

Pain judgements in other people: investigating underestimation of pain

Schmerzbeurteilungen: Untersuchungen zur Unterschätzung von Schmerzen

Dissertation

der Fakultät für Informations- und Kognitionswissenschaften

der Eberhard-Karls-Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Dipl. - Psych. Judith Kappesser

aus Mainz

Tübingen

2004

Tag der mündlichen Qualifikation:	19.01.2005
Dekan:	Prof. Dr. Michael Diehl
1. Berichterstatter:	Prof. Dr. Martin Hautzinger
2. Berichterstatter:	Amanda C. de C. Williams, PhD (University College London)

Contents

Abstract	1
German abstract	3
Introduction	
1. Communication of pain	6
1.1 When is pain communication successful? The sociocommunications model	7
1.2 How do observers decode the pain expression? The lens model	9
2. Importance of pain behaviour cues: Which cues are important and how do they affect judgements?	11
2.1 How do pain behaviour cues affect pain intensity judgements?	12
2.1.1 Verbal pain behaviour	12
2.1.2 Non-verbal pain behaviour	12
2.1.2.1 Facial expression of pain	13
2.1.3 Relationship between verbal and non-verbal pain behaviours ..	13
3. Importance of pain behaviour cues: How do observers perceive cues? ...	14
4. Agreement: How well do judges agree with patients' self-reported pain intensity ratings?	15
4.1 Direct comparisons in clinical settings	16
4.1.1 Criteria for evaluation of studies	17
4.1.2 Description of the selected studies and their results	23
4.1.2.1 Definition of agreement	24
4.1.2.2 Judges' relation to patients and patients' diagnoses	25
4.1.2.3 Available cues	26
4.1.3 Summary	27
4.2 Vignette studies	27
5. Explanations for underestimation	29
5.1 Theories proposed to explain underestimation	29
5.1.1 The hypothesis of developed insensitivity	29
5.1.2 Representativeness heuristic	30
5.2 Single variables suggested to explain underestimation	31

5.2.1 Base rate of deception	31
5.2.2 Supporting medical evidence	32
5.2.3 Relationship between patients and judges	33
5.2.4 Time spent with patients	34
5.3 Summary and conclusion	34
5.4 Alternative possibilities to explain underestimation	34
5.4.1 Missing cue	35
5.4.2 Evolutionary psychology	35
5.4.2.1 Social exchange theory	36
5.4.2.2 The cheating detection mechanism and pain judgements	36
5.4.2.2.1 Alerting the cheating detection mechanism	37
6. Aims and objectives	38

Study 1

Which cues are used to judge pain?

Perceptions of health care professionals

Aims and objectives

1. The concept of importance: verbal report as missing cue	40
2. The concepts of cautiousness and manipulation: alerted cheating detection mechanism	41

Method

1. Design	41
2. Participants	42
3. Materials	43
3.1 List of cues	43
3.1.1 Choice of context cues	43
3.1.2 Presentation of cues	44
4. Procedure	44
5. Statistical analyses	45
5.1 Rank ordering of the cues	45
6. Feedback	46

Results

1. Description of the sample	46
1.1 Qualitative background variables	46
1.2 Quantitative background variables	47
2. Results regarding the importance of cues	47
2.1 Importance rankings	48
2.1.1 Kendall's coefficient of concordance for importance rankings ...	49
2.1.2 Differences between cues for importance rankings	49
2.1.3 Missing importance cues	50
2.2 Spatial similarities of cues: multidimensional scaling (MDS) proce- dures	50
2.2.1 Introduction	50
2.2.2 Multidimensional scaling of importance rankings	51
2.2.2.1 Number of dimensions	51
2.2.2.2 Stimulus space	52
2.2.2.3 Interpretation of the stimulus space	52
3. Results regarding the ease of manipulation of cues	53
3.1 Ease of manipulation rankings	53
3.1.1 Kendall's coefficient of concordance for ease of manipulation rankings	54
3.1.2 Differences between cues for ease of manipulation	54
3.2 Multidimensional scaling of ease of manipulation rankings	55
3.2.1 Number of dimensions	55
3.2.2 Stimulus space	55
3.2.3 Interpretation of the stimulus space	55
4. Results regarding cautiousness	56
4.1 Cautiousness rankings	56
4.1.1 Kendall's coefficient of concordance for cautiousness rankings	56
4.1.2 Differences between cues for cautiousness rankings	57
4.1.3 Missing cautiousness cues	57
4.2 Multidimensional scaling of cautiousness rankings	58
4.2.1 Number of dimensions	58
4.2.2 Stimulus space	58
4.2.3 Interpretation of the stimulus space	59

Discussion

1. Discussion of results	60
2. Criticism	62
3. Outlook	63

Study 2

Examining two accounts for underestimation of pain

Aims and objectives

1. Verbal report as missing cue	64
2. Presence of a context cue making judges suspicious	64

Method

1. Design	65
2. Participants	65
3. Materials	66
3.1 Facial expressions of pain	66
3.1.1 Laptop	67
3.2 Scales to assess the dependent variables	68
3.2.1 Scale for pain judgements	68
3.2.2 Scale for confidence ratings	68
3.3 Background variables	68
4. Procedure	69
5. Statistical analyses	70
6. Feedback	70

Results

1. Description of the sample	71
1.1 Effect of the faking instruction (group 3)	71
2. Results of comparisons between groups	72
2.1 Descriptive statistics	72
2.2 Results of the ANOVA	72
2.3 Results of the Kruskal-Wallis test	73
3. The chain graph	74

3.1 Construction principles of chain graphs	74
3.2 Order of variables	75
3.3 Test for non-linearities	75
3.4 Summary of the final model	76
3.5 Association between difference in pain intensity and confidence ratings	78
Discussion	
1. Discussion of results	80
1.1 Accounts of underestimation	80
1.2 Effect of additional variables on underestimation	80
1.3 Association between agreement and confidence	82
2. Criticism	82
3. Outlook	84

Study 3

Which cues affect pain judgements of patients' relatives?

Aims and objectives

1. Choice of cues	85
-------------------------	----

Method

1. Design	87
1.1 Independent variables	87
1.2 The use of vignettes	88
1.3 Background variables	88
2. Participants	89
3. Materials	90
3.1 Vignettes	91
3.2 Scales to assess the dependent variables	92
3.2.1 Scale for pain intensity ratings	92
3.2.2 Scale for fairness ratings	92
4. Procedure	92
5. Statistical analyses	93

6. Feedback	93
-------------------	----

Results

1. Description of the sample	94
2. Descriptive statistics of quantitative and qualitative variables	94
3. Analyses of variance	96
3.1 Fairness of behaviour as dependent variable	96
3.2 Pain intensity ratings as dependent variable	97
3.2.1 Main effects	97
3.2.2 Interactions	97
3.3 Differences between reported and rated pain intensities	98
3.4 Examination of covariation	99
3.4.1 Impact of one dependent variable on the other	99
3.4.2 Impact of background variables	100

Discussion

1. Discussion of results	101
1.1 Pleasant and unpleasant tasks	101
1.2 Results of medical investigations	102
1.3 Reported pain intensity	102
1.4 Background variables	103
2. Criticism	104
3. Outlook	106

Discussion

1. Discussion of results	107
1.1 Accounts for pain underestimation	107
1.2 Impact of selected cues on pain judgements of patients' relatives	109
2. Criticism	110
2.1 Studies on agreement between patients and judges in the clinical context	110
2.2 Neglect of other accounts for underestimation	110
3. Outlook	112

References

114

Appendices

Appendix I: Introduction

I.1 Excluded studies in the clinical context	126
I.2 Scales measuring pain intensity	141
1.1 Verbal Rating Scales (VRS)	141
1.1 Visual Analogue Scales (VAS)	142
1.1 Numerical Rating Scales (NRS)	142
I.3 Bibliography for I.1 and I.2	144

Appendix II: Study 1

II.1 Correspondence with Ethics Committee: initial letter and reply	150
II.2 Poster asking for volunteers	152
II.3 Questionnaire	153
II.4 Feedback sheet	156
II.5 Post-hoc analysis of differences between cues regarding their impor- tance: results of Wilcoxon signed rank test	157
II.6 Summary of missing importance cues as provided by participants	158
II.7 Post-hoc analysis of differences between cues regarding their ease of manipulation: results of Wilcoxon signed rank test	159
II.8 Post-hoc analysis of differences between cues regarding their cautious- ness: results of Wilcoxon signed rank test	160
II.9: Summary of missing cautiousness-inducing cues as provided by partici- pants	161

Appendix III: Study 2

III.1 Letter and poster by which volunteers were contacted	162
1.1 Example of letters to contact oncology doctors	162
1.2 Example of poster to contact A&E doctors and nurses	163
III.2 Questionnaires	164
2.1 Questionnaire for condition 1: facial expression only	164
2.2 Questionnaire for condition 2: facial expression and verbal report	167

2.3 Questionnaire for condition 3: facial expression, verbal report and opioids	170
III.3 Feedback sheet	173
III.4 Correlations between all variables enclosed in the chain graph	174
III.5 Regression equations for variables excluded from the chain graph	175
III.6 Correlations between confidence and all explanatory variables	177
III.7 Regression equations for variables found not to be associated with con- fidence.....	178
 Appendix IV: Study 3	
IV.1 Ethics Committee application form including consent form	179
IV.2 Succeeding correspondence	186
IV.3 Letter to chronic pain patients	188
IV.4 Questionnaire	190
IV.5 Feedback sheet	198

Acknowledgements

I am deeply indebted to many people without whom I could and would have neither begun nor completed this PhD.

In particular, I would like to thank Amanda Williams. By accepting me for a placement when at Bath University and by introducing me to research on facial expressions of pain and evolutionary psychology it was she who made me think about doing a PhD at all. I would like to thank her for her on-going encouragement and amazing support throughout the whole time of my PhD. Her enthusiasm and help accompanied me during every stage of the project.

I am very grateful to Martin Hautzinger for facilitating this PhD, for his confidence in my capability of completing the PhD and for his assistance throughout the entire PhD.

Without the generous funding of the German National Merit Foundation I could not have realised this project and could have attended neither the 10th World Congress on Pain of the International Association of Pain in San Diego 2002 nor the Annual Scientific Meeting of the Pain Society in Glasgow 2003. The continuing support and interest of Ursel Theile, my German National Merit Foundation 'Vertrauensdozentin', was a great encouragement.

I wish to thank Ken Prkachin and Susan Mercer for so kindly providing the videos used in the second study as well as Ray Tait who, by kindly providing the original questionnaire he and his colleagues had used, facilitated the development of my own questionnaire for relatives. Further, I am grateful for discussion of and comments on ideas to Larry Fiddick, Ken Craig and Ken Prkachin and for discussion of methodology and assistance with statistical analyses to Reinhold Streit, Stefan Schmukle, Nanny Wermuth, Tony Coxtton and Matthew Hankins.

Special thanks go to the staff at INPUT Pain Management Unit, St Thomas' Hospital, especially to Gill Hancock and Jannie van der Merwe, for their support with data collection, their helpful comments and suggestions and their interest.

I am deeply indebted to the relatives of people in pain treated at INPUT Pain Management Unit, St Thomas' Hospital London as well as to staff at Guy's and St Thomas' Hospital who volunteered to take part in my studies.

And finally I wish to express my deep gratitude to my husband and parents for their invaluable support. They gave me space without which I could not have begun this project and gave me time by innumerable hours of child care which allowed me to complete this project.

Abstract

When people have to rate sufferers' pain intensity, they are confronted with the challenge of extracting from a diversity of cues those relevant to establish the presence and intensity of pain. In the literature a tendency is reported for health care professionals to underestimate pain and for relatives to overestimate pain. The aim of the present piece of work was to

- review studies investigating agreement between patients and judges (introduction)
- investigate how health care professionals perceive cues with regard to importance, manipulation and cautiousness (study 1)
- examine whether two explanations, one derived from a judgement and decision making model, the other one from an evolutionary psychology theory, can account for pain underestimation (study 2) and to
- investigate how selected cues impact on relatives' pain judgements (study 3).

When reviewing agreement studies, only 18 % of studies were methodologically sound according to the criteria set prior to searching for them. There appeared to be differences between studies depending on judges' relationship to patients and on patients' diagnoses. Furthermore, many studies were not explicit about which cues were available to judges, while others forbade judges to talk to patients.

Study 1 found that health care professionals perceive pain behaviours as more important and more cautiousness-inducing than contextual cues for pain judgements. However, importance of cues was closely related to cues' ease of manipulation.

Study 2 revealed that two accounts of pain underestimation, absence of an important cue (verbal report) and presence of a contextual cue (opioid abuse), led to a greater extent of underestimation than availability of verbal report and facial expression. In addition, expectations of pain exaggeration in patients affected agreement between judges and patients.

Study 3 showed that relatives (like health care professionals) were affected in their pain intensity ratings by verbal report. Although the effect was considerable, rela-

tives, were not affected by a contextual cue (medical evidence) which usually affects pain intensity ratings of professionals. Two additional variables concerning pain patients' behaviours (continuation of pleasant and unpleasant tasks) affected ratings of fairness of these behaviours and of pain intensity.

Taken together, results of the present piece of work largely support the presented hypotheses. However, they also suggest the need to improve the quality of studies investigating agreement, to establish a standard method for analysing agreement studies and to use the term underestimation of pain more selectively. Although more research is needed to further explore the appropriateness and usefulness of the lens model and social contract theory for pain judgement investigation, the results presented here seem at least promising. Both help to shift the research focus from evaluating pain judgements to explaining them. And only successful explanation of pain underestimation offers the chance to set up strategies capable of encountering its consequences.

Keywords: pain underestimation, judgement and decision making, evolutionary psychology

German abstract
Deutsche Zusammenfassung

Immer wenn jemand vor der Aufgabe steht, das Vorhandensein und die Intensität von Schmerzen eines Patienten zu beurteilen, muss er aus einer potentiell sehr großen Menge von Hinweisen (,cues') diejenigen auswählen, die für seine Aufgabe relevant sind. In der Literatur wird eine Tendenz beschrieben, dass Menschen, die beruflich Schmerzen beurteilen müssen (Ärzte, Krankenpfleger, Physio- und Ergotherapeuten), die Schmerzintensitäten der Patienten unterschätzen, während Angehörige von Patienten dazu neigen, die Intensität von Schmerzen des Patienten zu überschätzen. Ziel der vorliegenden Arbeit war es

- experimentelle Studien zusammenzustellen, in denen Übereinstimmung zwischen Beurteilern und Patienten untersucht wird, und ihre Ergebnisse zusammenfassend zu bewerten (Einleitung)
- zu untersuchen, wie Ärzte, Krankenpfleger und –schwestern sowie Physiotherapeuten verschiedene Hinweise bzgl. ihrer Wichtigkeit und ihrer Manipulierbarkeit für den Beurteilungsprozess sowie bzgl. des Auslösens von Misstrauen während des Beurteilungsprozess bewerten (1. Studie)
- zu untersuchen, ob mindestens eine von zwei vorgestellten Erklärungsmöglichkeiten (wobei eine Möglichkeit sich aus einem Modell der Urteilsforschung, die andere sich aus einer evolutionspsychologischen Theorie ableitet) die Neigung von Ärzten, Krankenschwestern und –pflegern, Schmerzen zu unterschätzen, vorhersagen kann (2. Studie) und
- zu untersuchen, wie sich vier Hinweise und wie sie sich auf die Schmerzbeurteilung bei Angehörigen von chronischen Schmerzpatienten auswirken (3. Studie).

Die Qualität der Studien, in denen Selbst- und Fremdeinschätzungen von Schmerzen verglichen wurden, wurden an Hand von Kriterien bewertet, die vor dem Suchprozess festgelegt wurden. Nur 18 % der Studien entsprachen diesen Kriterien. In den Ergebnissen dieser Studien gab es Unterschiede im Ausmaß der Unter- und Überschätzung von Schmerzen in Abhängigkeit von der Beziehung des Beurteilers zum Patienten und in Abhängigkeit von der Diagnose der beurteilten Patienten. Aus vielen der Studien ging weiterhin nicht eindeutig hervor, ob und welche Hinweise den Beurteilern zur Verfügung standen.

Die Ergebnisse der ersten Studie zeigen, dass Schmerzverhaltensweisen (verbales und non-verbales Verhalten) als wichtigere aber auch mehr Misstrauen auslösende Hinweise wahrgenommen wurden. Faktoren aus dem Kontext, in dem die Beurteilung stattfindet (z.B. ein möglicher sekundärer Krankheitsgewinn), wurden als weniger wichtig wahrgenommen. Weiterhin bestand ein enger Zusammenhang zwischen der Wichtigkeit von Hinweisen und ihrer Manipulierbarkeit.

Ergebnisse der zweiten Studie zeigen, dass die zwei vorgestellten Erklärungsmöglichkeiten, das Nichtvorhandensein eines wichtigen Hinweises (keine verbale Äußerung des Patienten zu seinen Schmerzen) und das Vorhandensein eines Hinweises auf sekundären Krankheitsgewinn (Schmerzmittelabhängigkeit), zu einer stärkeren Neigung führten, Schmerzintensitäten zu unterschätzen, als dies bei der Gruppe der Fall war, bei der der wichtige Hinweis vorhanden und der Hinweis bzgl. des sekundären Krankheitsgewinn nicht vorhanden war. Zusätzlich beeinflussten Annahmen der Beurteiler darüber, wie viele Patienten ihre Schmerzen übertrieben darstellen, das Ausmaß an Unterschätzung.

