheldof, E. L. M., Van den Heuvel, W. P. M., Vinck, J., Hidding, A., Vlaeyen, J. W. S., Crombez, G., & Mairiaux, P. (2001). *The relationship between fear-avoidance and occupational disability due to back pain*. Paper presented at the Fourth International Scientific Conference on Work-Related Musculoskeletal Disorders, Amsterdam.
- Godshall, F. J., & Elliott, T. R. (1997). Behavioral correlates of self-appraised problem-solving ability: Problem-solving skills and health-compromising behaviors. *Journal of Applied Social Psychology*, 27(11), 929-944.
- Gommans, I., Koes, B. W., & Van Tulder, M. W. (1996). Validity and responsiveness of the Dutch adaption of the Roland Disability Questionnaire. In M. W. Van Tulder & B. W. Koes & L. M. Bouter (Eds.), *Low back pain in primary care: Effectiveness of diagnostic and therapeutic interventions* (Vol. Chapter 6). Amsterdam: EMGO institute.
- Goossens, M. E., Rutten Van Molken, M. P., Kole Snijders, A. M., Vlaeyen, J. W., Van Breukelen, G., & Leidl, R. (1998). Health economic assessment of behavioural rehabilitation in chronic low back pain: a randomised clinical trial. *Health economics*, 7(1), 39-51.
- Goossens, M. E., Rutten van Molken, M. P., Vlaeyen, J. W., & van der Linden, S. M. (2000). The cost diary: a method to measure direct and indirect costs in cost-effectiveness research. *Journal of clinical epidemiology*, 53(7), 688-695.
- Goubert, L., Crombez, G., Vlaeyen, J., Van Damme, S., Van den Broeck, A., & Van Houdenhove, B. (2000). The Tampa Scale for Kinesiophobia: Psychometric Characteristics and norms. *Gedrag en Gezondheid*, 28(2), 54-62.
- Graves, T., Meyers, A. W., & Clark, L. (1988). An evaluation of parental problem-solving training in the behavioral treatment of childhood obesity. *Journal of Consulting and Clinical Psychology*, 56(2), 246-250.
- Green, T. D., Bailey, R. C., Zinser, O., & Williams, D. E. (1994). Causal attribution and affective response as mediated by task performance and self-acceptance. *Psychological reports*, 75(3 Pt 2), 1555-1562.
- Grossman, P. B., & Hughes, J. N. (1992). Self-control interventions with internalizing disorders: A review and analysis. *School Psychology Review*, 21(2), 229-245.
- Halford, W. K., Goodall, T. A., & Nicholson, J. M. (1997). Diet and diabetes (II): A controlled trial of problem solving to improve dietary self-management in patients with insulin dependent diabetes. *Psychology and Health*, 12(2), 231-238.
- Hasenbring, M., Marienfeld, G., Kuhlendahl, D., & Soyka, D. (1994). Risk factors of chronicity in lumbar disc patients. A prospective investigation of biologic, psychologic, and social predictors of therapy outcome. *Spine*, 19(24), 2759-2765.
- Hepner, P. P., & Krauskopf, C. J. (1987). An information-processing approach to personal problem solving. *Counseling Psychologist*, 15(3), 371-447.
- Hepner, P. P., Neal, G. W., & Larson, L. M. (1984). Problem-solving training as prevention with college students. *Personnel and Guidance Journal*, 62(9), 514-519.
- Hepner, P. P., Reeder, B. L., & Larson, L. M. (1983). Cognitive variables associated with personal problem-solving appraisal: Implications for counseling. *Journal of Counseling Psychology*, 30(4), 537-545.

- Heyneman, N. E., Fremouw, W. J., Gano, D., Kirkland, F., & et al. (1990). Individual differences and the effectiveness of different coping strategies for pain. *Cognitive Therapy and Research*, 14(1), 63-77.
- Hoevenaars, J. P., & Van Son, M. J. (1984). Drie Nederlandstalige zelfrapportagelijsten in de diagnostiek van depressie. / Three Dutch self-report inventories in the diagnosis of depression. *Gedragstherapie*, 17(3), 225-239.
- Horan, J. J., & Dellinger, J. K. (1974). "In vivo" emotive imagery: a preliminary test. *Perceptual and motor skills*, 39(1 Pt 2), 359-362.
- Hurley, D. A., Duso, T. E., McDonough, S. M., Moore, A. P., & Baxter, G. D. (2001). How effective is the acute low back pain screening questionnaire for predicting 1-year follow-up in patients with low back pain? *Clinical Journal of Pain*, 17(3), 256-263.
- Hussian, R. A., & Lawrence, P. S. (1981). Social reinforcement of activity and problem-solving training in the treatment of depressed institutionalized elderly patients. *Cognitive Therapy and Research*, 5(1), 57-69.
- Jadad, A. R. (1998). *Randomised controlled trials: A user's guide*. London: BMJ Books.
- Jannoun, L., Munby, M., Catalan, J., & Gelder, M. (1980). A home-based treatment program for agoraphobia: Replication and controlled evaluation. *Behavior Therapy*, 11(3), 294-305.
- Janssen, S. A., Spinhoven, P., & Brosschot, J. F. (2001). Experimentally induced anger, cardiovascular reactivity, and pain sensitivity. *Journal of psychosomatic research*, 51(3), 479-485.
- Jansson, L., & Oest, L. G. (1982). Behavioral treatments for agoraphobia: An evaluative review. *Clinical Psychology Review*, 2(3), 311-336.
- Jensen, M. C., Brant Zawadzki, M. N., Obuchowski, N., Modic, M. T., Malkasian, D., & Ross, J. S. (1994). Magnetic resonance imaging of the lumbar spine in people without back pain. *New England journal of medicine*, 331(2), 69-73.
- Jensen, M. P., Turner, J. A., & Romano, J. M. (1991a). Self-efficacy and outcome expectancies: relationship to chronic pain coping strategies and adjustment. *Pain*, 44(3), 263-269.
- Jensen, M. P., Turner, J. A., Romano, J. M., & Karoly, P. (1991b). Coping with chronic pain: A critical review of the literature. *Pain*, 47(3), 249-283.
- Johnson, S. M., & Greenberg, L. S. (1985). Differential effects of experiential and problem-solving interventions in resolving marital conflict. *Journal of Consulting and Clinical Psychology*, 53(2), 175-184.
- Kanfer, F. H., & Busemeyer, J. R. (1982). The use of problem solving and decision making in behavior therapy. *Clinical Psychology Review*, 2(2), 239-266.
- Kanfer, F. H., & Philips, J. S. (1970). *Learning foundations of behavior therapy*. New York: John Wiley & Sons, Inc.
- Kant, G. L., D'Zurilla, T. J., & Maydeu Olivares, A. (1997). Social problem solving as a mediator of stress-related depression and anxiety in middle-aged and elderly community residents. *Cognitive Therapy and Research*, 21(1), 73-96.
- Kazdin, A. E. (1997). Practitioner review: Psychosocial treatments for conduct disorder in children. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 38(2), 161-178.

- Kazdin, A. E., Bass, D., Siegel, T., & Thomas, C. (1989). Cognitive-behavioral therapy and relationship therapy in the treatment of children referred for antisocial behavior. *Journal of consulting and clinical psychology*, 57(4), 522-535.
- Kazdin, A. E., Esveldt Dawson, K., French, N. H., & Unis, A. S. (1987). Problem-solving skills training and relationship therapy in the treatment of antisocial child behavior. *Journal of Consulting and Clinical Psychology*, 55(1), 76-85.
- Kazdin, A. E., Siegel, T. C., & Bass, D. (1992). Cognitive problem-solving skills training and parent management training in the treatment of antisocial behavior in children. *Journal of consulting and clinical psychology*, 60(5), 733-747.
- Keefe, F. J., Lefebvre, J. C., Maixner, W., Salley, A. N., & Caldwell, D. S. (1997). Self-efficacy for arthritis pain: relationship to perception of thermal laboratory pain stimuli. *Arthritis care and research*, 10(3), 177-184.
- Kendall, N. A. S., Linton, S. J., & Main, C. J. (1997). *Guide to assessing psychosocial Yellow Flags in acute low back pain: Risk factors for long-term disability and work loss*. New Zealand: Accident Rehabilitation and Compensation Insurance Corporation of New Zealand and the National Health Committee.
- Klein, D. C., Fencil Morse, E., & Seligman, M. E. (1976). Learned helplessness, depression, and the attribution of failure. *Journal of personality and social psychology*, 33(5), 508-516.
- Klennerman, L., Slade, P. D., Stanley, I. M., Pennie, B., Reilly, J. P., Atchison, L. E., Troup, J. D., & Rose, M. J. (1995). The prediction of chronicity in patients with an acute attack of low back pain in a general practice setting. *Spine*, 20(4), 478-484.
- Knepper, S., & Feenstra, H. (1991). Ziekte of gedrag: basisbegrippen bij de behandeling van arbeidsongeschikten. *Nederlands tijdschrift voor geneeskunde*, 135(37), 1672-1676.
- Koes, B. W., Assendelft, W. J., van der Heijden, G. J., Bouter, L. M., & Knipschild, P. G. (1991a). Spinal manipulation and mobilisation for back and neck pain: a blinded review. *Bmj*, 303(6813), 1298-1303.
- Koes, B. W., Bouter, L. M., Beckerman, H., van der Heijden, G. J., & Knipschild, P. G. (1991b). Physiotherapy exercises and back pain: a blinded review. *Bmj*, 302(6792), 1572-1576.
- Kole Snijders, A. M., Vlaeyen, J. W., Goossens, M. E., Rutten van Molken, M. P., Heuts, P. H., van Breukelen, G., & van Eek, H. (1999). Chronic low-back pain: what does cognitive coping skills training add to operant behavioral treatment? Results of a randomized clinical trial. *Journal of consulting and clinical psychology*, 67(6), 931-944.
- Kole-Snijders, A. M. J., Vlaeyen, J. W. S., van Eek, H., Schuerman, J. A., & Groenman, N.H. (1990). Behavioral assessment of residual health behavior in chronic low back pain patients through the pain-rest contingency principle. *Pain*, 5, S394.
- Kopec, J. A., Esdaile, J. M., Abrahamowicz, M., Abenhaim, L., Wood Dauphinee, S., Lamping, D. L., & Williams, J. I. (1995). The Quebec Back Pain Disability Scale. Measurement properties. *Spine*, 20(3), 341-352.
- Kori, S. H., Miller, R. P., & Todd, D. D. (1990). Kinesiophobia: a new view of chronic pain behavior. *Pain Management* (jan/feb), 35-43.