Die dritte Studie ergab, dass Angehörige in ihrer Schmerzbeurteilung von der Schmerzbeurteilung des Patienten beeinflusst werden. Zwei Variablen, die das Verhalten des Patienten betreffen (ob er Tätigkeiten, die er gerne tut, weiter ausübt und ob er Tätigkeiten, die er nicht gerne tut, weiter ausübt), beeinflussten nicht nur die Schmerzintensitätsbeurteilung, sondern auch das Urteil darüber, wie fair das Verhalten des Patienten ist. Die Ergebnisse von medizinischen Untersuchungen dagegen hatten keine Auswirkungen auf die Beurteilung der Schmerzintensitäten.

Zusammengefasst unterstützen die Ergebnisse der Studien im Großen und Ganzen die aufgestellten Hypothesen. Sie legen allerdings auch nahe, dass die methodische Qualität der Studien verbessert werden muss, dass ein Standard bzgl. der Auswertung der betreffenden Studien etabliert werden sollte und dass der Begriff ‚Unterschätzung‘ differenzierter gebraucht werden sollte. Obwohl noch weitere Untersuchungen notwendig sind, sind die zwei hier untersuchten Erklärungsmodelle viel versprechend. Beide sind hilfreich, den Forschungsschwerpunkt vom Bewerten der Schmerzbeurteilung auf ihre Erklärung zu verlagern. Und nur die erfolgreiche Erklärung von Beurteilungen eröffnet die Möglichkeit, Strategien zu erarbeiten und zu imp-

lementieren, die die negativen Auswirkungen des Beurteilungsprozesses auffangen können.

Schlagworte: Unterschätzung von Schmerz, Urteilsbildung, Evolutionspsychologie

Introduction

1. Communication of pain

The experience of pain is commonly defined as private, highly personal and subjective (Merskey, Albe-Fessard, Bonica, Carmen, Dubner, Kerr, Lindblom, Mumford, Nathan, Noordenbos, Pagni, Renaer, Sternbach & Sunderland, 1979). This definition implies that despite its salience for the sufferer, the presence and intensity of pain are difficult to judge for other people. From sufferers' point of view pain is an experience that demands attention (Eccleston & Crombez, 1999) and prioritises escape, recovery and healing (Wall, 1999). From onlookers' viewpoints pain signals suffering and, thereby, enlists aid (Keefe & Dunsmore, 1992; Prkachin, 1992; Prkachin, Berzins & Mercer, 1994) as shown, for example, by physiological arousal and sympathetic non-verbal expression occurring in observers looking at others in pain (Vaughan & Lanzetta, 1980). In order to receive the help of onlookers, however, sufferers have to meet a requirement: It is necessary that others recognise their pain. Therefore, successful communication between sufferer and onlooker is of utmost importance (Williams, 2002).

But how can pain be communicated? Sufferers can, intentionally or not, communicate pain to their environments by a variety of observable behaviours, for which Fordyce (1976) established the term 'pain behaviours'. As pain behaviours make the private experience of pain public, onlookers can use them as signals for pain. However, pain behaviours occur among a potentially huge diversity of cues (pain-related ones as well as others) which is why observers of pain are confronted with the challenge to extract from this diversity the cues that can help them establish the presence of pain and its severity (Craig, Prkachin & Grunau, 2001; Hadjistavropoulos & Craig, 2002).

The variety of pain behaviours is usually classified into verbal and non-verbal behaviours (for instance Craig et al., 2001; Hadjistavropoulos, von Baeyer & Craig, 2001; Turk & Melzack, 2001). Verbal behaviours include pain self-reports such as complaining, describing pain, exclaiming or making a pain rating. Non-verbal pain behaviours range from paralinguistic vocalisations (such as crying or moaning), other non-verbal qualities of speech (such as timbre, volume or hesitations), physiological ac-

tivity (such as flushing, sweating or muscle tension) and bodily activity (including involuntary reflexes as well as purposeful actions) to facial expressions.

Despite the wide use of pain behaviour as a unitary concept in behavioural therapy and research, different behaviours may well serve different functions and their communicative value may be secondary to other functions (Prkachin, 1986, 2002). Physiological activity, for example, may be seen as an automatic response to the physical threat. Some of the bodily activities, such as reflexive withdrawal, might be a reaction to the physical threat, whereas others, such as rubbing a painful area, may mainly serve to cope with pain (Wilkie, Keefe, Dodd & Copp, 1992). Speech and facial expression can control pain only indirectly and function above all as social communications (Craig et al., 2001). Results of a study by Hadjistavropoulos and colleagues (Hadjistavropoulos, Craig, Grunau & Johnston, 1994), for instance, indicate that paravocalisations serve to get attention from people nearby, whereas the face serves as the source to communicate distress. Similarly, people in pain displayed facial expressions and vocalisations longer when an observer was present in the experimental situation (Sullivan, Adams & Sullivan, 2004). While the non-verbal behaviours occur immediately after the experience of pain, verbal report of pain comes into play comparatively late (Craig & Prkachin, 1983).

Hence, although pain is defined as a private, personal and subjective experience, pain can be made public by verbal and non-verbal pain behaviours. However, a successful communication is required whenever the sufferer needs the help of onlookers to escape and recover from pain. But how can pain be communicated successfully? A model that may be helpful answering this question is the 'sociocommunications model' of pain by Craig (2002).

1.1 When is pain communication successful? The sociocommunications model

The 'socio-communications model' of pain has recently been put together by Kenneth Craig (2002). The model is based on Rosenthal's conceptualisation of non-verbal communication (Rosenthal, 1982) and its adaptation to pain by Prkachin, Craig and Hadjistavropoulos (Hadjistavropoulos & Craig, 2002; Prkachin & Craig, 1995). Although it was originally used to explain pain in infants and children, it is my opinion that it can be transferred to all other pain sufferers.

As illustrated in Figure 1, the model directs attention to the dynamism and complexity of the information transmission process between sufferers and observers. Sufferers can encode their experience of pain (first box) in expressive behaviours (pain behaviours; second box). Expressions of pain in turn can enable the observer to draw inferences by decoding the messages (third box). Last but not least, the judge may make a decision about necessary actions (fourth box).

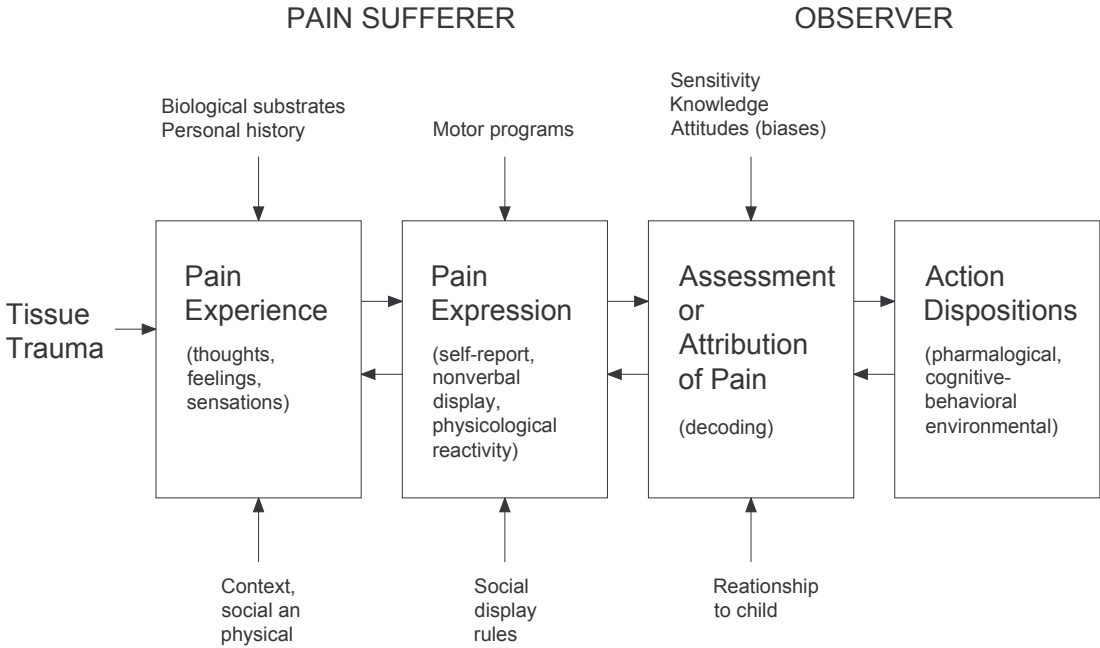


Figure 1. The socio-communications model by Craig (2002).

In terms of this model effective communication on part of the sufferer requires successful encoding of pain experience into pain expressions. On the part of the observer effective communication would imply successful decoding of pain. This sounds very simple, yet the model also suggests how complex this act of communication is and how easily it can be interrupted: The manner in which sufferers construct the significance of an injury, disease or physical threat, for instance, powerfully determines the overall nature of the pain experience. Or as Patrick Wall states: ‘Pure pain is never detected as an isolated sensation. Pain is always accompanied by emotion and meaning.’ (Wall, 1999), p. 38). Further, the relation between pain experience and expression is intricate: Some patients may not want to let others know that they are in pain and, therefore, try not to express their pain as for instance found in studies by Madison and Wilkie (1995) or Jacox (1979, 1980). Observers may not attend to or

appreciate pain expressions or interpret them in the context of other information creating distorting biases.

The advantage of the sociocommunications model is that it offers a broad perspective necessary to understand the complexity of the information transmission between sufferer and onlooker. Because of its generality, however, the model is fairly unspecific with regard to the single components e.g. 'How do sufferers encode their pain experience in expressions of pain (pain behaviours)?' or 'How do observers draw inferences by decoding the pain expressions?'. Whereas there has been quite a lot of research on how pain behaviours can be quantified as well as on their relationship with the experience of pain, 'the manner in which this information is decoded and used by observers remains uncertain' (Craig, 1992, p. 159). Nevertheless, another model illustrating how people form judgements seems useful as a theoretical framework and as a structure for available results: the 'lens model'.

1.2 How do observers decode the pain expression? The lens model

The 'lens model' was developed by Egon Brunswik (1952) and fairly early applied to research questions in the clinical and medical field by one of his students (Hammond, 1955, 1964; Kelly, 1964a, 1964b). Today it is widely used in the judgement and decision making research field. The model, presented in Figure 2, illustrates how people form judgements of an intangible criterion on the basis of a set of imperfect but tangible cues. It assumes that judgements are cognitive processes similar to inductive inferences. By making judgements (Y_s) judges draw inferences about a criterion that cannot be seen on basis of data (X_{1-i}) that can be seen (Connolly, Arkes & Hammond, 2000). Applied to the question of how observers decode pain expressions the model assumes that the pain experience as intangible event can only be referred to by using a set of tangible cues, e.g. the pain behaviours, that help observers to form judgements about the presence and intensity of pain.

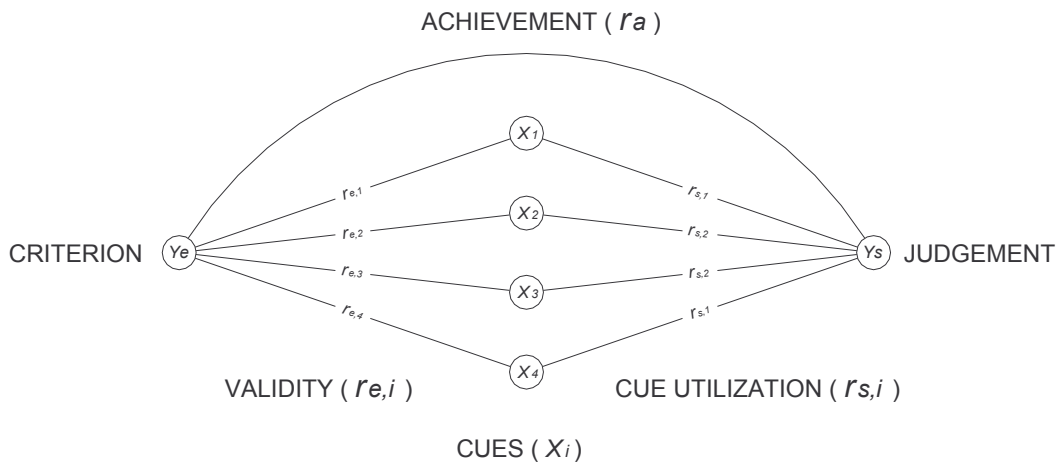


Figure 2. A schematic illustration of the lens model.

The lens model puts far more emphasis on the cues present in the judgement situation than the sociocommunications model. And rather than only introducing its three components (criterion, cues and judgement), the lens model also specifies their relations to each other. First, the fact that cues are of different values in making inferences about events is assessed by cue or ecological validity (r_e). The stronger the relationship between cue and event to be inferred, the greater the degree of validity. Second, the use of cues may differ depending on how important the judge considers the cue to be, a relation measured by cue utilisation (r_s). And third, achievement (r_a) indicates the agreement between judgements and criterion, i.e. the degree to which the judgement was correct.

According to the lens model there are several rather than one single question to ask. Instead of asking ‘How do observers decode pain expressions?’ the questions ‘How important are certain cues for judgements on pain experience?’, ‘How important do observers think certain cues are for judgements?’ and ‘How well do pain judgements agree with pain experience?’ seem more appropriate. Although the lens model has not been used explicitly in the pain field yet to investigate pain judgements, results of studies can and will be grouped according to these three questions in the following sections.

While the sociocommunications model is specific for communication of pain, the lens model can be applied to any situation in which a judgement is made. Consequently, there is no restriction (i.e. only pain expression) regarding the cues taken into account when applying the model to a certain situation.

2. Importance of pain behaviour cues: Which cues are important and how do they affect judgements?

The preponderance of studies examining the importance of cues explored the impact of characteristics of pain sufferers (e.g. age, sex, ethnicity, attractiveness) or judges (e.g. age, sex, profession, amount of working experience) on pain intensity judgements. The results of these studies are interesting and important. However, according to the sociocommunications model the cues of interest are the ones expressing pain i.e. the pain behaviours. Further, in her review of the literature Solomon (2001) concludes that 'many of the studies have serious methodological shortcomings including small samples, the adoptions of an atheoretical approach and use of survey instruments with poor or unknown psychometric properties' (p. 177). Therefore, I will only report results of studies investigating pain behaviours here and refer to sufferers' and judges' characteristics only briefly and where appropriate in the method sections of my studies.

Before reporting on the results, however, it is necessary to reflect briefly on the assessment of pain experience. The difficulty emerging is that most studies measure pain experience by asking sufferers about their pain. Or in the frequently cited words of Margo McCaffery (1980): 'Pain is what the experiencing person says it is, existing whenever he says it does.' (p. 26). This means, however, that self-report belongs to the group of pain expressions and is used to assess the pain experience. Or, using the language of the lens model, verbal report is the criterion as well as one of the cues.

It is important that verbal report can always only be an approximation for the pain experience, the biological substrates of which 'are a distributed system of multiple interacting sites associated with perception, affect, thought, language, motor control and expression' (Craig, 2002, p. 308). And although other measures of pain experience are available, such as measures of brain activity, they also have their limita-

tions. Brain scans, for example, cannot tell us as much about the presence, nature and intensity of pain as verbal report. In the context of pain research, verbal report is the most ecologically valid and, therefore, the most appropriate approximation available to assess the pain experience: If clinicians or relatives want to know whether sufferers are in pain and how intense the pain is, they will simply ask the patient. On the other hand, they will use sufferers' self-reports as one of the cues to draw an inference on the presence and intensity of pain. Therefore, although self-report is a pain behaviour cue, it is used and it is possible to use it as an approximation for the pain experience.

2.1 How do pain behaviour cues affect pain intensity judgements?

2.1.1 Verbal pain behaviour

Pain intensity as reported by the sufferer clearly has an impact on judges' pain ratings. A study by Todd and colleagues (Todd, Lee & Hoffman, 1994) found that patient pain assessment was a predictor for differences between patients' and judges' pain ratings. Further studies specified this relationship. For low reported pain intensities either no differences occurred between sufferers and judges (Chibnall & Tait, 1995) or judges rated the sufferers' pain higher than sufferers did (Chibnall, Tait & Ross, 1997; Krokosky & Reardon, 1989; Tait & Chibnall, 1997; Zalon, 1993). For medium reported pain intensities no systematic differences occurred between sufferers and judges (Chibnall et al., 1997) while for high reported pain intensities judges' ratings were lower than the sufferers' ones (Chibnall & Tait, 1995; Chibnall et al., 1997; Krokosky & Reardon, 1989; Tait & Chibnall, 1997; Zalon, 1993).

2.1.2 Non-verbal pain behaviour

In a study by von Baeyer and colleagues (von Baeyer, Johnson & McMillan, 1984) nursing students watched one of two videos showing a doctor-patient interaction. Exactly the same words were used by doctor and patient in both videos. However, the videos differed in how strongly the patient expressed her pain non-verbally. In one video she sat quietly, spoke in a normal tone of voice and maintained a neutral facial expression. In the other video, though, she was restless, frowned constantly, sighed heavily and spoke slowly. Nursing students who watched the doctor's interaction with the patient non-verbally expressing her pain evaluated the patient's pain to be more intense.

2.1.2.1 Facial expression of pain

Of all non-verbal behaviours, facial expressions seem to be especially important: a face in pain is the main focus of attention for observers and hard to ignore, probably because the other individual's distress could signal imminent personal threat (Craig et al., 2001; von Baeyer et al., 1984). Not only do observers consider movements in the face to be the most important determinants of their painfulness ratings (Prkachin, Currie & Craig, 1983), when using the Facial Action Coding System (FACS; Ekman & Friesen, 1978) research has established specific facial movements for pain, including lowering the brow, narrowing the eyes by tightening the lids and raising the cheeks or even fully closing the eyes, raising the upper lip, deepening the nasolabial fold and wrinkling the nose as well as opening the lips and mouth in varying degrees. This facial expression is consistent across a range of experimental pain modalities (Craig & Patrick, 1985; Galin & Thorn, 1993; LeResche, Dworkin, Wilson & Ehrlich, 1992; Patrick, Craig & Prkachin, 1986; Prkachin, 1992) and across different clinical pain conditions (Craig, Hyde & Patrick, 1991; Hadjistavropoulos & Craig, 1994; LeResche, 1982; LeResche & Dworkin, 1988; Prkachin & Mercer, 1989).

A study by Prkachin and colleagues (1994) found that observers' ratings of patients' pain were modestly correlated (coefficients ranged from - 0.09 to 0.75) with the overall degree (intensity and duration) of facial activity. Patrick and colleagues (1986) showed that judges are not only influenced by intensities of facial expressions but also by the frequency with which these expressions occur. The authors reported that in their study on average 55 % of variance in judges' pain ratings was attributable to changes in specific components of facial expressions of people in pain.

2.1.3 Relationship between verbal and non-verbal pain behaviours

Poole and Craig (1992) investigated the impact of facial expression and verbal report on pain judgements. They found that, although patient acknowledgement or denial of pain affected ratings based upon facial expressions, facial expressions were the more salient source of information: judges tended to discount the verbal report whenever it was discrepant with the facial expression. This finding is supported by a study reported by Jacox (1980) in which nurses were asked to describe situations in which it would be easy or difficult to assess pain. Their answers revealed that they regarded situations as difficult in which verbal report is discrepant with non-verbal

behaviour. Craig (1992) explains these findings by the fact that non-verbal behaviours seem less vulnerable to purposeful misrepresentation.

To summarise, studies show that verbal as well as non-verbal pain behaviour cues are important for pain judgements. However, the direction in which each affects judgements differs: whereas high reported pain intensity leads to lower pain ratings, higher non-verbal pain expressiveness leads to higher pain ratings in observers. Of all non-verbal pain behaviours the facial expression of pain seems to play an important role. The intensity and duration of as well as changes in specific facial components were shown to affect pain ratings. Studies that compared the impact of verbal and non-verbal pain behaviours found that non-verbal pain behaviours are more heavily weighted in pain judgements. One explanation for this finding is that non-verbal pain behaviours are regarded to be less vulnerable to misrepresentation.

3. Importance of pain behaviour cues: How do observers perceive cues?

Results of the study reported by Jacox (1980) revealed that the difficulty with which pain judgements are made varies with the kind of cues that are present in the judgement situation. Moreover, the results of a study by Ferrell and colleagues (Ferrell, Eberts, McCaffery & Grant, 1991) showed that the cues judges report to use most frequently are not the ones they consider as most important. When the authors asked nurses which cue they would use most often to assess pain, 91 % said they would ask the patient most frequently. However, only 45 % of them regarded this as most influential factor. But which cues are regarded as important by judges?

Several researchers let their participants rank order cues in order of their importance; unfortunately, the same cues were not included in all of these studies. When looking at studies including verbal as well as non-verbal pain behaviours, results are inconclusive: Whereas in the studies by Jacox (1980) and Oberst (1978) non-verbal pain behaviours were considered as more important than verbal ones, in the study by McKinley and Botti (1991) nurses regarded verbal communication as more important than non-verbal one.

4. Agreement: How well do judges agree with patients' self-reported pain intensity ratings?

The question of how well judges agree with patients has received particular attention in the pain literature since a considerable number of studies found that judges tend to underestimate patients' pain. Underestimation of pain was found to be associated with greater mood disturbance and less self-efficacy in patients (Cremeans-Smith, Stephens, Franks, Martire, Druley & Wojno, 2003; Miaskowski, Zimmer, Barrett, Dibble & Walhagen, 1997). It is further often referred to in association with another major clinical problem: inadequate pain relief. Studies revealed that a high proportion of patients experience pain for most or all of the time while in hospital (e.g. Lavies, Hart, Rounsefell & Runciman, 1992; Melzack, Abbott, Zackon, Mulder & Davis, 1987). In a study by Donovan and colleagues (Donovan, Dillon & McGuire, 1987), nearly 60 % of patients reported 'horrible' or excruciating' pain while being in hospital. Furthermore, the average amount of analgesics actually given to patients was less than 25 % of the amount originally prescribed, a finding confirmed by many other studies (e.g. Bartfield, Salluzo, Raccio-Robak, Funk & Verdile, 1997; Cleeland, Gonin, Baez, Loehrer & Pandya, 1997; Gauthier, Finley & McGrath, 1998; Marks & Sachar, 1973). Underestimation of pain is presumably associated with undermedication of pain since adequate pain assessment is a prerequisite for reliably adequate analgesic treatment (Lander, 1990a). However, underestimation is certainly not the only predictor for undermedication.

Studies designed to answer the question of how accurate judges are in their pain ratings compare the self-reported pain intensity of a sufferer with a pain intensity judgement made by another person. The difference between ratings is computed and judged by taking the self-report as the standard against which the judgement of the other person is evaluated. If both ratings match, the judgements are declared to agree. If there is a (significant) difference between both ratings, i.e. the judgement deviates from the self-report, the judgement by the other person is classified either as 'overestimation' (if the judgement is higher than the self-report) or as 'underestimation' (if the judgement is lower than the self-report).¹

¹ Especially the term 'underestimation' is frequently used in association with 'bias' in the pain literature. Since the term 'bias' in the judgement and decision making research field is regarded as very value-laden and as indicating systematic errors and lapses when reasoning, I would like to stress that I use the terms 'overestimation' and 'underestimation' throughout this piece of work without the suggestion of human irrationality.

It is important that the magnitude of the difference between ratings which is acceptable is a clinical rather than a statistical decision as 'we should ask whether the agreement is good enough for a particular purpose, not whether it conforms to some absolute, arbitrary criterion' (Bland & Altman, 1990, p. 339). Additionally, this difference should ideally be determined in advance. In many studies in the clinical context a judgement on a Visual Analogue Scale (VAS; a horizontal line of usually 10 cm length) that deviates up to 1 cm on either side from the patient's rating is defined as agreement. This criterion was established by lafrati (1986) and is based on clinical judgement since this criterion could mean the difference between interpretations of mild and moderate or moderate and severe pain (Miaskowski et al., 1997). This criterion is further supported by research investigating what patients regard as clinically significant changes in their pain. These studies found changes on a VAS that were regarded as clinically significant between 9 mm (Kelly, 1998) and 13 mm (Bird & Dickson, 2001; Gallagher, Liebman & Bijur, 2001; Todd, Funk, Fund & Bonacci, 1996).

Basically, two different approaches can be found investigating how well judges agree with patients: studies in the clinical context and vignette studies. The first type of study, which investigates the question in clinical contexts, compares the patients' self-reports of pain with pain judgements of health care professionals or patients' relatives. The second type of study uses vignettes in which fictional pain sufferers tell the reader about their painful condition. The reader is then asked to judge the pain from the information given. Whereas in vignette studies the judgement situation is fictional and access to information restricted, in studies in the clinical context the judgement situation is real and access to information less restricted.

4.1 Direct comparisons in clinical settings

Identifying studies in clinical settings proved to be far more difficult than looking for vignette studies. Reasons were that the key words in the articles, if given at all, were very unspecific and that for some studies the comparison was not the main focus of the paper and, hence, not referred to in title, key words and/or abstract. Further, several studies had been published in journals which were not referenced in either Med-Line or PsychLit. Nevertheless, 62 studies of very varying quality were found. Be-

cause of this diversity it seemed necessary to establish criteria by which the studies' relevance for the question of (dis)agreement between patients and observers could be evaluated and which would allow selection of relevant studies of good quality.

4.1.1 Criteria for evaluation of studies

The first selection criterion I chose was whether authors provided a definition of what they meant by agreement between patients' and judges' pain estimates. Only with such a definition the interpretation of results seems meaningful. Studies in which no definition was provided were excluded from the discussion of results later in this chapter. Further, for the purposes of this PhD it is necessary that not only the proportion of agreement but also the proportion of over- and underestimation was reported in the result section of the studies.

Two other criteria were methodological ones: since pain intensity can change rapidly, the time gap between ratings by patients and judges should not be too big. Consequently, only studies were included in which both ratings were made at least on the same day. Further, authors should report exactly what pain they assessed (e.g. average, current, worst, least pain). Therefore, studies without information which pain was assessed were excluded.

Last but not least, one of the main criticisms regarding the clinical studies was the reliance on a limited set of data analysis techniques (van der Does, 1989). However, the discussion about which techniques to use to obtain meaningful results is still ongoing. And although there are some suggestions, there is no standard yet. Nevertheless, the sole use of correlation coefficients and t-tests appears insufficient as both techniques suffer from flaws in determining the extent of agreement between two ratings: correlation coefficients could be perfect and highly significant, although the ratings between both groups differ consistently. T-tests are used to test for group differences. Yet, means of groups can be very similar despite considerable differences between single pairs of ratings (van der Does, 1989; Bland & Altman, 1986; 1990). Therefore, studies which only used correlation coefficients and t-tests to analyse data were excluded.

The mean absolute difference between the ratings of each patient-judge pair as well as the standard deviation between the paired observations and a plot of difference against the mean are considered valuable techniques in the context of agreement (van der Does, 1989; Bland & Altman, 1986; 1990). The mean absolute difference, however, does not allow to distinguish between over- and underestimation and, therefore, is useless when investigating pain underestimation.

To summarise, included in the following description of studies and their results are studies in which the authors defined agreement, reported agreement as well as over- and underestimation and used other data analysis techniques in addition to correlation coefficients and t-tests. The selection of studies according to these criteria led to the inclusion of eleven studies. A description of the excluded 51 studies can be found in Appendix I.1. As can be seen there, most of the studies were excluded because they did not meet several rather than only one of the selection criteria.

Table 1. Agreement between patients and observers in clinical contexts: summary of studies.

Authors	Sample	Research question	Pain intensity measure and time of measure	Relevant results	Comments
(1) Bowman, 1994	patients: abdominal hysterectomy, bilateral salpingectomy, gastroplasty, cholecystectomy, hemicolectomy, splenectomy and inguinal hernia repair (n = 16) <u>judges</u> : nurses assigned to patients (n = 13)	compare nurses' and patients' perception of patients' postoperative pain	<u>measure</u> : graphic rating scale (VAS with 5 points) <u>instruction</u> : current pain <u>time</u> : when patients complained of pain and requested pain medication; nurses after assessing the patient; data collected separately	<u>agreement</u> : 11 % <u>underest.</u> : 78 % <u>overest.</u> : 11 % <u>t-test</u> : nurses rated patients' pain sig. lower than patients (p=0.0002)	- agreement defined as ± 1 cm
(2) Brunelli, Costantini, Di Giulio, Gallucci, Fusco, Miccinesi, Paci, Peruselli, Morino, Piazza, Tamburini & Toscani, 1998	<u>patients</u> : terminal cancer patients (n = 147) <u>judges</u> : physician or nurse usually caring for patient	identify areas and specific symptoms for which health-care providers' ratings of patients' quality of life are accurate	<u>measure</u> : VRS (1-4; TIQ: Tamburini, Rosso, Gamba, Mencaglia, De Conno & Ventafredda, 1992) dichotomised as present (a lot, very much) and absent not at all, slight) <u>instruction</u> : past week <u>time</u> : at end of examination; data collected separately for patients and judges	<u>presence of pain</u> : 42.9 % <u>agreement</u> : 70.7 % <u>underest.</u> : 17.6 % <u>overest.</u> : 11.5 %	- agreement defined as matching categories
(3) Choiniere, Melzack, Girard, Rondeau & Paquin, 1990	<u>patients</u> : burn patients (n = 42) <u>judges</u> : nurses (n = 42) <u>judgement pairs</u> : 41 (during procedure), 42 (at rest)	assess accuracy with which nurses evaluate pain experience in burn patients	<u>measure</u> : VAS, VS (0-4) <u>instruction</u> : current pain <u>time</u> : twice (during painful procedure and rest) on one day every week until discharge; independent pain ratings at same time	<u>agreement</u> : 30 % <u>procedure</u> : 51 % <u>rest</u> : 51 % <u>underest.</u> : 43 % <u>procedure</u> : 33 % <u>rest</u> : 33 % <u>overest.</u> : 27 % <u>procedure</u> : 18 % <u>rest</u> : 18 %	- agreement defined as ± 1 cm on VAS

Authors	Sample	Research question	Pain intensity measure and time of measure	Relevant results	Comments
(3) continued				χ^2 -test: no sig. differences between number of correct and incorrect estimations <u>correlation (VAS): 0.47</u> <u>procedure: 0.47</u> (p<0.001) <u>rest: 0.33</u> (p<0.01)	
(4) Cremeans-Smith, Stephens, Franks, Martire, Druley & Wojno, 2003	<u>patients:</u> female osteoarthritis patients (n = 114) <u>judges:</u> spouses (n = 114), rheumatologists (n = 5)	examine dyadic agreement between patients and two role partners (spouse, rheumatologist)	<u>measure:</u> recent pain (during past month): VRS (1-5; AIMS2: Meenan, Mason, Anderson, Guccione & Kazis, 1992); typical pain (adapted from AIMS2): VRS (1-4) <u>instruction:</u> pain during past month, typical pain <u>time:</u> patient and spouse interviewed separately; physicians rated typical pain after appointment	<u>mean:</u> <u>recent pain:</u> 4.03 (p), 4.14 (s) <u>typical pain:</u> 2.76 (p), 2.40 (r) <u>agreement:</u> 55 % (s), 39 % (r) <u>underest.:</u> 18 % (s), 45 % (r) <u>overest.:</u> 27 % (s), 16 % (r)	– agreement defined as identical rating
(5) Everett, Patterson, Marvin, Montgomery, Ordonez & Campbell, 1994	<u>patients:</u> burn patients (n = 49) <u>judges:</u> nurses (n = 27) <u>judgement pairs:</u> 123	examine relations among nurse and patient pain ratings and investigate how different patient and nurse variables are related to pain reports and accuracy of nurses' estimates of pain	<u>measure:</u> VAS <u>instruction:</u> <u>patients:</u> worst, average pain during procedure <u>nurses:</u> overall pain <u>time:</u> shortly after wound care treatment (on average twice daily)	<u>agreement:</u> 53.7 %, <u>underest.:</u> 12.2 %, <u>overest.:</u> 34.1 % <u>mean abs. diff.:</u> 1.27 <u>correlation:</u> 0.70 (p<0.001)	– agreement defined as \pm 1 cm – nurses had additionally to rate pain-related behaviours during wound care which possibly increased sensitivity

Authors	Sample	Research question	Pain intensity measure and time of measure	Relevant results	Comments
(6) Iafrafi, 1986	patients: burn patients (n = 6) judges: nurses (n = 15) judgement pairs: 29	investigate correlation between pain that burn patient feels and burn nurse's perception of that pain	measure: VAS instruction: pain during wound cleansing/dressing change time: immediately after wound cleansing/dressing change; independently of each other	agreement: 31% underest.: 34.5 % overest.: 34.5 %	- agreement defined as ± 1 cm
(7) Madison & Wilkie, 1995	patients: pain due to lung cancer (n = 18) judges: family members (n = 18); judgement pairs: 18	compare patient and family member perceptions of pain	measure: VAS (once when videotaped while performing standardised behaviour, once when interviewed) instruction: current pain time: interview on same day for patient and family member; data collected separately	means: video: 24.6 (p), 37.3 (f) interview: 32.1 (p), 35.7 (f); agreement: video: 25 % interview: 18.75 % overest.: video: 62.5 % interview: 50 % underest.: video: 12.5 % interview: 31.25 % Wilcoxon: interview: ns video: p<0.06	- agreement defined as ± 1 cm
(8) Miaskowski, Zimmer, Barrett, Dibble & Walhaugen, 1997	patients: oncology outpatients (n = 78) judges: family caregivers (n = 78) judgement pairs: 78	examine whether differences in patients' and family caregivers' perception of pain experience are associated with negative outcomes for patients and caregivers	measure: VAS instruction: current pain time: when presenting in outpatient clinic; data collected separately	agreement: 29.5 % overest.: 52.9 %, aver. discrep.: 3.55 cm underest.: 17.6 %, aver. discrep.: 3.82 cm	- agreement defined as ± 1 cm

Authors	Sample	Research question	Pain intensity measure and time of measure	Relevant results	Comments
(9) Riemsma, Taal & Rasker, 2000	<u>patients</u> : rheumatoid arthritis (n = 188) <u>judges</u> : spouses of patients (n = 188)	examine differences in perceptions of patients and spouses about patients' health status and relation of differences to patients' and spouses' well-being	<u>measure</u> : 0-10 (converted from answers to single items) (Dutch version of AIMS2: Riemsma, Taal, Rasker, Houtman, van Paassen & Wiegman, 1996) <u>instruction</u> : during past month <u>time</u> : questionnaires were mailed to patients and spouses before patients attended education programme	<u>mean</u> : 5.32 (p), 6.02 (s) <u>agreement</u> : 20 % <u>underest.</u> : 23 % <u>overest.</u> : 57 % <u>t-test</u> : t=6.52 (p<0.001)	- agreement defined as identical ratings
(10) Rundshagen, Schnabel, Standl & Schulte am Esch, 1999	<u>patients</u> : undergoing major elective orthopaedic surgery (n=120) <u>judges</u> : nurses in post-anaesthesia care unit in charge of patients	investigate patients' and nurses' pain and anxiety assessments during postoperative pain treatment with 4 different analgesic regimens	<u>measure</u> : VAS <u>instruction</u> : current pain <u>time</u> : hourly ratings of pain during stay in post-anaesthesia care unit; nurses rated pain before patients	<u>agreement</u> : 48 % <u>underest.</u> : 51 % <u>overest.</u> : 1 % <u>correlation</u> : 0.69-0.89 for each of first 18 hours after surgery (all p's<0.001)	- agreement defined as ± 1 cm
(11) Zalon, 1993	<u>patients</u> : postoperative abdominal surgery patients (n = 119) <u>judges</u> : nurses responsible for the care of the patients (n =119) <u>number of paired judgments</u> : 119	compare nurses' assessments of pain with postoperative patients' actual pain ratings	<u>measure</u> : VAS <u>instruction</u> : when patient reported pain; nurses were asked to assess pain in their usual manner	<u>means</u> : 6.18 (p), 5.29 (n) <u>agreement</u> : 34.5 % <u>underest.</u> : 45.4 % <u>overest.</u> : 20.2 % <u>t-test</u> : nurses rated pain sig. lower than patients (p<0.05)	- agreement defined as ± 1cm

4.1.2 Description of the selected studies and their results

Table 1 summarises relevant information about the eleven selected studies regarding the

- sample with information on patients, judges and judgement pairs if reported by authors,
- main research question,
- pain intensity measures including instructions and time of measurement (in order not to expand this chapter unnecessarily, detailed description of the pain scales used are included in Appendix I.2)
- main results and
- my comments which refer mainly to the presence of an agreement definition and to whether the authors reported over- and underestimation.