- Kormann, A. K. (1977). *Organizational Behavior*. Englewood Cliffs, NJ: Prentice-Hall.
- Lamberts, H. (1991). *In het huis van de huisarts. Verslag van het Transitie Projekt (In the General Practitioner's house. Report of the Transition Project)*. Lelystad: Meditekst.
- Larson, L. M. (1990). A critique of problem-solving training: Where to from here? *Journal of Cognitive Psychotherapy*, 4(3), 257-265.
- Leino, P., & Magni, G. (1993). Depressive and distress symptoms as predictors of low back pain, neck-shoulder pain, and other musculoskeletal morbidity: a 10-year follow-up of metal industry employees. *Pain*, 53(1), 89-94.
- Lethem, J., Slade, P. D., Troup, J. D., & Bentley, G. (1983). Outline of a Fear-Avoidance Model of exaggerated pain perception-I. *Behaviour Research and Therapy*, 21(4), 401-408.
- Levine, F. M., Krass, S. M., & Padawer, W. J. (1993). Failure hurts: the effects of stress due to difficult tasks and failure feedback on pain report. *Pain*, 54(3), 335-340.
- Lilienfeld, S. O. (1996). Anxiety sensitivity is not distinct from trait anxiety. In R. M. Rapee (Ed.), *Current controversies in the anxiety disorders* (pp. 228-244). New York: The Guilford Press.
- Lilienfeld, S. O., Turner, S. M., & Jacob, R. G. (1993). Anxiety sensitivity: an examination of theoretical and methodological issues. *Advances in behaviour Research and Therapy*, 15, 147-183.
- Lindström, I., Öhlund, C., Eek, C., Wallin, L., Peterson, L. E., Fordyce, W. E., & Nachemson, A. L. (1992). The effect of graded activity on patients with subacute low back pain: a randomized prospective clinical study with an operant-conditioning behavioral approach. *Phys Ther*, 72(4), 279-290.
- Linton, S. J. (2000). A review of psychological risk factors in back and neck pain. *Spine*, 25(9), 1148-1156.
- Linton, S. J., Althoff, B., Melin, L., Lundin, A., Bodin, L., Mägi, A., Lindström, K., & Lihagen, T. (1994). Psychological factors related to health, back pain, and dysfunction. *Journal of Occupational Rehabilitation*, 4(1), 1-10.
- Linton, S. J., & Bradley, L. A. (1992). An 18-month follow-up of a secondary prevention program for back pain: help and hindrance factors related to outcome maintenance. *Clin J Pain*, 8(3), 227-236.
- Linton, S. J., & Bradley, L. A. (1996). Strategies for the prevention of chronic pain. In R. J. Gatchel & D. C. Turk (Eds.), *Psychological approaches to pain management: A practitioner's handbook* (pp. 438-457). New York, London: Guildford Press.
- Linton, S. J., Bradley, L. A., Jensen, I., Spangfort, E., & et al. (1989). The secondary prevention of low back pain: A controlled study with follow-up. *Pain*, 36(2), 197-207.
- Linton, S. J., & Buer, N. (1995). Working despite pain: factors associated with work attendance versus dysfunction. *Int J Behav Med*, 2, 252-262.
- Linton, S. J., Buer, N., Vlaeyen, J. W. S., & Hellsing, A.-L. (2000). Are fear-avoidance beliefs related to the inception of an episode of back pain? a prospective study. *Psychology and Health*, 14, 1051-1059.

- Linton, S. J., & Gotestam, K. G. (1984). A controlled study of the effects of applied relaxation and applied relaxation plus operant procedures in the regulation of chronic pain. *Br J Clin Psychol*, 23(Pt 4), 291-299.
- Linton, S. J., & Hallden, K. (1998). Can we screen for problematic back pain? A screening questionnaire for predicting outcome in acute and subacute back pain. *Clinical journal of pain, The*, 14(3), 209-215.
- Linton, S. J., & Ryberg, M. (2001a). A cognitive-behavioral group intervention as prevention for persistent neck and back pain in a non patient population: a randomized controlled trial. *Pain*, 90, 83-90.
- Linton, S. J., & van Tulder, M. W. (2001b). Preventive interventions for back and neck pain problems: what is the evidence? *Spine*, 26(7), 778-787.
- Loisel, P., Abenhaim, L., Durand, P., Esdaile, J. M., Suissa, S., Gosselin, L., Simard, R., Turcotte, J., & Lemaire, J. (1997). A population-based, randomized clinical trial on back pain management. *Spine*, 22(24), 2911-2918.
- Loscalzo, M. J., & Bucher, J. A. (1999). The COPE model: Its clinical usefulness in solving pain-related problems. *Journal of Psychosocial Oncology*, 16(3-4), 93-117.
- Lousberg, R., Schmidt, A. J., & Groenman, N. H. (1992). The relationship between spouse solicitousness and pain behavior: searching for more experimental evidence. *Pain*, 51(1), 75-79.
- Luteijn, F., Starren, J., & Van Dijk, H. (1985). *Handleiding bij de Nederlandse Persoonlijkheids Vragenlijst (Manual of the Dutch Personality Questionnaire, revised)*. Lisse, the Netherlands: Swets & Zeitlinger.
- Lynch, D. J., Tamburrino, M. B., & Nagel, R. (1997). Telephone counseling for patients with minor depression: preliminary findings in a family practice setting. *The Journal of family practice*, 44(3), 293-298.
- Main, C. J., Wood, P. L., Hollis, S., Spanswick, C. C., & Waddell, G. (1992). The Distress and Risk Assessment Method. A simple patient classification to identify distress and evaluate the risk of poor outcome. *Spine*, 17(1), 42-52.
- Marhold, C., Linton, S. J., & Melin, L. (2001). A cognitive-behavioral return-to-work program: effects on pain patients with a history of long-term versus short-term sick leave. *Pain*, 91(1-2), 155-163.
- Mayerson, N. H., & Rhodewalt, F. (1988). Role of self-protective attributions in the experience of pain. *Journal of Social and Clinical Psychology*, 6(2), 203-218.
- McCracken, L. M., & Gross, R. T. (1993). Does anxiety affect coping with chronic pain? *Clinical journal of pain, The*, 9(4), 253-259.
- McCracken, L. M., & Gross, R. T. (1998). The role of pain-related anxiety reduction in the outcome of multidisciplinary treatment for chronic low back pain: Preliminary results. *Journal of Occupational Rehabilitation*, 8(3), 179-189.
- McLeavey, B. C., Daly, R. J., Ludgate, J. W., & Murray, C. M. (1994). Interpersonal problem-solving skills training in the treatment of self-poisoning patients. *Suicide and life threatening behavior*, 24(4), 382-394.
- McNair, D. M., Lorr, M., & Droppleman, L. F. (1971). *EdITS Manual for the Profile Of Mood States*. San Diego (Cal.): Educational and Industrial Testing Service.

- Meerding, W. J., Bonneux, L., Polder, J. J., Koopmanschap, M. A., & van der Maas, P. J. (1998). Demographic and epidemiological determinants of healthcare costs in Netherlands: cost of illness study. *Bmj*, 317(7151), 111-115.
- Melzack, R. (1975). The McGill Pain Questionnaire: major properties and scoring methods. *Pain*, 1(3), 277-299.
- Melzack, R., & Wall, P. D. (1965). Pain mechanisms: a new theory. *Science*, 150(699), 971-979.
- Mendonca, J. D., & Siess, T. F. (1976). Counseling for indecisiveness: Problem-solving and anxiety-management training. *Journal of Counseling Psychology*, 23(4), 339-347.
- Mikulincer, M. (1986). Attributional processes in the learned helplessness paradigm: behavioral effects of global attributions. *Journal of personality and social psychology*, 51(6), 1248-1256.
- Mikulincer, M. (1988). The relation between stable/unstable attribution and learned helplessness. *Br J Soc Psychol*, 27(Pt 3), 221-230.
- Morley, S. (1989). Single case research. In G. Parry & F. N. Watts (Eds.), *Behavioural and mental health research: a handbook of skills and methods* (pp. 233-264). Hillsdale: Lawrence Erlbaum Associates.
- Morley, S., & Adams, M. (1989). Some simple statistical tests for exploring single-case time-series data. *British Journal of Clinical Psychology*, 28(1), 1-18.
- 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(1-2), 1-13.
- Mueser, K. T., Valenti Hein, D., & Yarnold, P. R. (1987). Dating-skills groups for the developmentally disabled: Social skills and problem-solving versus relaxation training. *Behavior Modification*, 11(2), 200-228.
- Murray, H. A. (1943). *Thematic Apperception Test (TAT), manual*. Harvard: The president and fellows of Harvard College, Harvard Psychological Clinic.
- Mynors Wallis, L. (1996). Problem-solving treatment: Evidence for effectiveness and feasibility in primary care. *International Journal of Psychiatry in Medicine*, 26(3), 249-262.
- Mynors Wallis, L., Davies, I., Gray, A., Barbour, F., & Gath, D. (1997). A randomised controlled trial and cost analysis of problem-solving treatment for emotional disorders given by community nurses in primary care. *The British journal of psychiatry*, 170(3), 113-119.
- Mynors Wallis, L. M., Gath, D. H., Lloyd Thomas, A. R., & Tomlinson, D. (1995). Randomised controlled trial comparing problem solving treatment with amitriptyline and placebo for major depression in primary care. *Bmj*, 310(6977), 441-445.
- Neumann, W., Kugler, J., Pfand Neumann, P., Schmitz, N., Seelbach, H., & Kruskemper, G. M. (1997). Effects of pain-incompatible imagery on tolerance of pain, heart rate, and skin resistance. *Perceptual and motor skills*, 84(3 Pt 1), 939-943.
- Nezu, A. M. (1986). Efficacy of a social problem-solving therapy approach for unipolar depression. *Journal of consulting and clinical psychology*, 54(2), 196-202.

- Nezu, A. M., & et al. (1986). Social problem solving as a moderating variable between negative life stress and depressive symptoms. *Cognitive Therapy and Research*, 10(5), 489-498.
- Nezu, A. M., & Nezu, C. M. (1993). Identifying and selecting target problems for clinical interventions: A problem-solving model. *Psychological Assessment*, 5(3), 254-263.
- Nezu, A. M., Nezu, C. M., Friedman, S. H., Faddis, S., & Houts, P. S. (1998). *A problem solving approach. Helping cancer patients to cope*. Washington DC: American Psychological Association.
- Nezu, A. M., & Perri, M. G. (1989). Social problem-solving therapy for unipolar depression: an initial dismantling investigation. *Journal of consulting and clinical psychology*, 57(3), 408-413.
- Nezu, A. M., & Ronan, G. F. (1985). Life stress, current problems, problem solving, and depressive symptoms: an integrative model. *Journal of Consulting and Clinical Psychology*, 53(5), 693-697.
- Nezu, A. M., & Ronan, G. F. (1988). Social problem solving as a moderator of stress-related depressive symptoms: A prospective analysis. *Journal of Counseling Psychology*, 35(2), 134-138.
- Nezu, C. M., Nezu, A. M., Friedman, S. H., Houts, P. S., DelliCarpini, L. A., Bildner, C., & Faddis, S. (1999). Cancer and psychological distress: Two investigations regarding the role of social problem-solving. *Journal of Psychosocial Oncology*, 16(3-4), 27-40.
- Nicholas, M. K., Wilson, P. H., & Goyen, J. (1991). Operant-behavioural and cognitive-behavioural treatment for chronic low back pain. *Behaviour Research and Therapy*, 29(3), 225-238.
- OECD, Statistics Directorate (1996). *Historical statistics 1960 - 1994*. Paris: OECD.
- Orbach, I. (1994). Dissociation, physical pain, and suicide: a hypothesis. *Suicide and life threatening behavior*, 24(1), 68-79.
- Osman, A., Barrios, F. X., Kopper, B. A., Hauptmann, W., Jones, J., & O'Neill, E. (1997). Factor structure, reliability, and validity of the Pain Catastrophizing Scale. *Journal of Behavioral Medicine*, 20(6), 589-605.
- Padawer, W. J., & Levine, F. M. (1992). Exercise-induced analgesia: fact or artifact? *Pain*, 48(2), 131-135.
- Parloff, M. B. (1986). Placebo controls in psychotherapy research: A sine qua non or a placebo for research problems? *Journal of Consulting and Clinical Psychology*, 54(1), 79-87.
- Patrick, D. L., Deyo, R. A., Atlas, S. J., Singer, D. E., Chapin, A., & Keller, R. B. (1995). Assessing health-related quality of life in patients with sciatica. *Spine*, 20(17), 1899-1908.
- Peters, M. L., Vlaeyen, J. W., & van Drunen, C. (2000). Do fibromyalgia patients display hypervigilance for innocuous somatosensory stimuli? Application of a body scanning reaction time paradigm. *Pain*, 86(3), 283-292.
- Pfiffner, L. J., Jouriles, E. N., Brown, M. M., Etscheidt, M. A., & et al. (1990). Effects of problem-solving therapy on outcomes of parent training for single-parent families. *Child and Family Behavior Therapy*, 12(1), 1-11.

- Philips, H. C. (1987). Avoidance behaviour and its role in sustaining chronic pain. *Behaviour Research and Therapy*, 25(4), 273-279.
- Philips, H. C., Grant, L., & Berkowitz, J. (1991). The prevention of chronic pain and disability: A preliminary investigation. *Behaviour Research and Therapy*, 29(5), 443-450.
- Picavet, H. S. J., Schouten, J. S. A. G., & Smit, H. A. (1996). *Prevalenties en consequenties van lage-rugklachten in het MORGEN-project 1993-1995* (263200004). Bilthoven: Rijksinstituut voor volksgezondheid en milieu (RIVM).
- Pincus, T., & Williams, A. (1999). Models and measurements of depression in chronic pain. *Journal of Psychosomatic Research*, 47(3), 211-219.
- Platt, J. J., Husband, S. D., Hermalin, J., Cater, J., & et al. (1993). A cognitive problem-solving employment readiness intervention for methadone clients. *Journal of Cognitive Psychotherapy*, 7(1), 21-33.
- Platt, J. J., Prout, M. F., & Metzger, D. S. (1987). Interpersonal cognitive problem-solving therapy (ICPS). In W. Dryden & W. L. Golden (Eds.), *Cognitive-behavioural approaches to psychotherapy* (pp. 261-290). Washington, DC: Hemisphere Publishing Corp.
- Reeson, K. A., & Craig, K. (1988). Medically incongruent chronic back pain: physical limitations, suffering and ineffective coping. *Pain*, 32, 35-45.
- Richards, J. S., Nepomuceno, C., Riles, M., & Suer, Z. (1982). Assessing pain behavior: the UAB Pain Behavior Scale. *Pain*, 14(4), 393-398.
- Riley, J. L., Robinson, M. E., Wise, E. A., Myers, C. D., & Fillingim, R. B. (1998). Sex differences in the perception of noxious experimental stimuli: a meta-analysis. *Pain*, 74(2-3), 181-187.
- Robinson, M. E., & Riley, J. L., III. (1999). The role of emotion in pain. In R. Gatchel & D. Turk (Eds.), *Psychosocial factors in pain: Critical perspectives*. New York: The Guilford Press.
- Roland, M., & Morris, R. (1983). A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. *Spine*, 8(2), 141-144.
- Romano, J. M., & Turner, J. A. (1985). Chronic pain and depression: does the evidence support a relationship? *Psychological bulletin*, 97(1), 18-34.
- Rosenstiel, A. K., & Keefe, F. J. (1983). The use of coping strategies in chronic low back pain patients: Relationship to patient characteristics and current adjustment. *Pain*, 17(1), 33-44.
- Salkovskis, P. M., Atha, C., & Storer, D. (1990). Cognitive-behavioural problem solving in the treatment of patients who repeatedly attempt suicide. A controlled trial. *The British journal of psychiatry; the journal of mental science*, 157, 871-876.
- Schmidt, N. B., & Cook, J. H. (1999). Effects of anxiety sensitivity on anxiety and pain during a cold pressor challenge in patients with panic disorder. *Behaviour research and therapy*, 37(4), 313-323.
- Schneider, W. (1988). Micro Experimental Laboratory: An integrated system for IBM-PC compatibles. *Behavior Research Methods, Instrumentation, and Computers*, 20, 206-217.

- Schoppink, L. E., van Tulder, M. W., Koes, B. W., Beurskens, S. A., & de Bie, R. A. (1996). Reliability and validity of the Dutch adaptation of the Quebec Back Pain Disability Scale. *Physical Therapy*, 76(3), 268-275.
- Schwartz, M. D., Lerman, C., Audrain, J., Cella, D., Rimer, B., Stefanek, M., Garber, J., Lin, T. H., & Vogel, V. (1998). The impact of a brief problem-solving training intervention for relatives of recently diagnosed breast cancer patients. *Annals of behavioral medicine*, 20(1), 7-12.
- Shaw, W. S., Feuerstein, M., Hauffer, A. J., Berkowitz, S. M., & Lopez, M. S. (2001). Working with low back pain: problem-solving orientation and function. *Pain*, 93, 129-137.
- Shure, M. B., & Spivack, G. (1981). The problem-solving approach to adjustment: A competency-building model of primary prevention. *Prevention in Human Services*, 1(1-sup-2), 87-103.
- Spinhoven, P., & Linssen, A. C. (1991). Behavioral treatment of chronic low back pain: I. Relation of coping strategy use to outcome. *Pain*, 45(1), 29-34.
- Spitzer, W. O., LeBlanc, F. E., & Dupuis, M. (1987). Scientific approach to the assessment and management of activity-related spinal disorders: report of the Quebec task Force on Spinal Disorders. *Spine*, 12(suppl 7), S1-S59.
- Spivack, G., Platt, J. J., & Shure, M. B. (1976). *The problem solving approach to adjustment*. San Francisco: Jossey-Bass.
- Spivack, G., & Shure, M. B. (1974). *Social adjustment of young children: A cognitive approach to solving real-life problems*. San Francisco: Jossey-Bass.
- Stark, K. D., Reynolds, W. M., & Kaslow, N. J. (1987). A comparison of the relative efficacy of self-control therapy and a behavioral problem-solving therapy for depression in children. *Journal of abnormal child psychology*, 15(1), 91-113.
- Staub, E., Tursky, B., & Schwartz, G. E. (1971). Self-control and predictability: their effects on reactions to aversive stimulation. *Journal of personality and social psychology*, 18(2), 157-162.
- Stegen, K. (1998). *Het verband tussen negatieve affectiviteit en psychosomatische klachten (The relationship between negative affectivity and psychosomatic complaints, doctoral thesis)*. University of Leuven, Belgium.
- Sternbach, R. A. (1986). Pain and 'hassles' in the United States: findings of the Nuprin pain report. *Pain*, 27(1), 69-80.
- Stratford, P. W., Binkley, J. M., Riddle, D. L., & Guyatt, G. H. (1998). Sensitivity to change of the Roland-Morris Back Pain Questionnaire: part 1. *Physical therapy*, 78(11), 1186-1196.
- Strayhorn, J. M. (1987). Control groups for psychosocial intervention outcome studies. *American Journal of Psychiatry*, 144(3), 275-282.
- Sullivan, M. J., & D'Eon, J. L. (1990). Relation between catastrophizing and depression in chronic pain patients. *Journal of abnormal psychology*, 99(3), 260-263.
- Sullivan, M. J. L., Bishop, S. R., & Pivik, J. (1995). The Pain Catastrophizing Scale: Development and validation. *Psychological Assessment*, 7(4), 524-532.
- Tellegen, A. (1982). *Brief manual for the Differential Personality Questionnaire, unpublished manuscript*. Minneapolis, MN: University of Minnesota.