What do results of these studies tell us about (dis)agreement? Figure 3 gives a graphical overview of proportions of agreement in all selected studies and shows that there is no general tendency to underestimate pain as it is often discussed in the literature. On the contrary, there is an enormous amount of variation between studies.

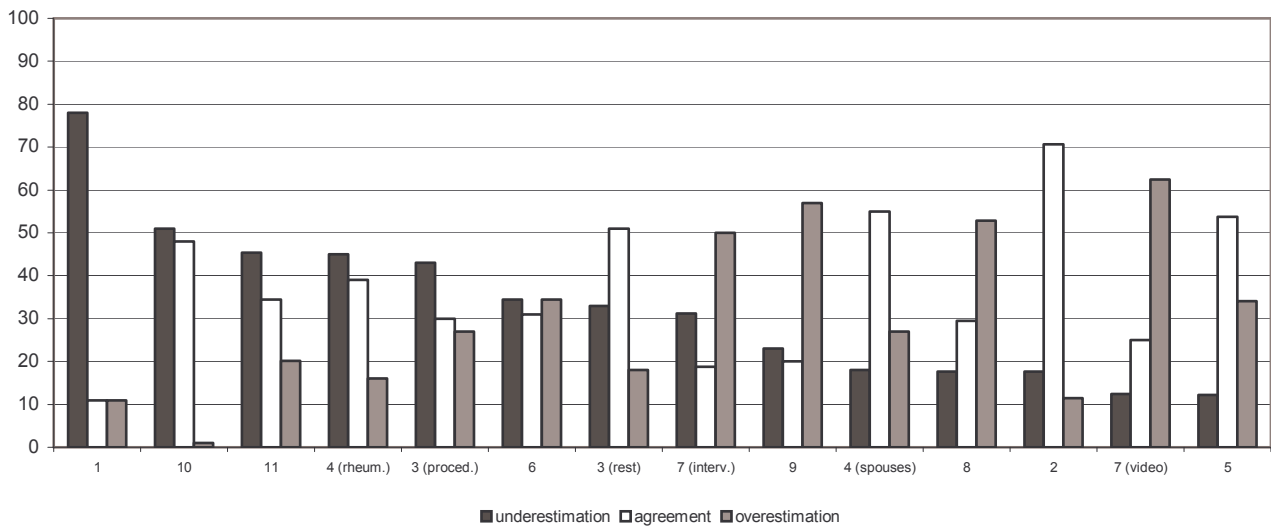


Figure 3. Proportions of agreement and disagreement obtained in selected studies.

The summary of studies was not only helpful to get an overview of proportions of (dis)agreement but also useful as it reveals new questions. Studies systematically vary with regard to judges' relationship to patients, to diagnoses of patients and to definition of agreement. Does any of these variables affect proportions of (dis)agreement? Further, which cues were available to judges in any of the selected studies? Each of these questions will be examined separately in the next sections.

4.1.2.1 Definition of agreement

Whereas the majority of studies (1, 3, 5, 6, 7, 8, 10, 11) defined agreement as within the range of ± 1 cm on a VAS scale, three studies (2, 4, 9) defined agreement as identical ratings of observers and judges on verbal rating scales (VRS) with four to five categories. Figure 4 shows the proportions of (dis)agreement depending on the definition of agreement applied and type of scale used by authors.

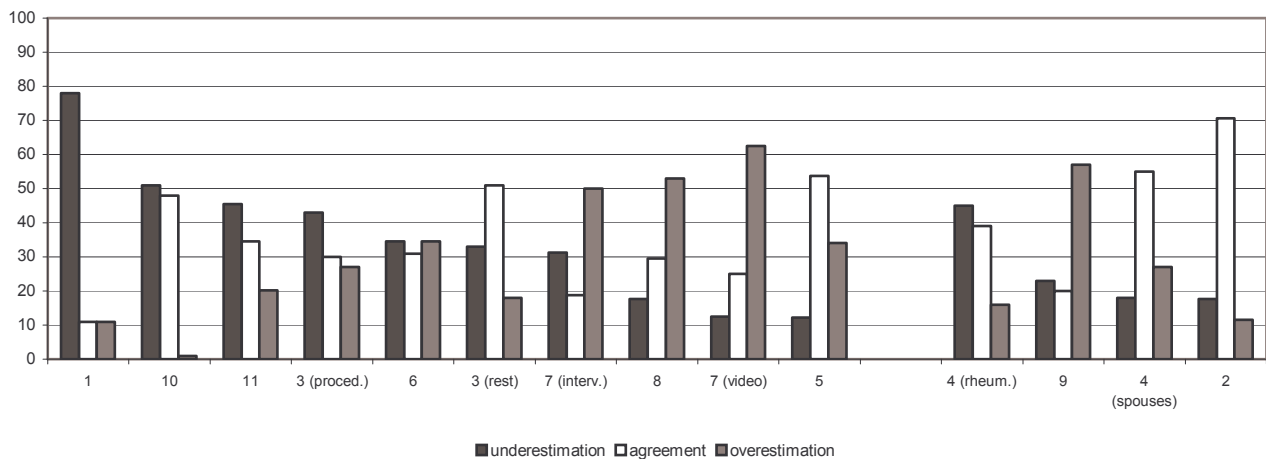


Figure 4. Distributions of agreement and disagreement depending on definition of agreement and type of scale: ± 1 cm (VAS) on left hand side, identical rating (VRS) on right hand side.

When considering the definition of agreement and type of rating scale, no clear picture emerges. There appear to be no systematic differences between studies depending on their definition of agreement or the rating scale used. At first I was surprised by this finding since I thought broader definitions of agreement rather than more narrow ones (e.g. ± 1 cm on the VAS rather than identical ratings on the VRS) more likely to produce higher proportions of agreement. Since the range of possible pain ratings, however, is larger for the VAS when compared to the VRS (10 cm rather than four or five adjectives), the definition of agreement on the VAS is not broader than the one on the VRS (2/10 compared to 1/4 or 1/5). Nevertheless, this finding

means that all studies can be included when looking at the impact of judges' relation to patients and patients' diagnoses on agreement.

4.1.2.2 Judges' relation to patients and patients' diagnoses

Proportions of agreement and disagreement depending on who judges the pain are shown in Figure 5. In seven studies (1, 2, 3, 5, 6, 10, 11) judges are health care professionals, in three ones relatives (7, 8, 9) and in one study both (4).

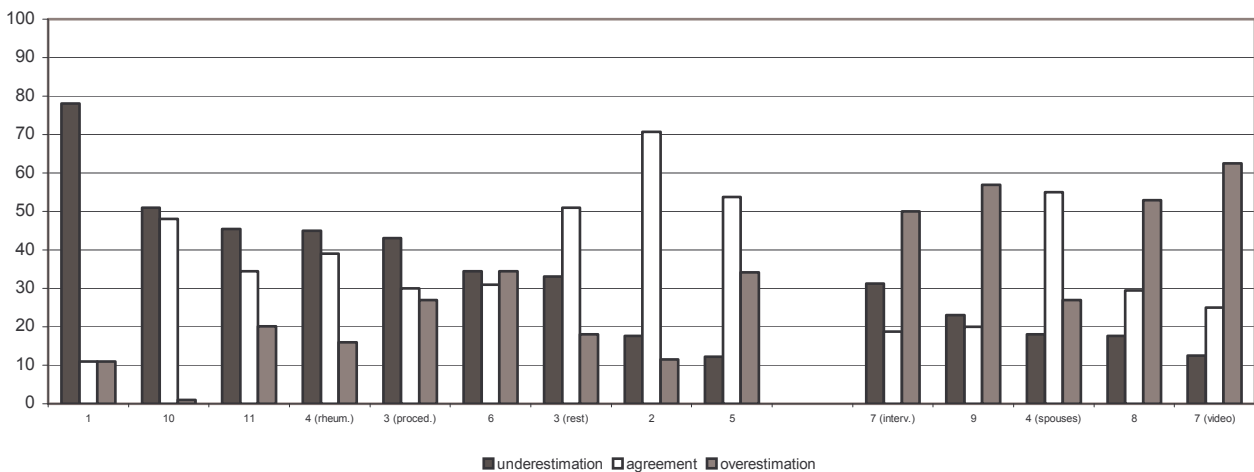


Figure 5. Distributions of agreement and disagreement depending on who judges pain: health care professionals on left hand side, family members on right hand side.

Results indicate that health care professionals are more likely to agree with or underestimate their patients' pain experience. Furthermore, they appear to be less likely to overestimate their patients' pain. Family members, on the other hand, seem to be more likely to overestimate and less likely to agree with or underestimate the pain sufferer's pain.

Proportions of (dis)agreement depending on patients' diagnoses are displayed in Figure 6. Of the eleven studies, three investigated acute pain conditions (1, 10, 11), three burn patients (3, 5, 6), three cancer patients (2, 7, 8), one osteoarthritis (4) and one rheumatoid arthritis (9).

Results indicate that there are differences depending on patients' diagnoses. Underestimation and agreement seem to be more likely for patients suffering from acute pain, overestimation for patients suffering from cancer and agreement for burn pa-

tients' pain. No clear picture emerges for patients suffering from osteoarthritis and rheumatoid arthritis pain.

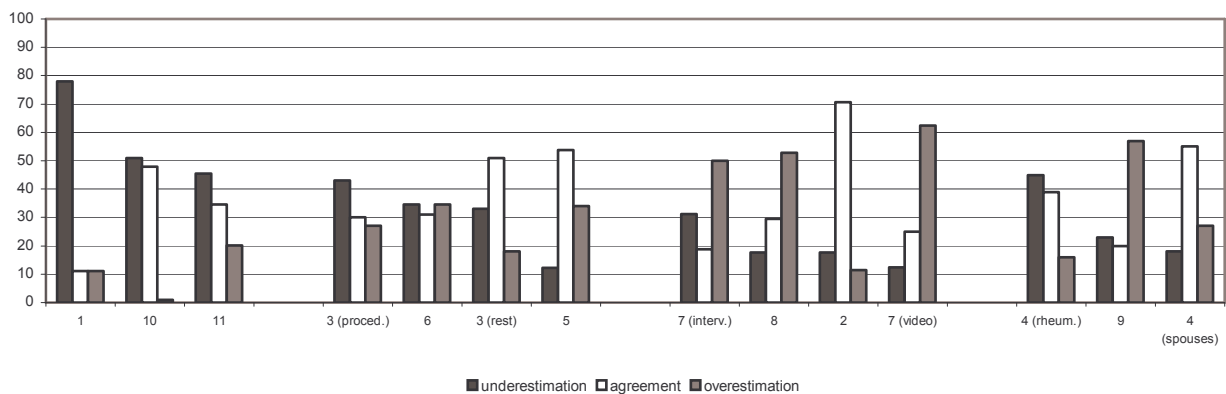


Figure 6. Distributions of agreement and disagreement depending on diagnosis of patients: first group: acute pain, second group: burn pain, third group: cancer pain and last group: osteoarthritis / rheumatoid arthritis.

Interpretation of these results, however, is constrained since judges and diagnoses are confounded in the studies: Patients whose pain was judged by health care professionals suffered primarily from postoperative or burn pain whereas patients whose pain was judged by family members suffered mainly from cancer pain.

4.1.2.3 Available cues

Of course, it is very difficult to control for cues in a non-standardised clinical situation. Accordingly, whenever information on cues was given, it was very limited. Nevertheless, differences between studies with regard to number and kind of available cues can be found.

No information on cues was given in one study (5). Interestingly however, judges in this study had additionally to rate pain-related behaviours which as the authors hypothesise might have increased nurses' sensitivity to pain. In study 11 judges were instructed to assess pain in their usual way. In study 4 it is not clear whether rheumatologists were aware of being asked about patients' general pain intensity after the medical appointment. Similarly, it is not clear whether patients and spouses in study (9) talked about the patient's pain before filling in the questionnaires.

In the remaining studies data were collected separately so that judges probably did not have the opportunity to talk to patients about patients' pain intensity. By not let-

ting judges talk to patients authors presumably wanted to avoid judges reporting exactly the intensity the patient had reported. Vignette studies, however, found significant differences between patients' and judges' ratings even though the exact pain intensity ratings of patients were available to judges (Chibnall & Tait, 1995; Chibnall et al., 1997; Tait & Chibnall, 1997). Further, by not allowing judges to talk to patients, study authors withheld a cue judges might have considered as important and used when judging pain.

4.1.3 Summary

What do these results tell us? First and before summarising the results, it is important to stress how difficult it was to find studies investigating the question of agreement. Moreover, of the 62 studies that could be found only a small proportion (18 %) was selected as useful according to criteria set prior to examining their results. Studies showed no general tendency to underestimate pain as it is often reported in the pain literature. On the contrary, there appeared to be differences depending on judges' relationship to pain patients as well as differences depending on patients' diagnoses. Since these variables were confounded, no conclusion can be drawn yet regarding the importance of each or both. Results further indicated that the definition of agreement did not affect the proportions of agreement and disagreement in the summarised studies. In the majority of studies judges were forbidden to talk to patients which made a possibly important cue, verbal report, unavailable to them.

4.2 Vignette studies

In addition to studies in the clinical context, vignette studies were used to investigate pain judgements in general and agreement between patients and judges. Agreement could be examined in vignette studies in which the pain self-report of a fictional pain sufferer was included which then can be compared to the reader's pain judgement. In the studies by Tait and Chibnall (1995; 1997) and colleague (1997) the amount of pain experienced by the sufferer is further systematically varied in order to see what effect this has on the pain judgements. A detailed description of the sample, the main research questions, measures and main results as well as comments on these studies are displayed in Table 2.

Table 2. Results of vignette studies.

Authors	Sample	Research questions	Measure	Main results
Chibnall & Tait, 1995	judges: undergraduate students (n = 80)	examine which contextual variables moderate perception and interpretation of chronic pain	<u>DV</u> : pain intensity judgement <u>measure</u> : NRS 0-10 <u>IV</u> : pain self-report, valence of relationship to person in pain, control over circumstances of pain onset, medical evidence	<u>t-tests</u> : when pain was low, no difference between pain judgements; however, when reported pain was high, less pain was allocated by judges ¹ <u>MANOVA</u> : sig. main effect of reported on rated pain intensity: the higher the reported, the higher the rated pain
Chibnall, Tait & Ross, 1997	judges: first-year medical students (n = 95)	investigate moderating influence of patient self-reported pain intensity on the effect of medical evidence availability	<u>DV</u> : pain intensity judgement <u>measure</u> : NRS 0-10 <u>IV</u> : pain self-report, medical evidence	<u>t-tests</u> : when reported pain was low, rated pain was slightly higher; when reported pain was moderate, no difference between both measures; when reported pain was high, rated pain was lower ¹ <u>MANOVA</u> : sig. main effect of reported on rated pain intensity: the higher the reported, the higher the rated pain
Tait & Chibnall, 1997	judges: internal medicine physicians (n = 80)	examine which characteristics of physician, patient and context may bias pain judgements	<u>DV</u> : pain intensity judgement <u>measure</u> : NRS 0-10 <u>IV</u> : pain self-report, diagnostic evidence, valence of patient-physician relationship	<u>t-tests</u> : when reported pain was low, rated pain was higher; when reported pain was high, rated pain was lower ¹ <u>MANOVA</u> : sig. main effect of reported on rated pain intensity: the higher the reported, the higher the rated pain

¹ low pain: 3/10, moderate pain: 5/10, high pain: 7/10

While in all three studies by Tait and colleagues (1995; 1997; 1997) different types of participants (e.g. students, doctors) were used, the results resemble each other: When the amount of pain experienced by the pain sufferer was low (3 on a scale from 0 to 10), there were either no differences between self-reports and judgements or the judgements were slightly higher than the self-reported pain. When the self-reported pain intensity was high (7 on a scale from 0 to 10), however, the participants judged the pain experience to be less severe than the person in pain.