- Tellegen, A., & Waller, N. G. (1989). Exploring personality through test construction: Development of the Multidimensional Personality Questionnaire. In S. R. Briggs & J. M. Cheek (Eds.), *Personality Measures: Development and Evaluation* (Vol. 1). Greenwich, CN: JAI Press.
- Tisdelle, D. A., & St Lawrence, J. S. (1986). Interpersonal problem-solving competency: Review and critique of the literature. *Clinical Psychology Review*, 6(4), 337-356.
- Toseland, R. W., Blanchard, C. G., & McCallion, P. (1995). A problem solving intervention for caregivers of cancer patients. *Social science and medicine*, 40(4), 517-528.
- Turk, D. C., Rudy, T. E., & Sorkin, B. A. (1993). Neglected topics in chronic pain treatment outcome studies: determination of success. *Pain*, 53(1), 3-16.
- Turner, J. A. (1982). Comparison of group progressive-relaxation training and cognitive-behavioral group therapy for chronic low back pain. *Journal of Consulting and Clinical Psychology*, 50(5), 757-765.
- Turner, J. A. (1996). Educational and behavioral interventions for back pain in primary care. *Spine*, 21(24), 2851-2857.
- Turner, J. A., & Clancy, S. (1988). Comparison of operant behavioral and cognitive-behavioral group treatment for chronic low back pain. *Journal of Consulting and Clinical Psychology*, 56(2), 261-266.
- Turner, J. A., Clancy, S., McQuade, K. J., & Cardenas, D. D. (1990). Effectiveness of behavioral therapy for chronic low back pain: A component analysis. *Journal of Consulting and Clinical Psychology*, 58(5), 573-579.
- Turner, J. A., Deyo, R. A., Loeser, J. D., Von Korff, M., & Fordyce, W. E. (1994). The importance of placebo effects in pain treatment and research. *Jama*, 271(20), 1609-1614.
- Turner, J. A., & Jensen, M. P. (1993). Efficacy of cognitive therapy for chronic low back pain. *Pain*, 52(2), 169-177.
- Turner, J. A., Jensen, M. P., & Romano, J. M. (2000). Do beliefs, coping, and catastrophizing independently predict functioning in patients with chronic pain? *Pain*, 85(1-2), 115-125.
- Van den Hout, J. H., Vlaeyen, J. W., Houben, R. M., Soeters, A. P., & Peters, M. L. (2001a). The effects of failure feedback and pain-related fear on pain report, pain tolerance, and pain avoidance in chronic low back pain patients. *Pain*, 92(1-2), 247-257.
- Van den Hout, J. H., Vlaeyen, J. W., Peters, M. L., Engelhard, I. M., & van den Hout, M. A. (2000). Does failure hurt? The effects of failure feedback on pain report, pain tolerance and pain avoidance. *European journal of pain*, 4, 335-346.
- Van den Hout, J. H. C., Vlaeyen, J. W. S., Heuts, P. H. T. G., De Vet, H. C. W., Sillen, W. J. T., Willen, A. J. E. H. L., Wijnen, J. A. G., & Passchier, J. (under review). Has problem solving therapy supplemental value when added to behavioral graded activity in non-specific low back pain patients? A randomized clinical trial.
- Van den Hout, J. H. C., Vlaeyen, J. W. S., Heuts, P. H. T. G., Sillen, W. J. T., & Willen, A. J. E. H. L. (2001b). Functional disability in non-specific low back pain: The role of pain-related fear and problem-solving skills. *IJBM*, 8(2), 134-148.

- Van den Hout, J. H. C., Vlaeyen, J. W. S., Heuts, P. H. T. G., Zijlema, J. H. L., & Wijnen, J. A. G. (accepted pending revision). Secondary prevention of work-related disability in non-specific low back pain: Does problem-solving therapy help? A randomized clinical trial.
- Van den Hout, J. H. C., Vlaeyen, J. W. S., Kole-Snijders, A. M. J., Heuts, P. H. T. G., Willen, J. E. H. L., & Sillen, W. J. T. (1998). Graded activity and problem solving therapy in sub-acute non-specific low back pain. *Physiotherapy*, 84, 167.
- Van den Velden, J., De Bakker, J. H., Claessens, A. A. M. C., & Schellevis, M. C. (1991). *Een nationale studie naar ziekten en verrichtingen in de huisartspraktijk (A national survey of diseases and operations in General Practice)*. Utrecht: NIVEL.
- Van der Giezen, A. M., Bouter, L. M., & Nijhuis, F. J. (2000). Prediction of return-to-work of low back pain patients sicklisted for 3-4 months. *Pain*, 87(3), 285-294.
- Van der Heijden, G. J., Beurskens, A. J., Koes, B. W., Assendelft, W. J., de Vet, H. C., & Bouter, L. M. (1995). The efficacy of traction for back and neck pain: a systematic, blinded review of randomized clinical trial methods. *Physical therapy*, 75(2), 93-104.
- Van der Kloot, W. A., Oostendorp, R. A., van der Meij, J., & van den Heuvel, J. (1995). De Nederlandse versie van 'McGill pain questionnaire': een betrouwbare pijnvragenlijst. *Nederlands Tijdschrift Voor Geneeskunde*, 139(13), 669-673.
- Van der Ploeg, H. M., Defares, P. B., & Spielberger, C. D. (1980). *Handleiding bij de Zelfbeoordelingsvragenlijst ZBV. Een Nederlandse bewerking van de Spielberger State-Trait Anxiety Inventory STAI-DY (Manual of the ZBV. A Dutch version of the Spielberger STAI-DY)*. Lisse: Swets & Zeitlinger.
- Van Tulder, M. W., Assendelft, W. J., Koes, B. W., & Bouter, L. M. (1997a). Method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for Spinal Disorders. *Spine*, 22(20), 2323-2330.
- Van Tulder, M. W., Assendelft, W. J., Koes, B. W., & Bouter, L. M. (1997b). Spinal radiographic findings and nonspecific low back pain. A systematic review of observational studies. *Spine*, 22(4), 427-434.
- Van Tulder, M. W., Koes, B. W., & Bouter, L. M. (1995). A cost-of-illness study of back pain in The Netherlands. *Pain*, 62(2), 233-240.
- Van Tulder, M. W., Koes, B. W., & Bouter, L. M. (1997c). Conservative treatment of acute and chronic nonspecific low back pain. A systematic review of randomized controlled trials of the most common interventions. *Spine*, 22(18), 2128-2156.
- Van Tulder, M. W., Ostelo, R. W., Vlaeyen, J. W., Linton, S. J., Morley, S. J., & Assendelft, W. J. (2000). Behavioural treatment for chronic low back pain. *Cochrane database of systematic reviews*, The(2), Cd002014.
- Van Veldhoven, M., & Meijman, T. (1994). Ontwikkeling van de VBBA (the development of the questionnaire Perception and Evaluation of work), *Het meten van psychosociale werkbelasting. De Vragenlijst Beleving en Beoordeling van Arbeid (VBBA) (Measurement of psychosocial work load)* (pp. 71-97). Amsterdam: NIA.
- Vingerhoets, A. J. J. M., & Van Tilburg, M. A. L. (1994). *Alledaagse Problemen Lijst, APL (Dutch Manual of the Everyday Problem Checklist, EPCL)*. Lisse: Swets & Zeitlinger.

- Vlaeyen, J. W., Haazen, I. W., Schuerman, J. A., Kole Snijders, A. M., & van Eek, H. (1995a). Behavioural rehabilitation of chronic low back pain: comparison of an operant treatment, an operant-cognitive treatment and an operant-respondent treatment. *Br J Clin Psychol*, 34(Pt 1), 95-118.
- Vlaeyen, J. W., Kole Snijders, A. M., Boeren, R. G., & van Eek, H. (1995b). Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain*, 62(3), 363-372.
- Vlaeyen, J. W., & Linton, S. J. (2000). Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. *Pain*, 85(3), 317-332.
- Vlaeyen, J. W., Pernot, D. F., Kole Snijders, A. M., Schuerman, J. A., Van Eek, H., & Groenman, N. H. (1990). Assessment of the components of observed chronic pain behavior: the Checklist for Interpersonal Pain Behavior (CHIP). *Pain*, 43(3), 337-347.
- Vlaeyen, J. W., Seelen, H. A., Peters, M., de Jong, P., Aretz, E., Beisiegel, E., & Weber, W. E. (1999). Fear of movement/(re)injury and muscular reactivity in chronic low back pain patients: an experimental investigation. *Pain*, 82(3), 297-304.
- Vlaeyen, J. W., Teeken Gruben, N. J., Goossens, M. E., Rutten van Molken, M. P., Pelt, R. A., van Eek, H., & Heuts, P. H. (1996). Cognitive-educational treatment of fibromyalgia: a randomized clinical trial. I. Clinical effects. *The Journal of rheumatology*, 23(7), 1237-1245.
- Vlaeyen, J. W. S. (1991). *Chronic low back pain: Assessment and treatment from a behavioral rehabilitation perspective*. Lisse, Netherlands: Swets and Zeitlinger.
- Vlaeyen, J. W. S., de Jong, J., Geilen, M., Heuts, P. H. T. G., & van Breukelen, G. (2001). Graded exposure in vivo in the treatment of pain-related fear: A replicated single-case experimental design in four patients with chronic low back pain. *Behaviour Research and Therapy*, 39, 151-166.
- Vlaeyen, J. W. S., Kole-Snijders, A. M. J., Rotteveel, A., Ruesink, R., & Heuts, P. H. T. G. (1995c). The role of fear of movement/(re)injury in pain disability. *Journal of Occupational Rehabilitation*, 5, 235-252.
- Waddell, G. (1987). 1987 Volvo award in clinical sciences. A new clinical model for the treatment of low-back pain. *Spine*, 12(7), 632-644.
- Waddell, G. (1998). A new clinical model of low back pain and disability, *The back pain revolution* (pp. 223-240). London: Churchill Livingstone.
- Waddell, G., Newton, M., Henderson, I., Somerville, D., & Main, C. J. (1993). A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain*, 52(2), 157-168.
- Waddell, G., & Turk, D. C. (1992). Clinical assessment of low back pain. In D. C. Turk & R. Melzack (Eds.), *Handbook of pain assessment*. New York, London: The Guilford Press.
- Wade, J. B., Price, D. D., Hamer, R. M., Schwartz, S. M., & Hart, R. P. (1990). An emotional component analysis of chronic pain. *Pain*, 40(3), 303-310.
- Wald, F. D. M., & Mellenberg, G. J. (1990). De verkorte versie van de Nederlandse vertaling van de Profile of Mood States (POMS) (The shortened version of the Dutch translation of the Profile of Mood States). *Nederlands Tijdschrift voor de Psychologie*, 45, 86-90.

- Ware, J. J., Kosinski, M., & Keller, S. D. (1995). A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Med Care*, 34, 220-233.
- Watson, D., & Clark, L. A. (1984). Negative affectivity: the disposition to experience aversive emotional states. *Psychological bulletin*, 96(3), 465-490.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology*, 54(6), 1063-1070.
- Watson, D., & Pennebaker, J. W. (1989). Health complaints, stress, and distress: exploring the central role of negative affectivity. *Psychological review*, 96(2), 234-254.
- Weisenberg, M., Tepper, I., & Schwarzwald, J. (1995). Humor as a cognitive technique for increasing pain tolerance. *Pain*, 63(2), 207-212.
- Wilkinson, P., & Mynors Wallis, L. (1994). Problem-solving therapy in the treatment of unexplained physical symptoms in primary care: a preliminary study. *Journal of psychosomatic research*, 38(6), 591-598.
- Williams, A. C. (1995). Pain measurement in chronic pain management. *Pain Reviews*, 2(1), 39-63.
- Williams, A. C., Pither, C. E., Richardson, P. H., Nicholas, M. K., Justins, D. M., Morley, S., Diamond, A., Linton, S., Vlaeyen, J., Nilges, P., & Eccleston, C. (1996). The effects of cognitive-behavioural therapy in chronic pain. *Pain*, 65(2-3), 282-284.
- Yu, P., Harris, G. E., Solovitz, B. L., & Franklin, J. L. (1986). A social problem-solving intervention for children at high risk for later psychopathology. *Journal of Clinical Child Psychology*, 15(1), 30-40.
- Zelman, D. C., Howland, E. W., Nichols, S. N., & Cleeland, C. S. (1991). The effects of induced mood on laboratory pain. *Pain*, 46(1), 105-111.
- Zillmann, D., de Wied, M., King Jablonski, C., & Jenzowsky, S. (1996). Drama-induced affect and pain sensitivity. *Psychosomatic medicine*, 58(4), 333-341.

SUMMARY

Low back pain is one of the most widespread and expensive problems of the western world. Especially chronic pain problems are responsible for high costs in terms of work loss and disability pensions. In order to prevent chronic pain disability, the development and evaluation of preventive interventions are of socio-economic importance. During the last decades, empirical studies contributed to a growing knowledge that psychosocial variables play an important role in the development and maintenance of pain problems. Consequently, cognitive-behavioral techniques were introduced in the management of pain and shown to be effective. Nevertheless, there is still room for improvement with regard to 1) the secondary prevention of pain, i.e. to intervene early before pain becomes chronic, and 2) to maintain positive treatment effects in the long term.

In this thesis, pain-related fear and problem solving are introduced as mechanisms that might be important in the prevention of chronic pain. From a theoretical point of view, avoidance learning is postulated to be one of the mechanisms that explain the association between pain, stress and disability levels. Several studies have reported on the efficacy of interventions aimed at reducing the avoidance of activities in pain patients, of which the graded activity approach is the most known. Additionally, there is evidence that the effects of graded activity can be improved when the treatment also focuses on resolving daily hassles and stress for which the patient perceives a lack of control. Problem-solving therapy is alleged to increase a sense of control in situations that are perceived as stressful, and with regard to the negative consequences of low back pain. The studies in this thesis will address questions regarding the role of problem solving and problem-solving therapy in addition to the contemporary theories and treatment of low back pain.

Chapter 1 introduces the themes that play the leading parts in this thesis: interpersonal problem solving, low back pain, pain-related fear, failure experiences, and secondary prevention. Next to the state-of-the-art, hypotheses are formulated about relations between the key themes of this thesis. In particular the role of problem solving in pain and the application of theoretical models of pain in corresponding therapies are described. This introductory chapter concludes with an outline of this thesis and phrases the research questions that will be addressed in the subsequent chapters.

Chapter 2 systematically reviews the literature regarding the efficacy of problem solving therapy in mental health care, including behavioral medicine. Nineteen randomized controlled trials were systematically selected to be included in a methodologically quality assessment and a structured data-extraction and -analysis procedure. Selected studies cover a wide range of populations and measures. Internal validity of the studies turned out to be high in only a small part of the studies (26%). The problem solving therapy was found to be

effective in most studies (72%), whereas effects were maintained in 80% of the studies including a follow-up assessment. Remarkably, highest success rates of the problem-solving therapy were found when the methodological quality was high. Domains in which the therapy yielded the highest success-rates were problem-solving skills, quality of life, domain specific outcomes, and outcomes referring to general improvement. Moreover, those protocols including more than 9 sessions of problem-solving therapy, were the most effective in nearly 100% of the comparisons.

Chapter 3 and 4 test the hypothetical model that underlies the randomized clinical trial that is described in chapter 6 and 7. It is hypothesized that lack of control, with regard to non-pain-related situations and the experience of failure resulting from it would result in increased pain and pain behavior.

Chapter 3 replicates a laboratory experiment addressing the question "Does failure hurt?". Furthermore, the role of negative affectivity was investigated 1) as a mood state (mediator), and 2) as a personality trait (moderator). Fifty-four female students were randomly assigned to one of three conditions: (1) failure feedback, (2) success feedback, and (3) a neutral control task. Failure and success experiences were manipulated by means of an intelligence task, after which false failure or success feedback was given. Following the manipulation, a second 'blind' experimenter conducted a cold pressor task in order to obtain pain (intensity, tolerance and avoidance) measures. Outcomes were in line with the original study, and showed that failure feedback was followed by increased pain report and decreased pain tolerance as compared to success feedback. In comparison with the neutral control task, however, no statistical significant differences were found. The role of negative affectivity, either as a mediator or moderator, was not confirmed in this study.

Chapter 4 describes a laboratory experiment, in which the role of non-pain-related failure and pain-related fear in chronic low back pain patients was investigated. Again, the potential role of negative affectivity in the relation between failure experiences and pain was examined. Seventy-six patients were divided into high and low pain-related fear groups, and within each group they were randomly assigned to either failure or success feedback condition. In the first part of the experiment patients conducted a 'social empathy test', followed by manipulated false failure or success feedback. Subsequently, a second experimenter, who was blind to the condition, conducted two lifting tasks in order to obtain pain report, pain tolerance and pain avoidance measures. With regard to the influence of failure feedback, outcomes were opposite to the a priori suggested hypothesis: Failure feedback resulted in decreased pain report and pain avoidance, and increased pain tolerance. Furthermore, trait negative affectivity moderated the effects of failure on pain avoidance, in that only those patients without negative affectivity, showed decreased avoidance during the second lifting task. In the presence of negative affectivity, neither failure nor success influenced subsequent pain behavior. The mediating role of negative affect was not confirmed. In line with current

literature, pain-related fear resulted in increased pain report and lowered pain tolerance. Pain intensity preceding the lifting task, was the strongest predictor with regard to all pain measures. The role of pain-related fear and the unexpected findings regarding non-pain-related failure are discussed, as well as some clinical implications of the findings.

Chapter 5 describes the results of a cross-sectional study investigating the role of problem-solving, exclusively and next to other psychological variables in the model of functional disability. In addition, it was investigated whether problem-solving moderated the influence of daily stress on functional disability. One-hundred-and-twenty-three employees with a recent new episode of sick leave as a result of back pain were included in the study. Psychological variables, i.e. pain-related fear, pain intensity, pain catastrophizing, daily stress and problem solving, all were individually related to functional disability. The supplemental contribution of problem solving variables to the predictive model of functional disability, however, was not found, neither could the moderating role of problem solving be confirmed. Pain intensity and pain catastrophizing turned out to be the strongest predictors. Results were discussed in the light of theoretical implications of the findings as well as their importance to clinical practice.

Chapter 6 and 7 address the question whether problem-solving therapy has supplemental value when added to graded activity in non-specific low back pain patients. A randomized clinical trial was conducted in employees with a recent new episode of sick leave as a result of non-specific low back pain. Patients were referred to the rehabilitation center by general practitioner, occupational physician, or rehabilitation physician. In total, 115 low back pain patients were included in the study and randomly assigned to 1) graded activity + problem solving therapy (GAPS; $n = 58$) or graded activity + group education (GAGE; $n = 57$). Graded activity is an operant behavioral treatment that aims to increase activity levels using quota systems. Problem solving therapy is a cognitive behavioral therapy in which problem-solving skills are taught according to the model by D'Zurilla and Goldfried (1971). In the current study, the problem-solving therapy consisted of 10 ninety-minute sessions. Group education (attention control) consisted of 10 ninety-minute lessons about issues related to the back and back pain. Patients were assessed pre-treatment, post-treatment, and at 6- and 12-month follow-up. Next to assessments at the rehabilitation center, patients kept a cost diary during the year following the intervention, and data were retrieved from the occupational health services regarding sick leave and work status the year preceding and following the intervention. Primary outcome measures were functional disability, physical health activities, health care utilization, days of sick leave, and work status.

Chapter 6 in particular focuses on the outcomes of self-report measures, the cost diary, and behavioral observations. Next to the effect-evaluation, a process-evaluation was conducted and described in this chapter. The manipulation check showed that the manipulation succeeded in that problem-solving

skills were exclusively changed in the GAPS condition. At post-treatment, both conditions significantly improved compared to pre-treatment with respect to primary and secondary outcome measures. At 6-month follow-up functional disability still decreased relative to post-treatment, as well as pain behavior in the GAPS condition. Improvements sustained at 6- and 12-month follow-up. Next, between-group differences were evaluated. The supplemental value of problem-solving therapy became clear in the long term, at 12-month follow-up: Patients in the GAPS condition reported significantly less functional disability (RDQ), and relatively more patients in this condition had a clinically relevant improvement in relation to pre-treatment functional disability scores. Moreover, according to the cost diary, a higher proportion of patients in the GAPS condition performed physical health activities during the second half-year follow-up. However, regarding health care utilization and days of sick leave, both assessed by means of the cost diary, the hypothesized supplemental value of the problem solving therapy was not found. Intention-to-treat and per-protocol analyses, confirmed findings regarding primary outcomes. The process evaluation found patients in GAPS and GAGE conditions equally content about therapy and their personal improvement. Nevertheless, more patients in the GAGE condition agreed that the therapy came up to their expectations, and that therapy goals were attained. Finally, methodological issues and suggestions for further research and clinical practice are discussed in the final part of this chapter.