5. Explanations for underestimation

The tendency to underestimate pain has received much attention in the pain literature. Researchers were (and still are) puzzled by its incidence and, hence, are looking for explanations. Although various explanations have been discussed, results were not conclusive (Williams, 2002). Two theories and several single variables that have been proposed to explain the occurrence of underestimation are introduced and evaluated in the next sections.

5.1 Theories proposed to explain underestimation

5.1.1 The hypothesis of developed insensitivity

Commonly, underestimation is explained by what Amanda Williams (Williams, 2002) calls the 'hypothesis of developed insensitivity'. It states that with increasing work experience health care professionals are repeatedly exposed to patients' pain. The hypothesis further proposes that because of this repeated exposure health care professionals develop a relative insensitivity to pain and/or a scale with a higher upper limit which in turn leads to underestimation of many patients' pain. Supporting evidence for this theory could stem from studies in which the amount of working experience of judges was investigated in association with underestimation.

Negative associations between length of experience and agreement with patients' pain judgements were found in studies by Choiniere and colleagues (1990), Mason (1981), Perry and Heidrich (1982) and Lenburg and colleagues (Lenburg, Glass & Davitz, 1970). They showed that less experienced judges rated pain intensities to be higher than more experienced judges. In a study by Halfens and colleagues (Halfens, Evers & Abu-Saad, 1990), in which three groups were contrasted, a curvilinear association was found: least experienced judges inferred least pain, most experienced

judges a middle pain intensity and medium experienced judges inferred most pain. However, no association between working experience and underestimation was found in studies by Dudley and Holm (1984), Everett and colleagues (1994), Hamers and colleagues (Hamers, van den Hout, Halfens, Huijer Abu-Saad & Heijltjes, 1997), Oberst (1978) and van der Does (1989). Zalon (1993) even found a trend towards a positive association. The hypothesis of developed insensitivity is further not supported by results of a study by Prkachin and colleagues (Prkachin, Solomon, Hwang & Mercer, 2001) in which judges with chronic pain patients in their families - judges who presumably are repeatedly exposed to pain too - were more accurate when judging pain than health care professionals.

In summary, there is some support for the hypothesis of developed insensitivity, although no clear picture emerges from studies that have been done so far.

5.1.2 Representativeness heuristic

Janice Lander (1990a) in a review article on clinical judgements in pain management recommended applying the clinical decision making model as proposed by Daniel Kahneman and Aron Tversky (Kahneman & Tversky, 1982; Tversky & Kahneman, 1974, 2000) to pain judgements. These authors assume that when making judgements or decisions 'people rely on a limited number of heuristic principles which reduce ... complex tasks ... to simpler judgmental operations. In general, these heuristics are quite useful, but sometimes they lead to severe and systematic errors' (Tversky & Kahneman, 2000, p.35). Misapplying simple judgemental strategies to complex inferential problems such as pain judgements can lead to inferential biases or errors of judgement. Lander (1990a) argues that among others the representativeness heuristic is applied in the pain judgement context.

Representativeness heuristic is a judgemental strategy in which objects or events are evaluated by the degree to which they are representative of a class. Errors are likely to occur when judgements are made without considering the frequency of the event in the population (the base rate).

Therefore, the question emerges whether judges underestimate pain because they ignore the base rate of patients malingering (according to Kahneman and Tversky

(1982a) an 'error of application'). This question is difficult to answer and has not yet been tested empirically. The real incidence of deception is for several reasons difficult to establish: the nature of the actions is secretive, distinguishing between conscious and unconscious motivation a challenge and skills people exercise to succeed with their lies vary (Craig, Hill & McMurtry, 1999). Consequently, it is unlikely that there is a 'true' base rate of deception that people could ignore when making their judgements. Rather than ignoring base rates it is also possible that judges underestimate pain because they do not know the real incidence (what Kahneman and Tversky (1982a) would call an 'error of comprehension'). However, when asking clinicians for their base rates they easily come up with their expectations of malingering and deception that in general are low (Leavitt & Sweet, 1986; Mendelson, 1992b).

Since the representativeness heuristic to me seems more like a post-hoc explanation than a theory helpful to explain cognitive processes underlying decision making, I find it hard to apply it to the question of agreement between patients and judges. Moreover, the research programme by Kahneman and Tversky has recently been criticised by Gigerenzer and his research group (Gigerenzer, 2000; Gigerenzer, Todd & Group, 1999) for 'too narrowly drawn norms for evaluating reasoning and too vague heuristics which have directed attention away from detailed models of cognitive processes and toward post-hoc accounts of alleged errors' (Gigerenzer, 1996, p.592).

5.2 Single variables suggested to explain underestimation

5.2.1 Base rate of deception

Although I find it hard to apply the concept of base rate as presented in the representativeness heuristic to the question of agreement, I think the general base rate of deception of observers is worth considering. The American Psychiatric Association (1994) defined malingering as 'intentional production of false or grossly exaggerated physical or psychological symptoms, motivated by external incentives such as avoiding military duty, avoiding work, obtaining financial compensation, evading criminal prosecution or obtaining drugs' (p. 683). Accordingly, Leavitt and Sweet (1986) found that symptoms surgeons considered as important regarding malingering could be categorised along two dimensions: exaggeration and incongruity. Interestingly, dissimulation (minimisation or concealment of pain) has only rarely been researched in the pain field.

Although the base rate expectancies of malingering are generally thought to be low (Fishbain, Cutler, Rosomoff & Rosomoff, 1999), individual estimates can be high (Leavitt & Sweet, 1986; Simmonds, Barlow & Kreth, 1996). Research on the effect of the deception base rate on judgements has shown that it affects the detection of deception (Ekman & O'Sullivan, 1991) and the judgements observers make (Drayer, Henderson & Reidenberg, 1999; Faust, Hart, Guilmette & Arkes, 1988; Krivo & Reidenberg, 1996; Poole & Craig, 1992).

Ekman and O'Sullivan (1991) explained the group differences in detecting liars, which they had found, by differences in the base rate of deception the varying groups expected. Faust and colleagues (1988) found that increasing the base rate expectation of deception had only little effect on success in detection of deception. However, they could also show that forewarning judges of malingering led clinicians to raise their threshold for diagnosing impairment caused by head injury from a neuropsychological test battery. Similarly, judges in a study by Poole and Craig (1992) consistently underestimated pain when they expected some deception.

From results of two of their studies (Drayer et al., 1999; Krivo & Reidenberg, 1996) Reidenberg and colleagues concluded that health care professionals think that patients overstate their pain intensity. The authors asked health care professionals to rate patients' pain intensity and to say what patients' own ratings would be. When comparing both judgements, health care professionals ratings of patients' pain was consistently lower than patients' ratings. However, health care professionals' ratings of how the patients would rate their pain were accurate. The conclusion of Reidenberg and colleagues is further supported by two additional studies: Harrison (Harrison, 1993) found that pain of patients who were perceived to understate their pain was not underestimated. And results of Dar and colleagues (Dar, Beach, Barden & Cleeland, 1992) showed that spouses overestimated pain for patients whom they judged to have a 'stoic attitude'.

5.2.2 Supporting medical evidence

The presence or absence of supporting medical evidence was found repeatedly to be associated with effects on estimation (Dudley & Holm, 1984; Halfens et al., 1990; Oberst, 1978; Taylor, Skelton & Butcher, 1984). Studies by Chibnall, Tait and col-

league (Chibnall & Tait, 1995; Chibnall et al., 1997; Tait & Chibnall, 1997), for example, revealed that the presence of medical evidence led to comparable pain ratings in sufferers and judges. In the absence of medical evidence, however, judges rated sufferers' pain intensity to be lower than sufferers' ratings.

5.2.3 Relationship between patients and judges

Chibnall and Tait (1995) investigated the impact the relationship as perceived by judges on agreement. When judges considered their relationship to the patient as positive, they agreed with patients' self-report. Underestimation occurred when the judge considered the relationship as negative.

With regard to patients' relatives two studies took into account satisfaction with relationships as an independent variable. However, their results were contradictory. Whereas Riemsma and colleagues (2000) found no relation between disagreement between patients and spouses and marital commitment scores, Miaskowski and colleagues (1997) found that disagreement was associated with less interpersonal well-being in patients. Further support of the importance of satisfaction with the relationship to the patient stems from research on solicitousness of chronic pain patients' spouses. In a review Newton-John (2002) concluded that marital satisfaction is one mediator of the association between spouses' responses to patients' behaviour and patients' functioning.

Not only the perceived relationship was found to be associated with underestimation but also the more objective relation between patient and judge. Results of a study by Prkachin and colleagues (2001) showed that three groups of judges (judges with family members suffering from chronic pain, judges without family members suffering from chronic pain and health care professionals working with pain patients) all underestimated patients pain. Yet, underestimation was greatest for health care professionals, followed by judges without pain patients in their families. Underestimation was least in the group of judges who had chronic pain patients in their families. Miaskowski and colleagues (1997) found differences even within the group of patients' relatives: When they compared spouses or children as judges with other relatives of pain patients, patients agreed more frequently with their spouses or children than with other relatives.

This further supports the finding in 4.1.2.2 that extent of agreement may vary depending on who judges the patients' pain.

5.2.4 Time spent with patients

The time judges spend with patients was found to be related to their accuracy in rating patients' pain. For family members as judges agreement was higher if judges lived with and claimed familiarity with the patient (O'Brien & Francis, 1988) and lower if judges worked full time (Miaskowski et al., 1997). For health care professionals as judges a study by Schuler and colleagues (Schuler, Neuhauser, Hauer, Oster, Razu & Hacker, 2001) showed that the less judges had seen the patient, the more uncertain they felt of making a pain judgement.

5.3 Summary and conclusion

In summary, single variables and two theories have been suggested to explain underestimation. Research with regard to the single variables was frequently atheoretical. Further, the single variables often provided post-hoc explanations. Research lends some support to the hypothesis of developed insensitivity although no clear picture emerges from studies. The theory suggested by Lander (1990a) has not been tested empirically yet. Furthermore, the question emerges whether application of the representativeness heuristic is wise, since the research programme by Kahneman and Tversky has recently been criticised by Gigerenzer and his research group (Gigerenzer, 1996, 2000; Gigerenzer et al., 1999).

To conclude, Lander's criticism from more than ten years ago seems to be still valid. She stated that 'the superficiality of the research means that it has very little to offer in the way of guidance toward understanding the causes of poor pain management or guidance toward developing solutions to the problem' (1990a, p. 21).

5.4 Alternative possibilities to explain underestimation

My puzzlement with underestimation of pain as well as my dissatisfaction with the attempts to explain it made me think about alternative reasons why underestimation might occur. One possibility became apparent to me when I studied the lens model (see 1.2) which stresses the necessity of cues without which observers cannot draw inferences. The second alternative occurred to me when I came across evolutionary

psychology, an approach to psychology rather than a theory, in which 'knowledge and principles from evolutionary biology are put to use in research on the structure of the human mind' (Cosmides, 2002). Both alternatives and their application to pain judgements are outlined in the following two sections.

5.4.1 Missing cue

The lens model emphasises the requirement of cues without which observers cannot make their judgements. Going back to the studies in which agreement between patients and judges in the clinical context was determined, in the majority of studies judges were not allowed to talk to patients about their pain intensity due to the study design. This means that verbal report of pain was not available to judges as a cue when making their judgements. If verbal report is considered to be an important cue by judges, underestimation of pain may be a result of verbal pain behaviour as cue being unavailable to judges.

This hypothesis can be supported by results of Madison and Wilkie (1995) who divided their patient group into those who reported that they would tell others that they are in pain and those who reported that they tried not to express their pain to others. The authors found that judges are more likely to correctly rate patients' pain intensity for patients who did tell others that they experienced pain.

5.4.2 Evolutionary psychology

One basic assumption of evolutionary psychologists such as John Tooby and Leda Cosmides is that the human mind constitutes of a set of information-processing devices, located in the brain, responsible for all mental activity and generating behaviour (Cosmides, 2002). Each device needs to integrate motivation, perception, thinking, emotion and behaviour in one functional unit (Gigerenzer, 2000). Evolutionary psychologists further assume that these neural circuits were designed by natural selection to solve specific adaptive problems (such as attachment development, mate search, parenting, social exchange, coalition formation) which our hunter-gatherer ancestors were faced with (Cosmides & Tooby, 2000). Whenever information-processing devices were adaptive, their genes were selected for in evolution which is why these devices are still present today and enable us to solve these problems, no matter whether they are still important for us today.

5.4.2.1 Social exchange theory

One of the topics evolutionary psychology has dealt with is social exchange, situations in which two or more individuals cooperate for mutual benefit by exchanging goods or services. In this exchange an 'individual is required to pay a cost (or meet a requirement) to an individual (or group) in order to be eligible to receive a benefit from that individual' (Cosmides, 1989, p.197). However, 'always cooperate' would not be an evolutionarily stable strategy since then cheaters (individuals who take benefits without reciprocating) could invade the group of cooperators and outreproduce them. On the other hand, 'always cheat' would not be evolutionary stable either as a group of cheaters could be invaded by people who cooperate selectively (rather than indiscriminately). The vital point which was demonstrated by evolutionary game theory is that selective cooperation would not work without a cognitive heuristic for detecting cheaters – or, more precisely, a heuristic for directing an individual's attention to information that could reveal that it is being cheated (Cosmides & Tooby, 1992; Gigerenzer, 2000). Empirically, studies carried out by Cosmides (1989) and Gigerenzer and Hug (1992) lent empirical support to the existence of a cheating detection device. Not only humans but also animals are able to recognise and expel cheaters from their buddy systems as, for example, Wilkinson (1990) demonstrated for vampire bats.

5.4.2.2 The cheating detection mechanism and pain judgements

What could the cheating detection mechanism have to do with underestimation of pain? Situations in which health care professionals judge patients' pain could be regarded as social exchange situations as they offer help ('provide a benefit') under the condition that patients are in pain ('meet a requirement'). Yet, should health care professionals become suspicious of patients not satisfying the requirements and purposefully exaggerate or even fake their pain in order to receive the benefit (for example in form of pain relief) the cheating detection mechanism might be alerted and lead to scepticism and more conservative judgements of patients' pain intensity. Accordingly, purposeful exaggeration of pain is usually responded to with anger and alarm by health care professionals (Craig et al., 1999).

5.4.2.2.1 Alerting the cheating detection mechanism

Which factors may lead to health care professionals becoming suspicious of patients' pain? In general terms, health care professionals might expect cheating whenever they suspect or infer secondary gain factors as their presence is frequently equated with malingering (Fishbain, 1994). Despite problems with the definition of secondary gain and its abuse in medical practice (Fishbain, 1994), potential secondary gains are discussed and investigated in the pain field.

In a review of scientific evidence for secondary gain in chronic pain patients Fishbain and colleagues (Fishbain, Rosomoff, Cutler & Rosomoff, 1995) found that studies in this area can be divided into two areas of inquiry. The first one concerns the sick role itself [for an explanation of the sick role concept see Parsons (1951)] or marital reinforcement of the sick role [often termed spouse solicitousness – see Fordyce (1976)] which is supposed to motivate the patient to maintain the sick role. The second area is the one in which compensation status is considered as a secondary gain factor in the maintenance of disability.

What other factors could be secondary gains? For patients suffering from acute pain receiving analgesic medication may be a secondary gain if they are addicted (which is something health care professionals were found to be afraid of, compare Cleeland, Cleeland, Dar & Rinehardt, 1986; Lander, 1990b; Lavies et al., 1992; Marks & Sachar, 1973) or if the medication has desirable psychoactive effects such as sedation or euphoria. Another cue that might make judges suspicious (and was shown to be associated with underestimation – see 5.2.2) when encountering acute as well as chronic pain patients is the absence of medical findings that support patients' pain reports which might lead to the conclusion that pain is psychogenic, i.e. that it only exists in the patient's mind.

Although these and other sources of secondary gain have received great attention in the pain field, it should be emphasised again here that there are many difficulties associated with the concept. For instance, the relationships between pain and financial compensation are fairly complex which argues against a simplistic attribution of malingering based only on the availability of compensation (Mendelson, 1992a, 1992b). Yet, the sole presence of potential financial rewards is usually equated with malingering.

ing (Fishbain, 1994). Or discrepancies between pain behaviour and physical pathology are found for a large proportion of chronic pain patients (Waddell, 1991) and do not indicate the genuineness of pain since substantial pathology can often be detected in people who are asymptomatic (Turk, 1996; Wall, 1999) and ongoing research discloses new pathophysiological sources of pain (Teasell & Merskey, 1997; Wall, 1999).

But let us return to underestimation. In order to explain underestimation, an activated cheating detection algorithm must lead to a more conservative judgement and greater underestimation in observers. This effect was shown by Poole and Craig (1992) who asked nursing and psychology students to watch the videotaped faces of patients who were undergoing a medical examination. Subsequently, participants were asked to rate the patients' pain. The authors found that judges consistently estimated the pain to be less severe when they had been primed in advance to expect some deception.

To summarise, the cheating detection device may be activated by the presence of certain context cues (such as the absence of supporting medical evidence) possibly associated with secondary gains (such as marital reinforcement or compensation status). As a result, these factors lead to more conservative judgements of pain, i.e. underestimation.