Chapter 7, also addresses the supplemental value of problem-solving therapy, but in this study, data from an objective data source (occupational health services) are used to evaluate outcomes regarding days of sick leave and work status. Data from 84 of 115 patients included in the study were retrieved from the occupational health services' registration. Baseline characteristics of the patients involved in this part of the evaluation did not differ from the original sample neither was the randomization violated. Although both conditions decrease days of sick leave during the first half-year following the intervention, this decrease is exclusively continued in the GAPS condition during the second half-year. Patients who received problem solving therapy in addition to graded activity, had significantly less days of sick leave, both as a result of back pain and in general, than those who did not receive the additional problem-solving therapy. Sensitivity analyses confirmed these findings. Finally, the work status of patients is described one week before the intervention, at 6- and 12-month follow-up. In line with previous findings, differences between treatment conditions are most striking one year after the intervention, favoring the GAPS condition. First, 86% of patients in this condition made a 100% return-to-work as compared to 67% in the GAGE condition. Second, a remarkably lower proportion of patients in the GAPS condition received disability pensions after one year, i.e. 9 versus 21% in the GAGE condition. Both chapters 6 and 7 reveal that the problem solving therapy demonstrated to have supplemental value in addition to graded activity. With regard to days of sick leave, data of the cost

diary were not in agreement with those from the occupational health services, in that an effect was exclusively found in the latter. The difference is discussed. The chapter closes with some recommendations in the light of the secondary prevention of pain and further research.

Chapter 8 summarizes the answers to research questions addressed in this thesis. Of each chapter, conclusions and major points of discussion are described. With regard to the randomized controlled trial, methodological issues, external validity, generalization to clinical practice, and limitations are discussed. Furthermore, suggestions are made for questions to be addressed in further research. Finally, this chapter concludes with a general consideration putting problem solving in a broader perspective regarding its contribution to theoretical models of pain, as well as to the treatment of pain patients.

SAMENVATTING

Lage rugpijn is één van de meest wijdverbreide en dure gezondheidsklachten van onze westerse samenleving. Met name de chronische rugpijn is verantwoordelijk voor hoge kosten in termen van ziekteverzuim en arbeidsongeschiktheid. Ten einde chronische pijn te voorkomen, is de ontwikkeling en evaluatie van preventieve interventies van individueel en sociaal-economisch belang. In de afgelopen decennia heeft empirisch onderzoek aangetoond dat psychosociale factoren een belangrijke invloed hebben op pijn, en het chronisch worden van pijn. Hieruit voortvloeiend werden met succes, cognitief-gedragsmatige technieken geïntroduceerd in de behandeling van pijn. Desalniettemin is er ruimte voor verbetering wat betreft 1) de secundaire preventie van pijn, ofwel het vroegtijdig interveniëren voordat het proces onomkeerbaar wordt, en 2) het behoud van effecten op de langere termijn.

In dit proefschrift worden pijn-gerelateerde vrees en het oplossen van problemen geïntroduceerd als potentieel belangrijke mechanismen in het verklarende model van chronische pijn. In het biopsychosociale model van pijn, wordt vermijdingsgedrag gezien als een belangrijke link tussen pijn, stress en functionele beperkingen. Verschillende studies hebben inmiddels aangetoond dat interventies die erop gericht zijn om de vermijding van activiteiten te doorbreken, zoals *graded activity*, effectief zijn. Tevens werd gevonden dat de effectiviteit van *graded activity* vergroot kan worden wanneer de behandeling zich ook zou richten op stressvolle gebeurtenissen en alledaagse problemen voor welke de patiënt een gebrek aan controle ervaart. Van de probleemoplossende vaardigheidstraining wordt beweerd dat deze het gevoel van controle vergroot in situaties die als stressvol worden ervaren, zoals bijvoorbeeld de negatieve consequenties van lage rugpijn. De studies in dit proefschrift hebben allen tot doel om antwoorden te vinden op vragen betreffende de rol van probleem-oplossen en de probleemoplossende vaardigheidstraining in de bestaande theoretische modellen en behandelingen van lage rugpijn.

Hoofdstuk 1 introduceert de thema's die een hoofdrol spelen in dit proefschrift: interpersoonlijke problem solving, lage rugpijn, pijn-gerelateerde vrees, faal ervaringen, en secundaire preventie. Naast de *state-of-the-art*, worden hypothesen gevormd over mogelijke relaties tussen de hoofdthema's van dit proefschrift. In het bijzonder de rol van probleem oplossen bij pijn, en de toepassing van theoretische pijn-modellen in corresponderende therapieën, worden beschreven. De algemene introductie wordt besloten met een overzicht van het gehele proefschrift en formuleert de onderzoeksvragen die in de achtereenvolgende hoofdstukken aan bod zullen komen.

Hoofdstuk 2 geeft een systematisch overzicht van de literatuur betreffende de effectiviteit van probleemoplossende vaardigheidstraining in de geestelijke gezondheidszorg, met inclusie van de medische psychologie. Negentien gerandomiseerde en gecontroleerde trials werden systematisch geselecteerd.

Vervolgens werd een methodologische kwaliteitsmeting en een gestructureerde data-extractie en -analyse procedure uitgevoerd. De geselecteerde studies omvatten een brede verzameling van populaties en uitkomstmaten. De methodologische kwaliteit van de studies bleek slechts in een klein gedeelte van de studies (26%) als 'hoog' te worden gekwalificeerd. De probleemoplossende vaardigheidstraining kwam in de meeste studies als meest effectief naar voren (72%), terwijl in 80% van de studies waarin ook een follow-up meting was opgenomen, effecten behouden bleven op de langere termijn. Opvallend werden de hoogste successcores behaald in die studies die tevens een hoge methodologische kwaliteit hadden. Uitkomstmaten waarop de probleemoplossende vaardigheidstraining de beste effecten liet zien waren: probleemoplossende vaardigheden, kwaliteit van leven, domein specifieke uitkomsten, en uitkomsten die refereerden aan een algemene vooruitgang. Tenslotte bleken protocollen waarin de probleemoplossende vaardigheidstraining bestond uit meer dan 9 sessies, het meest effectief in bijna alle vergelijkingen.

Hoofdstukken 3 en 4 testen een aspect van de theoretische gedachte die ten grondslag ligt aan de gerandomiseerde en gecontroleerde trial welke in de hoofdstukken 6 en 7 beschreven zal worden. Er wordt verondersteld dat een gebrek aan ervaren controle met betrekking tot niet aan pijn gerelateerde situaties en de faalervaring die daarvan het gevolg is, zou leiden tot een toename van pijn en pijngedrag.

Hoofdstuk 3 is een replicatie van een laboratorium experiment welke een antwoord wil geven op de vraag "Doet falen pijn?". Verder wordt de rol van negatieve affectiviteit onderzocht, ten eerste in de betekenis van stemming op een bepaald moment (mediator), en ten tweede als persoonlijkheidstrekk (moderator). Vierenvijftig eerstejaars studentes werden at random verdeeld over 3 condities: (1) faal feedback, (2) succes feedback, en (3) een neutrale controle taak. Faal- en succeservaringen werden gemanipuleerd door middel van een intelligentietest, waarna 'valse' faal- of succes-feedback gegeven werd. Onmiddellijk na de manipulatie, voerde een tweede geblindeerde experimentator een 'cold pressor' taak uit om pijnmaten (intensiteit, tolerantie, en vermijding) te verkrijgen. De uitkomsten waren in overeenstemming met de uitkomsten zoals die gerapporteerd werden in de originele studie. Er werd gevonden dat faal-feedback inderdaad leidde tot verhoogde pijnrapportage en verlaagde pijn tolerantie in vergelijking tot succes-feedback. Echter, in vergelijking tot studentes die een neutrale controle taak hadden gehad, werden geen statistisch significante verschillen gevonden. De veronderstelde rol van negatieve affectiviteit, als moderator of mediator, werd niet gevonden in deze studie.

Hoofdstuk 4 beschrijft een laboratorium experiment waarin de invloed van niet aan pijn gerelateerd falen en pijn-gerelateerde vrees werd onderzocht in een groep van chronisch lage rugpijn patiënten. Opnieuw werd de rol van negatieve affectiviteit, als mediator en moderator, in relatie tot faalervaringen en pijn bestudeerd. Zesenzeventig patiënten werden eerst onderverdeeld in 2

groepen met respectievelijk een hoge en een lage mate van pijn-gerelateerde vrees. Binnen elke groep werden patiënten vervolgens at random verdeeld in een faal- dan wel een succes-feedback conditie. In het eerste gedeelte van het experiment voerden de patiënten een 'sociale empathie test' uit, gevolgd door gemanipuleerde 'valse' faal- of succes-feedback. Vervolgens voerde een tweede experimentator, die 'geblindeerd' was, 2 tiltaken uit met als doel pijn-rapportage, pijntolerantie, en pijnvermijding te scoren. Met betrekking tot de invloed van faal-feedback, waren uitkomsten in tegenstelling tot de a priori gestelde hypothesen: Faal-feedback leidde tot een verminderde pijnrapportage en -vermijding, en een verhoogde pijntolerantie. Verder werd gevonden dat negatieve affectiviteit de effecten van falen op pijnvermijding modereerde, dat wil zeggen dat alleen patiënten die niet gehinderd werden door de aanwezigheid van trek negatieve affectiviteit, verminderde vermijding toonden tijdens de 2e tiltaak. Echter, in de aanwezigheid van negatieve affectiviteit, beïnvloedde noch falen noch succes, het daarop volgende pijngedrag. De mediërende rol van een negatieve stemming werd niet gevonden. In aansluiting op de bestaande literatuur, bleek pijn-gerelateerde vrees te resulteren in een verhoogde pijnrapportage, en een verlaagde pijntolerantie. Opmerkelijk was dat de pijn die gerapporteerd werd voorafgaande aan de tiltaak, de sterkste voorspeller was van alle pijnmaten. In de discussie worden de rol van pijn-gerelateerde vrees, en de onverwachte bevindingen met betrekking tot niet aan pijn gerelateerd falen besproken, evenals enkele klinische implicaties van de bevindingen.