6. Aims and objectives

In addition to studies investigating proportions of agreement between patients and judges with regard to pain, there has been plenty of research on measurement of pain behaviour cues and on contextual cues (such as secondary gain factors) in the pain field. However, 'research on how these ... cues are interpreted by health professionals is still in its infancy' Skevington (1995, p. 187). One of the reasons for this might be that not very many studies have focussed on the various cues (pain behaviours as well as context) available to judges in the judgement situation and on how judges use these cues to make their judgement (i.e. how judges draw inferences about pain). Further, how judges use and interpret available cues is not only unclear for health care professionals as judges but even more so for other groups of judges such as family members.

Hence two questions emerge which should be addressed in this piece of work:

1. Can underestimation of pain be accounted for by one or both of two alternative explanations (verbal report as missing cue; alerted cheating detection mechanism)?
2. How do selected cues impact on pain judgements of patients' relatives?

The next three major sections outline three studies, each of them investigating one of the questions beyond. The first study is concerned with the question of whether health care professionals regard verbal report as an important cue and what cues make them cautious when judging pain. Its results are important prerequisites for the second study in which the hypotheses are tested that underestimation can be explained by verbal report as missing cue and / or by an alerted cheating detection mechanism. The aim of the third study was to examine whether and how certain cues affect the judgements by relatives of chronic pain patients. These three studies are followed by an overall discussion.

Study 1

Which cues are used to judge pain? Perceptions of health care professionals

Aims and objectives

The lens model emphasises the importance of cues to make judgements, and many cues are available in the clinical situation in which pain is judged. However, not very much is known on how health care professionals weigh the various cues discussed in the literature and potentially available in the clinical encounter. Due to the alternative accounts for underestimation I proposed in the introduction, I had several concepts in mind (importance, manipulation, cautiousness) in the light of which I wanted to explore health care professionals' perceptions of cues. Further, I was interested to explore the relations between these concepts.

1. The concept of importance: verbal report as missing cue

The lens model emphasises cue utilisation, i.e. how important judges consider cues to be for their judgement, as one important concept when investigating judgements. During the last three decades the importance of pain patients' verbal report has been emphasised in a number of articles. Yet, in many empirical studies comparing pain judgements, verbal report as cue was either explicitly not given by experimenters or it is not clear from the articles which cues were or were not available for judges. Therefore, one hypothesis to explain underestimation was that verbal report as cue was not made available to people although they regard it as important when judging pain. A prerequisite for this hypothesis, however, is that verbal report is not only considered as important in the literature but also by judges.

There were three studies asking health care professionals to order cues regarding their importance and including verbal as well as non-verbal pain behaviours. All three yielded different results: only one found verbal report to be the most important cue (McKinley & Botti, 1991), the other two ones found non-verbal pain behaviours to be more important than verbal report (Jacox, 1980; Oberst, 1978). Therefore, the first objective was to ask health care professionals about the importance cues have when they have to judge pain.

2. The concepts of cautiousness and manipulation: alerted cheating detection mechanism

The second account for underestimation proposed was an alerted cheating detection mechanism. However, there are no studies about which cues possibly could alert this device. Two possibilities came to mind from social contract theory. One is more explicit: possibly there are cues, such as secondary gain factors, that make judges suspicious. More implicit is the possibility that manipulation of cues might cause suspiciousness in judges too. Possibly the cheating detection device is also alerted whenever cues in a judgement situation are discrepant. In this case judges should rely more on cues that are more difficult rather than on those easier to manipulate. And, in fact, a study by Poole and Craig (1992) could show that judges tended to rely more on non-verbal behaviour whenever verbal and non-verbal pain behaviour were discrepant. Hence, the second and third objective was to ask health care professionals about the ease of manipulation of cues and about the cautiousness cues may cause.

To summarise, the study's three objectives were to explore the following questions in the context of pain judgements:

1. Which cues do health care professionals believe to be important? (Objective 1)
2. Which cues do health care professionals believe to be easily manipulated? (Objective 2)
3. Which cues do health care professionals believe make them cautious? (Objective 3)

Method

1. Design

To investigate these questions an exploratory study with a cross-sectional design was set up. The study consisted of three questions according to which participants had to rank order up to ten cues:

1. rank order according to the **importance** cues have for the pain judgement,
2. rank order according to how easy the cues are to **manipulate** and
3. rank order according to how **cautious** cues make participants when judging pain.

The cues which had to be rank ordered were presented in random order. The order of the three questions (importance, manipulation and cautiousness) was not varied. Although this risks response effects which cannot be controlled, presenting the three questions randomly risks difficulty understanding the task and, consequently, introduces unreliability. The order of the questions followed their inherent logic with importance being a basic task, followed by ease of manipulation which possibly leads to cautiousness.

In a pilot study with six participants I tested the use of a numerical rating scale to rate cues for each of the questions. However, many of the volunteers simply grouped the cues into the two groups the cues were taken from (pain behaviour vs. contextual cues) without differentiating within these groups. Therefore, I decided to replace the numerical rating scale as a response mode by the rank ordering task.

2. Participants

An initial approach to the Ethics Committee to explore whether full ethical approval was needed for this study established that there was no need. The inquiry and the Ethic Committee's reply are included in Appendix II.1. As the study was exploratory, no power was calculated a priori. A rule of thumb for multidimensional scaling procedures suggests that at least 25 people should be included in the analysis (Morley, personal communication). To improve power I aimed for 60 volunteers altogether.

Next, permission was sought from the Medical Directors for Accident and Emergency (A & E), Orthopaedics, Oncology and Palliative Care, the Physiotherapy Outpatient Department and Rheumatology. These specialties were selected as many patients who attend these departments are in pain (Daut & Cleeland, 1982; Donovan et al., 1987; Johnston, Gagnon, Fullerton, Common, Ladores & Forlini, 1998; Winefield, Katsikitis, Hart & Rounsefell, 1990). All but Rheumatology agreed to participate in the study. After Medical Directors had given their permission, one or several consultants and nurses were approached to ask which times would be most convenient for staff to take part. To explain aim and content of the study and ask for participation, posters were put up in the staff's coffee rooms (A & E, Oncology), or letters (Orthopaedics) and emails (Palliative Care, Physiotherapy) containing the same information as the

poster were sent to possible volunteers. One example poster is included in Appendix II.2.

3. Materials

A questionnaire containing all materials described below is included in Appendix II.3.

3.1 List of cues

In recognition of participants' time and task complexity I set an upper limit of ten cues. Half of the cues were the five categories of pain behaviours as established in the literature (Craig et al., 2001; Hadjistavropoulos et al., 2001; Turk & Melzack, 2001): verbal self-report, paralinguistic vocalisations, physiological activity, bodily activity and facial expression. The other half were context cues.

3.1.1 Choice of context cues

The limit on the total number of cues necessitated a choice of possible context cues. Since they should alert the cheating detection device, cues were chosen that are either discussed in association with underestimation of pain or as secondary gains.

The presence of medical evidence has been shown to be related to pain underestimation (Chibnall & Tait, 1995; Chibnall et al., 1997; Tait & Chibnall, 1997). Associated with presence or absence of medical evidence as context factor is the still present distinction of psychogenic versus organic pain [even though a large proportion of chronic pain complaint has little or no relationship to detectable organic pathology, substantial pathology can often be demonstrated in people who do not complain about pain and ongoing research can disclose new pathophysiological sources of pain (Craig et al., 1999)]. To sample the distinction between psychogenic as compared to organic pain I included the history of a psychiatric illness as context cue.

Probably the most frequently discussed secondary gain in the pain field is the presence of a disability claim (Mendelson, 1992a, 1992b). Patient history of drug abuse was chosen since doctors and nurses may be reluctant to prescribe and deliver an adequate dose of opioid analgesic because of their fear that a patient may become addicted (Cleeland et al., 1986; Lander, 1990b; Lavies et al., 1992; Marks & Sachar,

1973). Further, history of drug abuse appeared as a factor that made A & E staff suspicious of pain complaints in a previous study (Kappesser & Williams, 2002).

Since Fordyce (1976) established the concept of pain behaviour and its importance in the treatment of chronic pain patients, attention from family (e.g. a solicitous spouse) has been considered as positive reinforcement and therefore as a factor which maintains pain behaviours (Flor, Turk & Rudy, 1989; Lousberg, Schmidt & Goenman, 1992; Romano, Turner, Friedman, Bulcroft, Jensen, Hops & Wright, 1992). Therefore, it was included as a contextual cue.

To summarise, the five chosen contextual cues were: results of medical investigation, history of psychiatric illness, disability claim, history of drug abuse and attention from family. As only a limited number of context cues could be chosen, the choice of context cues although theoretically justifiable must be incomplete (e.g. patients' ethnic background was left out). Hence, after rank ordering the cues participants were asked to list any missing cues they regarded as important for the question.

3.1.2 Presentation of cues

The cues were presented in random order for each of the three questions. For the importance question participants were given all ten cues when present to rank order (e.g. verbal report of pain, facial expression of pain, supporting medical results, history of drug abuse). For the manipulation question participants were given the five pain behaviour cues as nonverbal pain behaviours are perceived being less easily manipulated than verbal report (Craig & Prkachin, 1983; Johnson, 1977). For the cautiousness question participants were given all ten cues to rank order: the five pain behaviour cues and the results of medical investigations when absent (e.g. no self-report, no facial expression, no supporting results) and the remaining four contextual cues when present.

4. Procedure

Volunteers were approached before the beginning of, during or after the end of their shift. Aim and content of the study were explained: I said that I was interested in how people try to judge pain, that in the literature various factors would be discussed as important when judging pain and that I wanted to know from people who have to

judge pain on a daily basis what they think about the factors discussed in the literature. Then participants were verbally asked for their consent.

After the participants had agreed to participate in the study, age, sex, qualification, overall length of work experience, specialty and length of work experience in their specialty were recorded.

The first task was introduced to people as a list of factors which would be discussed as important when judging pain. First, participants were asked to 'please order the following factors according to the **importance they have for you when judging pain**. Assign 1 to the most important, 2 to the second most important and so on'. Correspondingly, they were asked to 'please order the following factors according to **how easy they are for the patient to control or manipulate**' and to 'please order the following factors according to **how cautious they make you when judging pain**'. After the importance and the cautiousness ranking task, participants were asked whether anything was missing which is important for them or makes them cautious when judging pain.

After participants completed the last task, they were offered the chance to ask questions concerning content and aim of the study.

Overall, filling in the questionnaire took participants between ten and twenty minutes.

5. Statistical analyses

5.1 Rank ordering of the cues

In addition to descriptive statistics (median, 25- and 75-percentile, range and skewness), the results of the rank ordering task were analysed using Kendall's coefficient of concordance (W). The values of this coefficient expand between 0 and 1 and indicate whether or not judges essentially apply the same standard. The higher a value the more volunteers are concordant (Siegel, 1956). Differences between cues were investigated using the Friedman Test for several related samples and the Wilcoxon signed rank test for post-hoc analyses of possible occurring differences.

Further, a nonmetric multidimensional scaling procedure (PROXSCAL, implemented in SPSS 11) was applied to the data in order to represent spatially the similarities of and dissimilarities between objects. I chose PROXSCAL as SPSS is the programme most widely known and used in social science research.

6. Feedback

After the data collection and the statistical analyses were completed, a feedback letter with the main results was sent to the Medical Directors and the Senior Nurse Managers. It was further put up as a poster in the staff's coffee room or sent to participants as letter or as an email attachment. The method of feeding back the results depended on the way participants had been approached to explain aim and content of the study. A sample feedback sheet is included in Appendix II.4.

Results

The results are divided into four sections: First, the sample of participants will be described. Further, and separately for each of the questions of interest (importance, ease of manipulation and cautiousness), the results of the rank ordering tasks and the multidimensional scaling solutions will be outlined.

1. Description of the sample

Altogether 65 nurses, doctors and physiotherapists working for one of two big London teaching hospitals took part in the study. Nurses and doctors worked for one of the following specialties: A & E ($n = 15$), Orthopaedics ($n = 19$), Oncology and Palliative Care ($n = 15$). All physiotherapists ($n = 16$) worked in the outpatient unit.

Among the 65 participants were four students and one radiographer. Three students were on placement on the orthopaedic ward, one in the physiotherapy department. The radiographer worked for the Oncology Department and due to her additional qualifications was also involved in patient treatment as a nurse.

1.1 Qualitative background variables

Among the participants were more women (78.5 %) than men. Nurses (75.5 %) outnumbered doctors. There was an association between sex and profession: more doc-

tors were male (63.6 %) and more nurses female (88.2 %). Likewise, the majority (80 %) of physiotherapists was female. These results were typical of National Health Service hospital staffing.

In terms of seniority, of the doctors, 36.4 % were juniors, 18.2 % more advanced and 45.5 % seniors. In the group of nurses, 8.8 % were health care assistants without a nursing diploma, 47.1 % had junior nursing grades, 35.3 % senior nursing grades and 8.8 % had the most senior grades. Among the physiotherapists 20 % were juniors, 60 % intermediate and 20 % seniors.

1.2 Quantitative background variables

The youngest participant in the sample was 19 years, the oldest 58 years old. On average participants' age was 31 years and three months. Most of the health care professionals had been working in their profession for three years and about 50 % had been working for six or less years in their profession. Of the participating nurses and doctors, about 50 % had been working in their specialty for approximately four years.

2. Results regarding the importance of cues

The five pain behaviour and five context cues were ranked by importance, ease of manipulation by patients and eliciting caution when judging pain.

Participants found the rank order task unusual but not too difficult, frequently adding comments about the difficulties health care professionals face when having to judge pain from the task. One volunteer refused to rank order the cues regarding importance and cautiousness on the grounds that ranks would vary across patients. However, she had no problem ranking the cues regarding manipulation. Another participant added two cues for the importance and cautiousness rankings which he considered essential. Therefore, the sample size for the importance and cautiousness rankings was $n = 63$, the one for the manipulation rankings $n = 65$.

Not all distributions of the ranks assigned to each of the cues for each of the three questions appeared to be normally distributed. Therefore, skewness was reported in addition to medians, modes, ranges, 25th - and 75th -percentile. Then Kendall's coef-

ficient of concordance was reported, followed by an investigation of differences between cues and cues listed as missing by participants.

2.1 Importance rankings

Z-values for skewness for the single cues ranged from -4.13 to 5.95 . Interestingly, all pain behaviour cues took positive skewness values, i.e. they had a tendency to be skewed to the left (from 0.52 for 'posture' to 5.95 for 'verbal report'). On the contrary, all context cues, with the exception of 'results of medical investigations', took negative skewness values, i.e. they had a tendency to be skewed to the right (from -1.28 for attention of family to -4.13 for disability claim). Results of medical investigations ($z = -0.06$) was neither skewed to the left nor to the right.

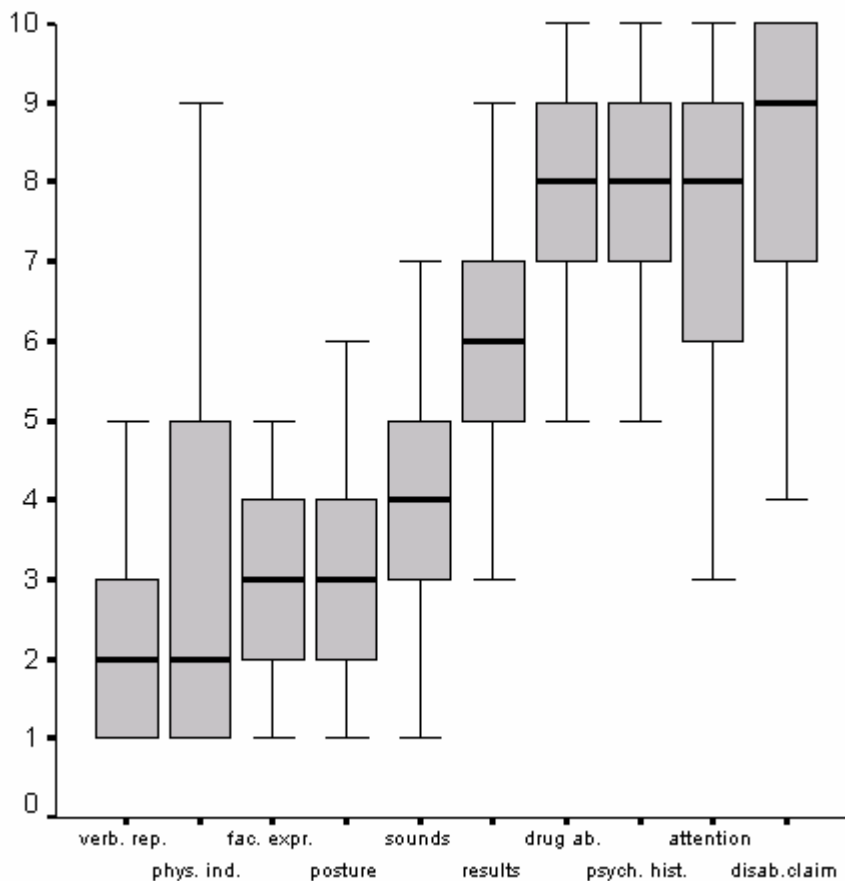


Figure 1. Boxplots for importance rank ordering.

Figure 1 displays the range, 25th percentile, median and 75th percentile for each cue (boxplots). The box is the portion of distribution falling between 25th and 75th percen-

tiles, the horizontal line in the box represents the median. The vertical lines outside each box connect the largest and smallest values.