Hoofdstuk 5 beschrijft de resultaten van een cross-sectionele studie waarin de invloed van probleem oplossen op functionele beperkingen wordt bestudeerd. Zowel de afzonderlijke bijdrage van probleem oplossen, als de toegevoegde waarde aan een psychologisch model worden onderzocht. Daarnaast wordt onderzocht of de invloed van alledaagse problemen op functionele beperkingen wordt gemodereerd door probleem oplossen. Honderden-drieëntwintig werknemers die niet langer dan 20 weken het werk verzuimden als gevolg van lage rugpijn, werden geïncludeerd in deze studie. Psychologische variabelen, namelijk pijn-gerelateerde vrees, pijn intensiteit, pijn catastroferen, alledaagse problemen, en probleem oplossen, bleken allen afzonderlijk gerelateerd aan de mate van functionele beperkingen. De toegevoegde bijdrage van probleem oplossen aan het voorspellende model van functionele beperkingen, werd echter niet gevonden. Tevens werden geen aanwijzingen gevonden voor de veronderstelde modererende rol van probleem oplossen. Pijn intensiteit en pijn catastroferen bleken de sterkste voorspellers in het model. Deze resultaten worden besproken in het licht van theoretische overwegingen als ook wat betreft de klinische betekenis van de bevindingen.

Hoofdstukken 6 en 7 richten zich op de vraag of probleemoplossende vaardigheidstraining een meerwaarde heeft wanneer deze wordt toegevoegd aan graded activity, bij specifieke lage rugpijn patiënten. Een gerandomi-

seerde en gecontroleerde trial werd uitgevoerd bij werknemers die recentelijk het werk waren gaan verzuimen als gevolg van lage rugpijn. De werknemers werden verwezen door huisarts, bedrijfsarts, revalidatiearts, of een combinatie. In het totaal werden 115 lage rugpijnpatiënten geïnccludeerd in de studie en at random toegewezen aan één van twee condities: 1) graded activity + probleemoplossende vaardigheidstraining (GAPS; $n = 58$), of 2) graded activity + groepseducatie (GAGE; $n = 57$). Graded activity is een operant gedragsmatige behandeling welke gericht is op het stapsgewijs verhogen van activiteiten-niveaus volgens een van tevoren vastgesteld individueel schema. Probleemoplossende vaardigheidstraining is een cognitieve gedragstherapie waarin probleemoplossende vaardigheden worden getraind volgens een model van D'Zurilla en Goldfried (1971). In de huidige studie werd de probleemoplossende vaardigheidstraining in een groep uitgevoerd en bestond uit 10 sessies van 1,5 uur. Groepseducatie (aandachts-controle) bestaat eveneens uit 10 groepssessies van 1,5 uur waarin thema's gerelateerd aan de rug en aan rugpijn worden besproken. Patiënten werden gemeten voor- en onmiddellijk na de behandeling, en tijdens 6- en 12-maanden follow-up metingen. Behalve de metingen in het revalidatiecentrum, hielden de patiënten een kostendagboek bij gedurende het jaar na de interventie, en werden data opgevraagd bij de betreffende ARBO-diensten met betrekking tot ziekteverzuim en werkstatus gedurende het jaar vóór en na de behandeling. Primaire uitkomstmaten waren de mate van functionele beperkingen, fysieke gezondheidsactiviteiten, medische consumptie, verzuimdagen, en werkstatus.

Hoofdstuk 6 beschrijft in het bijzonder de resultaten betreffende de zelf-rapportage maten, het kostendagboek, en de observatiematen. Naast de effect-evaluatie wordt in dit hoofdstuk ook een procesevaluatie beschreven. De manipulatie check toonde aan dat de manipulatie geslaagd was, dat wil zeggen dat probleemoplossende vaardigheden alleen in de GAPS conditie toegenomen waren. Ten tijde van de nameting waren beide condities significant veranderd ten opzichte van de voormeting met betrekking tot primaire en secundaire effectmaten. Zes maanden na de interventie is de mate van functionele beperkingen nog steeds afgenomen ten opzichte van de nameting, evenals pijngedrag in de GAPS conditie. De verbeteringen blijven in beide condities behouden op de langere termijn, een jaar na de interventie. Vervolgens wordt het verschil tussen beide condities getoetst. De meerwaarde van probleemoplossende vaardigheidstraining wordt duidelijk tijdens de 2e follow-up meting, 12 maanden na de interventie: patiënten in de GAPS conditie rapporteerden significant minder functionele beperkingen (RDQ), relatief meer patiënten in deze conditie hadden een klinisch relevante verbetering ondergaan ten opzichte van de voormeting. Bovendien rapporteerden een groter percentage van patiënten in de GAPS conditie dat zij fysieke gezondheidsactiviteiten (kostendagboek) hadden uitgevoerd tijdens het 2e halfjaar na de interventie. Met betrekking tot medische consumptie en ziekteverzuim (kostendagboek), bleef de veronderstelde meerwaarde van de probleemoplossende vaardigheidstraining uit. Zowel

de intention-to-treat analyse als de per-protocol analyse bevestigden de eerder gevonden meerwaarde van de probleemoplossende vaardigheidstraining. De procesevaluatie toonde aan dat patiënten in beide condities even tevreden waren over de therapie en hun persoonlijke verbetering. Desalniettemin, bleek ook dat meer patiënten in de GAGE conditie aangaven dat de therapie aan hun verwachtingen had voldaan en dat therapiedoelen waren bereikt. Ten slotte worden methodologische issues, en suggesties voor verder onderzoek en voor de klinische praktijk besproken in de laatste paragraaf van dit hoofdstuk.

Hoofdstuk 7 richt zich, zoals het vorige hoofdstuk, op de vraag of probleemoplossende vaardigheidstraining een meerwaarde heeft wanneer deze wordt toegevoegd aan graded activity. In dit hoofdstuk echter, worden data van een objectieve bron (de ARBO-diensten) gebruikt om uitkomsten met betrekking tot verzuimdagen en werkstatus te evalueren. Data van 84 van de oorspronkelijk 115 geïncludeerde patiënten werden verzameld uit de verzuimregistratie van de ARBO-diensten. Demografische variabelen en scores op de voormetingen verschilden niet tussen de groep betrokken in dit deel van de studie, en de oorspronkelijke groep. Ook de randomisatie check wees uit dat de patiënten in beide condities niet van elkaar verschilden voorafgaande aan de interventie. Ondanks het feit dat beide condities een afname laten zien van het aantal verzuimdagen in het eerste halfjaar na de interventie ten opzichte van de voormeting, zet deze afname alléén door in de GAPS conditie tijdens het 2e halfjaar na de interventie. Wanneer rekening wordt gehouden met de scores in het halfjaar voorafgaande aan de interventie, hadden patiënten die probleemoplossende vaardigheidstraining ontvingen in aanvulling op graded activity, significant minder verzuimdagen, zowel als gevolg van rugpijn en in het algemeen, dan patiënten die de probleemoplossende vaardigheidstraining niet ontvingen. Sensitiviteitsanalyses bevestigden deze resultaten. Tenslotte werd de werkstatus van patiënten beschreven in de week voorafgaande aan de interventie, en 6, en 12 maanden ná de interventie. Overeenkomstig met andere resultaten, zijn de verschillen tussen de condities ook nu weer het opvallendst een jaar na de interventie, ten voordele van de GAPS conditie. Ten eerste blijkt dat 86% van de patiënten in deze conditie weer volledig terug aan het werk is ten opzichte van 68% in de GAGE conditie. Ten tweede, maar hoogst relevant, blijkt dat veel minder patiënten in de GAPS conditie arbeidsongeschikt zijn geraakt (WAO) na één jaar, dan in de GAGE conditie, respectievelijk 9 versus 21%. Beide hoofdstukken, 6 en 7, onthullen dat de probleemoplossende vaardigheidstraining een meerwaarde heeft voor lage rugpijnpatiënten, als toevoeging aan graded activity. Met betrekking tot verzuimdagen, waren de uitkomsten van het kostendagboek niet in overeenstemming met die van de ARBO-diensten. De meerwaarde van de probleemoplossende vaardigheidstraining werd uitsluitend gevonden in de laatste. Het verschil in uitkomsten wordt besproken. Tenslotte worden in de discussie enkele aanbevelingen gedaan ten aanzien van de secundaire preventie van pijn en vervolgonderzoek.

Hoofdstuk 8 vat de antwoorden op de verschillende onderzoeksvragen die

aan bod kwamen in dit proefschrift samen. Per hoofdstuk worden de conclusies en de belangrijkste discussiepunten besproken. Met betrekking tot de gerandomiseerde en gecontroleerde trial worden, i) methodologische issues, ii) de externe validiteit en de generalisatie naar de klinische praktijk, en iii) de tekortkomingen van de studie, besproken. Verder worden suggesties gedaan betreffende vragen voor toekomstig onderzoek. Tenslotte besluit dit hoofdstuk met een algemene beschouwing waarbij 'probleem oplossen' in een breder perspectief worden geplaatst betreffende de bijdrage in theoretische modellen van pijn, maar ook wat betreft de behandeling van pijnpatiënten.

DANKWOORD

Een onderzoek als dit, en de promotie zelf zouden niet mogelijk zijn geweest zonder de steun, inzet en het enthousiasme van velen. Bovendien zou het zonder hen allen een stuk minder leuk zijn geweest. Daarom nu een welgemeend: DANK JULLIE WEL!

Beste Johan, nu aan mij de eer om jou als eerste te bedanken. Een betere copromotor kon ik me niet wensen: onovertroffen. Er waren momenten dat ik twijfelde of het wel allemaal goed zou komen met Better Back. Juist op die momenten waren jouw optimisme en bekrachtiging van cruciaal belang. De inhoudelijke discussies en jouw verfrissende ideeën zette me altijd weer op het goede spoor. Je betrokkenheid gold niet alleen inhoudelijk, maar zeker ook persoonlijk. Met regelmaat kreeg ik een mooi kaartje van je met bemoedigende teksten die een enorme stimulans voor me waren en nog steeds zijn. Johan, bedankt voor alles en ik hoop dat onze wegen zich in de toekomst nog vaak zullen kruisen.

Beste Marcel, veel bewondering heb ik voor jouw wetenschappelijke gedrevenheid. Het siert je als promotor dat je aanzet tot kritisch denken. Bovenal wil ik je bedanken omdat ik lid mocht zijn van jouw club: DMKEP. Onder jouw bevlogen leiding waren er veel mogelijkheden, en een grote vrijheid om wetenschappelijke ideeën te onderzoeken. Het enthousiasme waarmee onderzoek wordt bedreven binnen het departement getuigt hiervan en ik was er maar al te graag bij.