As shown, variation was confined: 50 % of participants chose and assigned one of three ranks to most cues and four or five ranks to 'physiological indices', 'disability claim' and 'attention from family'. Furthermore, participants seemed to distinguish between the pain behaviour and the context cues: three quarters of participants assigned a rank from 1 to 5 to each of the pain behaviour cues and a rank from 5 up to 10 to each of the context cues.

2.1.1 Kendall's coefficient of concordance for importance rankings

Technically speaking, the Kendall coefficient of concordance (W) is a measure of the relation among several rankings each of a different rater and may take values between 0 (no agreement among rankings) and 1 (perfect agreement among rankings). A high value of W may, therefore, be interpreted as meaning that judges are essentially applying the same standard in ranking the objects under study (Siegel, 1956).

The coefficient of concordance for importance was fairly high and significant ($W = 0.65$, $p < 0.001$), indicating that judges essentially applied the same standard when they ranked the ten cues regarding their importance.

2.1.2 Differences between cues for importance rankings

The Friedman Test for several related samples indicated that there were statistically significant differences in importance between the cues ($\chi^2 = 370.34$, $p < 0.001$). For post-hoc analysis a Wilcoxon signed rank test was used to examine between which pairs of cues the differences are significant. Because of the multiple comparisons a more stringent p -value ($p < 0.05/45 = 0.001$) was chosen.

The only significant difference within the pain behaviour cues occurred between verbal report and sounds ($z = -4.64$, $p < 0.001$). However, all pain behaviour cues differed significantly from all context cues (all p -values < 0.001). Within the group of context cues, 'results of medical investigations' differed significantly from all other context cues (all p -values < 0.001). The results are summarised in Appendix II.5.

2.1.3 Missing importance cues

Forty-two participants (65 %) did not add any missing cue. For those who added missing cues, answers were grouped according to similarity of content. A summary of all answers can be found in Appendix II.6 as only answers who were given by at least two participants are reported below (percentages were calculated by dividing the number of participants mentioning any missing cue by the number of participants answering at all):

- impact of pain on daily life, function, quality of life (27.3 %),
- other, more specific objective signs (e.g. muscular atrophy, other medical conditions) (22.8 %),
- temporal pain characteristics (chronic, time since onset of pain; 18.2%),
- medication (previous, current, dependency; 13.6 %),
- compensation (9.1%),
- cultural, ethnic background (9.1%),
- mental, emotional state (9.1%),
- behaviours during examination (e.g. very quiet, not complaining; 9.1%)

2.2 Spatial similarities of cues: multidimensional scaling (MDS) procedures

2.2.1 Introduction

MDS allows the user to represent similarities of objects spatially by recovering underlying structures among them. Depending on the underlying model applied, the graphical representation of the space varies: in distance models, such as the one implemented in SPSS, the proximity of points in the space is used to represent their empirical similarity or dissimilarity. Thereby, each object is represented by a point in a multidimensional space. The points are arranged in this space (stimulus space) so that two similar stimuli are represented by points which are close together. Consequently, two dissimilar objects are represented by two points which are far apart.

MDS programmes calculate coordinates for the stimuli at a starting configuration. Distances of these coordinates are calculated from the data, computed and iterative adjustments are made to the coordinates to achieve the best fit. How well the distances fit the data is quantified by badness of fit indices (such as stress indices) and goodness of fit indices (such as congruence coefficients).

To decide how many dimensions are needed to best represent the data, certain criteria have been proposed (Borg & Groenen, 1997; Schiffman, Reynolds & Young, 1981): First, changes in dimensionality need to be examined with associated

changes in the fit indices. Generally, badness of fit indices will decrease whereas goodness of fit indices will increase with an increasing number of dimensions. In scree plots, badness of fit indices (stress values) associated with different dimensionalities are plotted against the dimensionalities. Usually, the curve in a scree plot takes a convex form (monotonically decreasing at an increasingly slower rate with increasing number of dimensions). An indicator for the appropriate dimensionality to be chosen is an elbow in this curve, i.e. a point where the decrements in stress begin to be less pronounced. Another criterion is the interpretability, as commonly interpretation of the stimulus space becomes more complicated with increasing number of dimensions.

Due to the rank ordering task an ordinal measurement level was chosen for all three ranking tasks. In all scree plots the normalised raw stress is shown as badness of fit index as this is the index PROXSCAL minimises when trying to best fit the distances to the data.

2.2.2 Multidimensional scaling of importance rankings

2.2.2.1 Number of dimensions

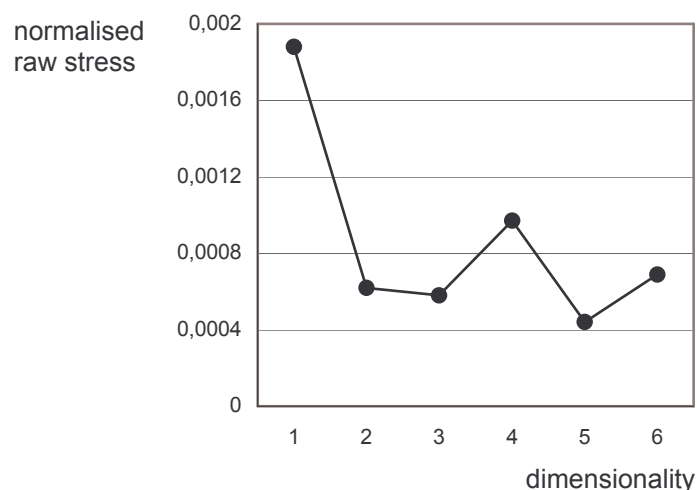


Figure 2. Normalised raw stress plotted against dimensionality for the importance ranking data.

As shown in the stress plot in Figure 2, the curve in the scree plot was not convex since stress increases for four and six dimensions. Nevertheless, there appeared to be an elbow at dimensionality 2 which was why a two-dimensional solution was chosen.

2.2.2.2 Stimulus space

Figure 3 and Table 1 show the coordinates and the resulting stimulus space for the two-dimensional solution.

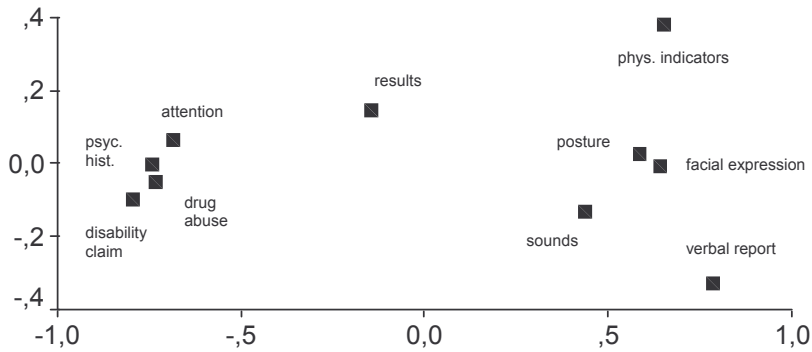


Figure 3. Similarities and differences between the pain behaviour and the context cues in the two-dimensional solution.

Table 1. Stimulus coordinates for the two-dimensional solution for importance.

cues	dimension	
	1	2
verbal report	0.785	- 0.327
facial expression	0.640	- 0.006
sounds	0.436	- 0.133
postures, movements	0.582	0.027
physiological indicators	0.653	0.382
results of med. investigations	- 0.143	0.145
disability claim	- 0.794	- 0.099
drug abuse history	- 0.732	- 0.050
attention from family	- 0.684	0.065
psychiatric history	- 0.744	- 0.005

2.2.2.3 Interpretation of the stimulus space

In the stimulus space there were two groups of cues, one on the right hand side, the other one on the left hand side with a single cue in the middle. The group on the right hand side consisted of the five pain behaviour cues, the single cue in the middle was 'results of medical investigations' and the group on the left hand side were the remaining four context cues.

The order of cues on the first dimension corresponded to the rank order of cues regarding their importance with the pain behaviour cues ('verbal report', 'physiological

indicators', 'facial expression', 'posture' and 'sounds') being more important than the context cues ('results of medical investigations' followed by the remaining four context cues that are very close together).

In addition, correlation coefficients for the medians in the importance ranking and cue coordinates on both dimensions supported this interpretation: medians correlated highly with the first dimension (Pearson's $r = -0.99$, $p < 0.001$), but not with the second dimension (Pearson's $r = -0.06$, $p = 0.862$). Therefore, the first dimension could be labelled 'importance'.

The order of cues on the second dimension was not that conclusive. However, if only the pain behaviour cues and the 'results of medical investigations' were considered, the order of the cues appeared to be similar to the rank order regarding ease of manipulation. The cues that were ranked as easy to manipulate (such as 'verbal report' and 'sounds') fell more towards the bottom of the y-axis while the cues considered difficult to manipulate (such as 'physiological indicators' and 'results of medical investigations') were more towards the top of the y-axis.

This interpretation was confirmed by high correlations between the five pain behaviour cues when ranked regarding their ease of manipulation and their coordinates. Correlation between the cue coordinates on the second dimension and the rank medians was very high (Pearson's $r = 0.96$, $p = 0.01$), correlation with coordinates on the first dimension, however, was low (Pearson's $r = -0.15$, $p = 0.812$). Hence, the second dimension could be labelled as 'ease of manipulation'

3. Results regarding the ease of manipulation of cues

3.1 Ease of manipulation rankings

Z-values for skewness for the single cues ranged from -9.73 to 6.54 . Whereas 'physiological indicators' (-9.73) and 'postures' (-3.79) were skewed to the right, 'sounds' (3.11) and 'verbal report' (6.54) were skewed to the left. The z-value for skewness for 'facial expression' (-0.77) was fairly close to zero.

As shown in the boxplots in Figure 4, there was little variation: for four of the cues 50% of ranks were either of two consecutive ranks. For 'physiological indicators' half

the participants agreed on one rank. A further indicator of high agreement was that the median of each cue was identical to the mode.

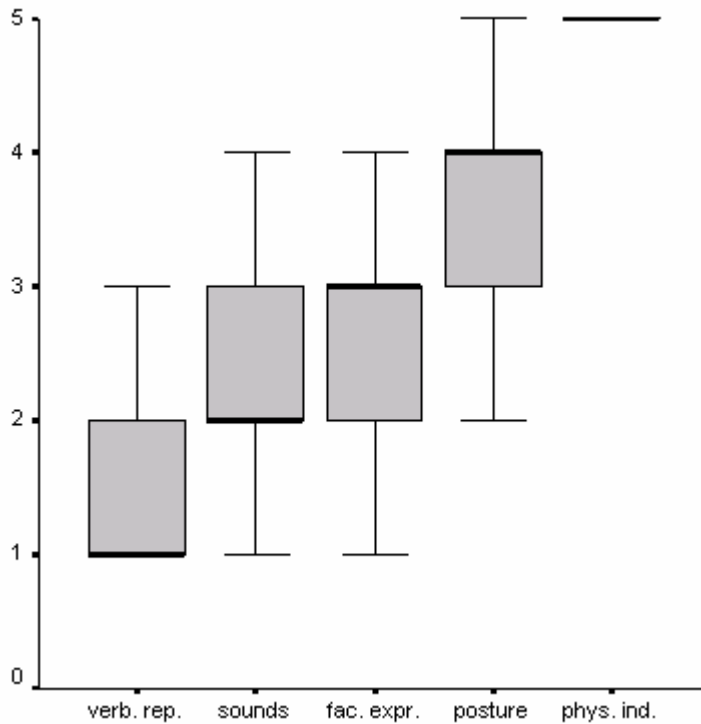


Figure 4. Boxplots for ease of manipulation rankings.

3.1.1 Kendall's coefficient of concordance for ease of manipulation rankings

As for the importance rankings the concordance coefficient was high over all groups ($W = 0.59, p < 0.001$), indicating that judges applied the same standard when ranking the six cues.

3.1.2 Differences between cues for ease of manipulation rankings

The Friedman Test for several related samples indicated that again there were statistically significant differences in ease of manipulation ranks between the cues ($\chi^2 = 154.29, p < 0.001$). For the Wilcoxon signed rank test a p -value of < 0.003 ($p < 0.05/15 = 0.003$) was chosen. Results (in detail summarised in Appendix II.7) showed that significant differences between all cue pairs occurred (with $p = 0.001$ or $p < 0.001$).

3.2 Multidimensional scaling of ease of manipulation rankings

3.2.1 Number of dimensions

The scree plot for the stress values (Figure 5) showed that stress values for one-, three- and four-dimensional solutions were fairly low. Because of the number of cues and ease of interpretation a one-dimensional solution seemed most appropriate to represent similarities of the data.

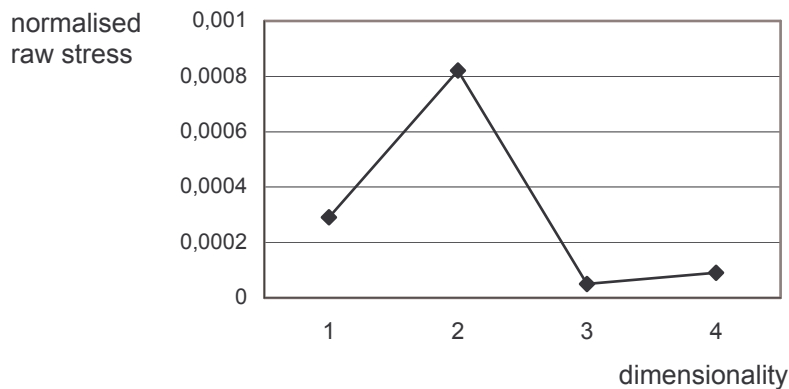


Figure 5. Normalised raw stress plotted against dimensionality for the five pain behaviour cues only.

3.2.2 Stimulus space

Table 2 shows the coordinates of the cues.

Table 2. Stimulus coordinates for the one-dimensional solution for the manipulation ranking data.

	<u>dimension</u>
<u>cues</u>	<u>1</u>
verbal report	- 0.884
sounds	- 0.440
facial expression	0.000
postures, movements	0.391
physiological indicators	0.933

3.2.3 Interpretation of the stimulus space

The five pain behaviour cues on this dimension are in exactly the same order as the median ranks for the rating tasks: 'verbal report' (considered easiest to manipulate), 'sounds', 'facial expression', 'posture' and 'physiological indicators'. Moreover, this was also reflected by a nearly perfect Pearson's correlation coefficient between the median ranks and the cue coordinates for this dimension ($r = 0.986$, $p = 0.002$).

4. Results regarding cautiousness

4.1 Cautiousness rankings

Only the z-value for skewness for absence of verbal report (2.96) showed substantial skewness to the left. All other values were fairly close to zero with values ranging from - 1.66 (attention from family) to 1.8 (no postures).

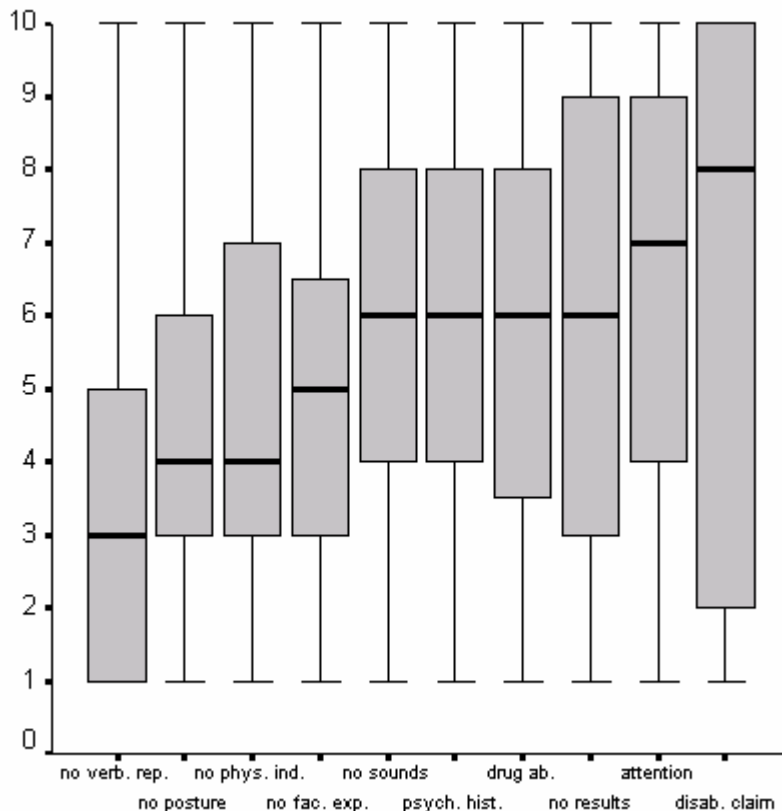


Figure 6. Boxplots for cautiousness rankings.

The distributions of the single rankings, as displayed in Figure 6, showed greater variation than the distributions for importance or ease of manipulation rankings and ranks for all cues ranged from 1-10. Medians, however, covered a smaller range (3-7). Despite the higher variation of the ranks assigned to cues, participants in general rated pain behaviour cues as causing more caution than context cues, a result similar to the importance ranking order.

4.1.1 Kendall's coefficient of concordance for cautiousness rankings

The concordance coefficient for cautiousness was very low despite the χ^2 -test indicating significant difference from 0 ($W = 0.08$, $p < 0.001$). The low coefficient indi-

cated that judges applied different standards when ranking the ten cues according to how cautious they would make them.

4.1.2 Differences between cues for cautiousness rankings

Again, the Friedman Test for several related samples indicated that there were statistically significant differences in cautiousness ranks between the cues, however, the χ^2 -value was not as high as it was for importance or ease of manipulation rankings ($\chi^2 = 44.44$, $p < 0.001$). For the Wilcoxon signed rank test a p -value of < 0.001 ($p < 0.05/45 = 0.001$) was chosen.