Geen onderzoek zonder deelnemers. Graag wil ik op deze plek alle deelnemers aan Better Back bedanken voor de tijd en moeite die gestoken is in het trouw invullen van vragenlijsten en dagboeken.

De projectgroep Better Back bestond uit een club van mensen die vernieuwing en de hindernissen die daarbij moesten worden genomen niet schuwden. Peter Heuts, Annelies Willen, en Wil Sillen bedankt voor de fijne samenwerking in de projectgroep én in het team.

Ook alle andere Better Back behandelaars van het team 'reuma en pijn', i.h.b. Rob Pelt, Monique Joosten, Noël Dortu, Herman Mulder, en Ank Kole-Snijders, wil ik bedanken voor hun enthousiasme en inzet in de behandeling. Het was en is een genoegen om met jullie te werken. HRC, HCA, en Fitplan bedank ik voor de geboden faciliteiten.

Ank-Kole-Snijders en Sylvia Geurts, problemen oplossen is jullie toevertrouwd! Bedankt voor de leuke samenwerking bij het ontwerpen en uitvoeren van "Hindernissen te lijf". Ank, jij verdient een extra woord van dank omdat ik altijd een beroep heb kunnen doen op jouw brede deskundigheid. Jouw ken-

nis en ervaring met de pijnrevalidatie maken je tot voorbeeld voor deze beginnende pijn-psychooloog.

Dank aan de begeleidingscommissie, Riekje de Vet, Jan Passchier, Jef Wijnen, en Peter Heuts. Tijdens de looptijd van het project waren er diverse problemen te overwinnen waarvoor, dankzij jullie inbreng, altijd een goede oplossing werd gevonden. Het was me een waar genoegen.

Heel veel dank gaat uit naar al die mensen die me hielpen om de data te verzamelen en zorgvuldig in te voeren: Marjon Dudzik, Wiel van 't Hooft, Anny Raven, Han Zijlema, Hugo Voncken en Vera van Dam, zonder jullie was het zeker niet gelukt, bedankt!

Ruud Houben, Aukje Soeters, Iris Engelhard en Madelon Peters, bedankt voor jullie hulp en enthousiasme bij het doen van de faalexperimenten.

Ik had 2 werkplekken en kon er maar moeilijk afscheid van nemen getuige de uitgestelde verhuizing. Collega's van iRv en DMKEP: bedankt voor een te gekke tijd!!

Twee werkplekken en twee kamergenoten: Roelof en Ruud bedankt voor de gedeelde werkweek die in jullie gezelschap altijd goed was door te komen. Dank je wel Jolanda voor het plezier, het luisterend oor, en de schaatspret. We promoveren bijna gelijk: heel veel succes! Bedankt Mariëlle voor de vriendschap en je deskundige blik. Caroline, jij werd al snel een fijne vriendin, bedankt voor de steun, de preuf, en de leuke fietstochten naar het iRv. En verder bedank ik in het bijzonder Yvonne, Janine, Olga, Erik, Henk, Marjon, Jeroen, Camiel, Geert, dames van Synchron, Marcel, en Luc voor de fijne collegialiteit.

Bij DMKEP wil ik in het bijzonder alle AIOs en ex-AIOs bedanken met wie ik leuke tijden heb beleefd op het werk, maar vooral ook tijdens de FRAIO-meetings en de EPP-uitjes. Marie-Anne en Anoeke, bedankt voor de fijne tijd en jullie aanmoediging. Sabine, mijn trouwe reismaatje naar de pijncongressen, ik heb ervan genoten. Wanneer gaan we weer? Bert en Theo, bedankt voor jullie technische hulp bij de faalexperimenten. En verder alle AIOs die de laatste tijd zo met me hebben meegeleefd: Mattijn, Ruud, Jeffrey, Anneke, Josefiën, Iris, Simkje, Guus, Marisol, Rosanne, Jorg, en Nicole, bedankt!

Met veel plezier kijk ik terug op de reisjes naar Barcelona, Kopenhagen, Manchester, Wenen, Hull, Gent, en in den lande. Alle reisgenoten die me daarbij vergezelden wil ik bedanken voor de ontzettend leuke tijd.

Wie zegt dat lotgenoten-contact niet werkt? Inez, Ellen, Dianne, Helga, en Manon, bedankt voor de gedeelde deskundigheid en smart waar het gaat om de evaluatie van therapieën.

Jacky, jij wist vast ook niet waaraan je begon toen je zei dat je mijn boekje wel wilde lay-outen. Maar het is een pracht exemplaar geworden en ik kan je niet

genoeg bedanken voor al het werk dat je daar in hebt zitten. Paula, bedankt voor de mooie verpakking die je aan mijn boekje meegaf. Is het toevallig dat je oranje concept op 2/2/2002 in mijn bus viel? Pap bedankt voor de mooie tekeningen van Rodin's denker.

Ingrid, een paar jaar deelden we lief en leed op weg naar Hoensbroek en terug. Bedankt voor al die kostbare en vertrouwde momenten.

Ellen, altijd al mijn paranimf. Wie anders zou ik willen vragen op mijn promotiedag dan jij. Je bent een vriendin, collega, maatje, steun en toeverlaat (zelfs in een 747 met een gat in de vleugel). Met jou in een tenue de ville komt het straks op 19 april helemaal goed.

Ellen, lieve 'hofdame', Heidi, Astrid, Jacqueline, Karien, Yolanda, en alle andere vrienden en vriendinnen die steeds maar weer vroegen wanneer we een feestje hadden. Bedankt voor die aanmoediging: het is zover!

Lieve balletvriendinnen, geacht vrouwenennetwerk, nog nooit begon het weekend zo goed als op al die héérlijke vrijdagavonden in jullie gezelschap. Jullie vriendschap is me veel waard, dank jullie wel. En nu dan eindelijk: Champagne!!

Romeo, trouwe kater, dank je wel voor het op me wachten boven aan de trap.

John, enig broertje, bedankt voor jouw steun, humor, huis-tuin-en-keuken psychologie, lieve zorg en vriendschap. Ik ben apetrots op een broer als jij en kan me geen betere paranimf toewensen.

Lieve pap en mam, bij al wat ik doe weet ik me door jullie twee gesteund. Jullie betrokkenheid heeft me altijd geholpen om te bereiken wat ik wil. Bedankt voor alles!

Anja

ABOUT THE AUTHOR

Anja van den Hout was born March 21, 1968 in Gemert, the Netherlands. In 1986 she obtained her VWO diploma (highschool) at the 'Macropedius College' in Gemert, and started her study at the Faculty of Health Sciences of the University of Maastricht. She specialized in Health Education and Mental Health Science, and graduated in 1992. From 1993 to 1995 she worked as a prevention functionary at the Regional Ambulatory Center for Mental Health Care (RIAGG), and at the Department of Medical Psychology, University of Maastricht, as a research assistant and assistant professor. In December of 1995, she started as a Ph.D. student at the Department of Medical, Clinical, and Experimental Psychology of the University of Maastricht, and was detached to the Institute for Rehabilitation Research in Hoensbroek. Next to her Ph.D. study she started a post-university education Behavior Therapy, and carried out cognitive-behavior therapies at the RIAGG in Maastricht. Furthermore, Anja worked as a tour-guide of cycling holidays in Europe during several summer holidays. Since November 2001 she is working as a post-doctorate researcher at the Centre for Nursing Research of the University of Maastricht, studying pain in children, and as a psychologist at the Department of Pain Rehabilitation and Rheumatology of the Hoensbroeck Rehabilitation Center. In 2002 the author will officially register as a behavior therapist (VGt), and she hopes to complete her official registration as a Health Care Psychologist in the near future.

PUBLICATIONS

Van den Hout, J.H.C., Vlaeyen, J.W.S., Heuts, P.H.T.G., Sillen, W.J.T., Willen, J.E.H.L. (2001). Functional disability in non-specific low back pain: the role of pain-related fear and problem-solving skills. *International Journal of Behavioral Medicine* 8(2), 134-148.

Van den Hout, J.H.C., Vlaeyen, J.W.S., Houben, R.M.A., Soeters, A.P.M., Peters, M.L. (2001). The effects of failure feedback and pain-related fear on pain report, pain tolerance, and pain avoidance in chronic low back pain patients. *Pain* 92, 247-257.

Van den Hout, J.H.C., Vlaeyen, J.W.S., Peters, M.L., Engelhard, I.M., Van den Hout, M.A. (2000). Does failure hurt? The effects of failure feedback on pain report, pain tolerance, and pain avoidance. *European Journal of Pain*, 4(4), pp. 335-346.

Van den Hout, J.H.C., & Kole-Snijders, A.M.J. (2000). Cognitieve behandelingen. In J.W.S. Vlaeyen & P.H.T.G. Heuts (Eds.), *Gedragsgeoriënteerde behandelingsstrategieën bij rugpijn* (pp. 83-103). Houten/Diegem: Bohn Stafleu Van Loghum.

Van den Hout, J.H.C., Vlaeyen, J.W.S., Kole-Snijders, A.M.J., Heuts, P.H.T.G., Willen, J.E.H.L., & Sillen, W.J.T. (1998). Graded activity and problem solving therapy in sub-acute non-specific low back pain. *Physiotherapy*, 84, 167.

Van den Hout, J.H.C., Arntz, A., Kunkels, F.H.J. (1995). Efficacy of a self-control therapy program in a psychiatric day-treatment center. *Acta Psychiatrica Scandinavica*, 92, 25-29.

Van den Hout, A., Kole-Snijders, A., Geurts, S. & Vlaeyen, J. (2001). Probleemoplossings Vaardigheidstraining bij a-specifieke lage rugpijn. *Gedrag & Gezondheid*, 29, 269-278.

Van den Hout, J.H.C., Vlaeyen, J.W.S. (in press). Problem-solving therapy and behavioral graded activity in the prevention of chronic pain disability. In S.J. Linton (Ed.), *New avenues for the prevention of chronic musculoskeletal pain and disability*.

Van den Hout J.H.C., Vlaeyen J.W.S., Heuts P.H.T.G., Zijlema J.H.L., Wijnen J.A.G. (accepted pending revision). Secondary prevention of work-related disability in non-specific low back pain: Does problem-solving therapy help? A randomized clinical trial. *Clinical Journal of Pain*.