Results are summarised in detail in Appendix II.8. They showed that significant differences occurred only for pairs in which 'no verbal report' as cue was included. 'No verbal report' differed significantly from 'no sounds' ($z = - 3.97$, $p < 0.001$), 'no results of medical investigations' ($z = - 3.59$, $p < 0.001$), 'no attention from family' ($z = - 3.52$, $p < 0.001$) and 'no psychiatric history' ($z = - 3.57$, $p < 0.001$).

4.1.3 Missing cautiousness cues

Again, 42 participants (65 %) did not add any missing cue. For those who added missing cues, answers were grouped according to similarity in context. A summary of all answers can be found in Appendix II.9 as only answers who were given by at least two participants are reported below (percentages were calculated by dividing the number of participants mentioning any missing cue by the number of participants answering at all):

- legal claims (22.7 %)
- behaviour associated with medication (dependence, failure to attempt self-medication, no response to analgesics at any stage; 18.2 %)
- culture, ethnic background ('Mediterranean disease'; 13.6 %)
- attitudes of patients regarding pain and diagnosis (13.6 %)
- meeting a patient repeatedly in A & E (9.1 %)
- patients' behaviours (fear avoidance, made no efforts themselves; 9.1 %)
- if pain has no impact on life, causes no functional disability (9.1 %)
- discrepancies (between physical signs and verbal report; between patient's and patient's family report; 9.1 %)

4.2 Multidimensional scaling of cautiousness rankings

4.2.1 Number of dimensions

In the scree plot for the cautiousness ranking data (Figure 7) a clear elbow emerged at a two-dimensional solution.

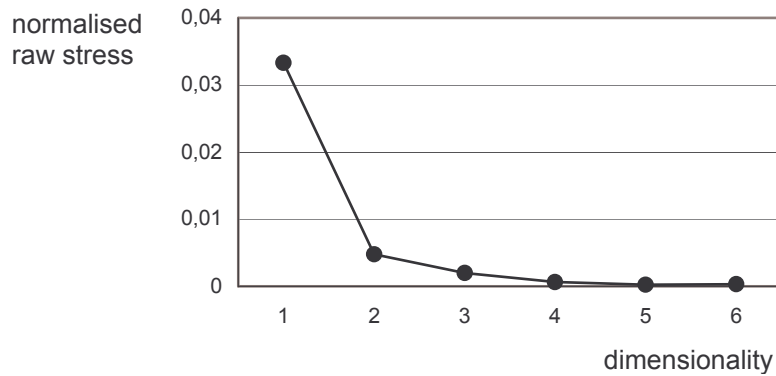


Figure 7. Normalised raw stress plotted against dimensionality for the cautiousness ranking data.

4.2.2 Stimulus space

Figure 8 and Table 3 display the stimulus space and show the coordinates of the cues.

Table 3. Stimulus coordinates for the two-dimensional solution for the cautiousness ranking data.

cues	dimension	
	1	2
no verbal report	- 0.909	0.395
no facial expression	- 0.385	0.086
no sounds	- 0.220	0.125
no postures, movements	- 0.576	- 0.043
no physiological indicators	- 0.436	- 0.218
no results of med. investig.	- 0.048	- 0.696
disability claim	1.102	0.166
drug abuse history	0.500	0.394
attention from family	0.517	- 0.301
psychiatric history	0.454	0.092

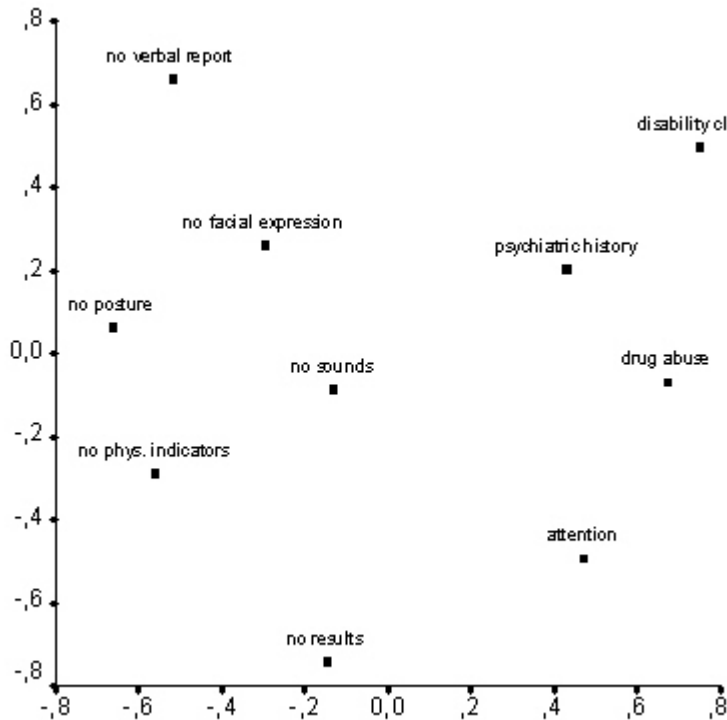


Figure 8. Similarities and differences between the pain behaviours and the context cues in the two-dimensional solution.

4.2.3 Interpretation of the stimulus space

An interesting picture emerged for the cues in this two-dimensional solution. There were two groups of cues and two single cues: four context cues on the right hand side ('disability claim', 'drug abuse', 'psychiatric history' and 'attention'), the cue 'no results of medical investigations' at the bottom in the middle, followed by a group of four pain behaviours more towards the left hand side ('no sounds', 'no facial expression', 'no physiological indicators' and 'no posture') and the cue 'no verbal report' at the top on the very left hand side.

This stimulus space noticeably resembled the importance stimulus space rotated by 180 degrees and wider distances. This was not surprising as rankings for importance and cautiousness were very similar. The order of cues on the first dimension was similar to the one for the importance ranking, with the cue that was most important and its absence most cautiousness-inducing on the right hand side. On the second dimension the pain behaviour cues easier to manipulate were nearer the top of the y-axis. Therefore, the first dimension could be labelled 'importance' whereas the second dimension could be labelled 'ease of manipulation'.

Again, this interpretation was supported by correlation coefficients between medians of rankings and coordinates in the stimulus space. Medians of cautiousness and importance rankings correlated strongly with coordinates on the first dimension (cautiousness: Pearson's $r = 0.93$, $p < 0.001$; importance: Pearson's $r = 0.957$, $p < 0.001$), but not on the second dimension (cautiousness: Pearson's $r = -0.16$, $p = 0.661$; importance: Pearson's $r = -0.03$, $p = 0.942$). The pain behaviour cue coordinates for the second but not for the first dimension correlated strongly with medians of the manipulation ranking (second dimension: Pearson's $r = -0.97$, $p = 0.005$; first dimension: Pearson's $r = -0.36$, $p = 0.552$).

Discussion

1. Discussion of results

Although participants were asked to rank order cues with regard to three different questions – importance, ease of manipulation and cautiousness - the results showed that these three different concepts were interrelated. The two cues considered as most important, 'verbal report' and 'physiological indicators', differed with regard to their ease of manipulation: 'verbal report' was considered as the pain behaviour easiest to manipulate, 'physiological indicators' as the pain behaviour most difficult to manipulate. This was also found in the resulting MDS solution. The two-dimensional stimulus space revealed that the concept 'importance' included 'ease of manipulation' as an underlying second dimension in addition to the first dimension identified as 'importance'.

The rank order regarding cautiousness was found to be fairly similar to the one for importance. However, whereas participants agreed to a very high extent on the importance and ease of manipulation rank order, they agreed to a far lesser extent on the cautiousness order as indicated by higher variability in the ranks as well as by a low concordance coefficient. Not only the cautiousness rank ordering but also the stimulus space was found to be very similar to one of importance. The only differences were that cues in the cautiousness space were spread wider apart and that the space itself was rotated by 180 degrees. Because of the similarity between the importance and the cautiousness stimulus space, both dimensions in the cautious-

ness space could be interpreted similarly to the ones in the importance space: the first one as 'importance' and the second as 'ease of manipulation'.

Another interesting result was that participants tended to distinguish between pain behaviour and context cues. This was reflected on one hand in the rank ordering of the cues and on the other hand in the stimulus spaces. In the importance rank order pain behaviours were judged to be more important than context cues. Within the group of context factors the cue 'results of medical investigations' was distinctive in that it was more important than all other cues. This grouping of cues was also reflected in the stimulus space which contained two groups of cues and a single cue: one group were the pain behaviours, the second one the context cues with the exemption of 'results of medical investigations' that was located between both groups.

The distinction between pain behaviours and context cues also emerged for the cautiousness rankings although variability in ranks was far higher than for importance rankings. Overall, health care professionals regarded the absence of pain behaviours as more cautiousness-inducing than the presence of context cues. However, due to the high variability differences between pain behaviour and context cues did not reach significance for all pairs. Again this result was also found in the stimulus space in which two groups of cues and two single cues emerged: one group contained the context cues with the exemption of 'results of medical investigations', the other group consisted of the pain behaviour groups except the 'absence of verbal report'. The cue 'no results of medical investigations' was placed between both groups, the cue 'no verbal report' a little apart from the pain behaviour group.

To summarise, although the three different concepts – importance, ease of manipulation and cautiousness – seem to differ in their complexity they all seem to be relevant for the question what might lead to pain underestimation. Ease of manipulation of pain behaviours as concept was very clear as indicated by low variability in ranks, significant differences between all cues, high agreement of participants on the rank order and a one-dimensional stimulus space. With regard to pain underestimation cues perceived to be easy to manipulate, possibly lead to a greater extent of underestimation.

Importance as concept was more complex since ease of manipulation emerges as an additional substantial underlying dimension to importance in the MDS solutions. Present pain behaviour cues are considered to be more important than context cues and differ with regard to their ease of manipulation (e.g. 'verbal report' and 'physiological indicators' are considered very important, but whereas 'verbal report' is perceived to be very easy to manipulate, 'physiological indicators' are perceived to be very difficult to manipulate). With regard to pain underestimation pain behaviours are perceived to be important cues for judges which is why their presence in judgement situations is crucial.

Cautiousness as importance was a fairly complex concept too, as indicated by high variability of and low agreement on ranks. However, it was also considered as closely related to importance, as indicated by the rank ordering and the MDS solution with 'importance' and 'ease of manipulation' as underlying dimensions. Important but absent pain behaviour cues are considered to be more cautiousness-inducing than present context cues and differ with regard to their ease of manipulation. Since absent important pain behaviours are cautiousness-inducing, their absence might lead to pain underestimation.

2. Criticism

There are several strengths as well as limitations affecting interpretation of results. The major methodological drawback certainly is that opinions of health care professionals do not necessarily tell us about their behaviour. It is therefore uncertain whether what health care professionals think about cues actually impacts on their way of judging pain in patients. Further, it is uncertain whether and to what an extent the results of this study can be generalised with regard to the sample and the cues used. Nonetheless, the sample size is fairly robust for statistical analyses and for interpretation of results. As cues were selected they could not possibly cover all cues impacting on pain judgements. However, the majority of participants did not think any cues were missing which is encouraging as this means the relevant cues were offered.

3. Outlook

Despite these limitations the ranking task provided rich information. Close relationships between the three concepts – importance, ease of manipulation and cautiousness – were found as was a distinction between groups of cues in pain behaviour and context cues. The results of the study further revealed that health care professionals consider ‘verbal report’ as important cue when judging pain which was the main research question. This result lends indirect support to the hypothesis that disagreement between health care professionals and patients could be due to verbal report as missing cue.

Study 2

Examining two accounts for underestimation of pain.

Aims and objectives

Although underestimation of pain has received much attention in the literature (Craig et al., 2001), various explanations for this tendency were not conclusive (Williams, 2002). Aim of this study was to test the two accounts for underestimation proposed in the introduction: whether underestimation can be accounted for by

1. verbal report as missing cue (Objective 1) and/or
2. a context cue making judges suspicious (Objective 2).

1. Verbal report as missing cue

When putting together the studies examining agreement, it occurred to me that in most studies judges were not allowed to talk to patients about their pain. However, the lens model emphasises the importance of cues to draw inferences and the previous study found that health care professionals considered verbal report as a very important indicator when judging pain. Therefore, the absence of verbal report as cue may be associated with pain underestimation. The hypothesis to test was that the absence of verbal report leads to underestimation.

2. Presence of a context cue making judges suspicious

Context cues hinting at secondary gains may activate the cheating detection device proposed by Cosmides and Tooby (Cosmides, 1989; Cosmides & Tooby, 1992). An activated cheating device may lead to more conservative judgements of pain. Results of the previous study cannot offer a definite answer which context factor makes judges most cautious. Yet, they have shown that obtaining opioid drugs was fairly cautiousness-inducing. Since this cue had also appeared as a suspicious making factor in another study (Kappesser & Williams, 2002), it was chosen as contextual cue for the present study. The hypothesis to test was that the presence of this cue leads to pain underestimation.

To summarise, the study's objectives were to answer the following questions in the context of pain judgements:

1. Does the unavailability of verbal report as cue lead to pain underestimation? (Objective 1)
2. Does the presence of a context cue lead to pain underestimation? (Objective 2)

Method

1. Design

To investigate these questions a study with a cross-sectional design was set up with number of cues as independent variable and pain judgement and confidence in pain judgement as dependent variables.

The number of cues presented was systematically varied across three groups. Group 1 judges had no verbal report available but had to base their pain judgements solely on facial expressions. Group 2 judges were provided with facial expression and patients' self-reported pain intensity. Group 3 judges were provided with the same cues as participants in group 2 but in addition were informed that some of the patients, whose faces they would see, were faking pain in order to obtain opioid drugs.

2. Participants

There was no need to apply for full ethical approval (see correspondence with Ethic Committee, Appendix II.1).

To my knowledge there has been no similar study in the existing literature that I could use to determine the sample size for the present study. Giving the limited number of doctors and nurses working in the hospitals' single departments, no single department within the hospitals was large enough to enable me to collect data only there. Hence, permission was sought from the Medical Directors of the two largest departments, A & E and Oncology, and granted. However, even with these two departments a sample size of only 120 volunteers altogether seemed realistic (40 participants per cell). Using a power calculation programme (Faul & Erdfelder, 1992), a power analysis for a compromise ANOVA showed that with a medium effect size of $f = 0.25$, a β/α ratio = 1 (assuming that both, α and β , are equally serious) and 3

groups a sample size of 120 would lead to an acceptable power of 0.8445, $\alpha = 0.1555$ and a critical $F(2, 117) = 1.8909$.

The sample was a convenient sample of participants who were assigned to the single conditions by chance: I started with condition 1 for the first participant, continued with condition 2 for the second participant and so on. Since participants were recruited from two specialties and professions, I kept numbers of people working for A & E and Oncology as well as numbers of doctors and nurses constant across conditions.

To explain aim and content of the study and ask for participation, posters were put up in the A & E coffee room as well as in the staff rooms on the oncology wards. Oncology doctors were contacted by letter. One example of each, poster and letter, is included in Appendix III.1.

3. Materials

A questionnaire containing the instructions and scales described below is included in Appendix III.2.

3.1 Facial expressions of pain

The videotapes with the facial expressions were very kindly provided by Prof. Kenneth Prkachin and Dr. Susan Mercer. Facial reactions of patients with shoulder pain were videotaped while they were undergoing a range of physiotherapeutic motion tests [for a full discussion of patient characteristics, tests and videotape characteristics, see Prkachin & Mercer (1989)]. In addition to the facial expressions, which were FACS-coded by Prof. Prkachin and colleagues, the pain intensity reported by the patients themselves was available.

From the available pool of videotaped patients a selection had to be made according to the following criteria. The first criterion was the limited time of volunteers which restricted study demands to no longer than 15 minutes. Although the videotaped sequences were very short, this criterion set an upper limit of eight to the number of faces shown. Since there is evidence that the patient's ethnic background and sex may impact on pain judgements (Carey & Mills Garrett, 2003; Davitz & Davitz, 1975; Hadjistavropoulos, McMurtry & Craig, 1996; Zborowski, 1952), only patients from a

Caucasian ethnic background [the majority in the study by Prkachin & Mercer (1989)] and an equal number of male and female patients were chosen. Further, the pain intensity as reported by the sufferer has been found to affect pain judgements (Chibnall & Tait, 1995; Chibnall et al., 1997; Tait & Chibnall, 1997) so for reasons of ecological validity a group of patients was selected who reported a wide range of pain intensities while undergoing the motion tests. Lastly, videotapes were selected in which only faces were shown and no additional movements of the arm, in order not to give judges additional cues of what was done to the patient.

Further characteristics of the selected four male and four female faces are summarised in Table 1.

Table 1. Summary of sex of, motion test undergone by and pain rating of patient and intensity of facial pain expression (FACS index).

face	sex	motion test (all passive)	sensory pain rating [scale by Heft, Gracely, Dubner & McGrath (1980)]	FACS index
1	male	external rotation	very intense	61.06
2	male	external rotation	slightly moderate	13.1
3	male	abduction	Strong	44.69
4	male	external rotation	very intense	3.2
5	female	internal rotation	very intense	29.3
6	female	external rotation	Intense	77.5
7	female	external rotation	slightly intense	29.3
8	female	internal rotation	extremely intense	41.4

The video sequences lasted on average 10.53 seconds (standard deviation 2.07) and volunteers were shown the eight faces in random order, randomised by a random number table.

3.1.1 Laptop

Videsequences were shown to participants using one of three portable computers with screen widths ranging between 30 and 36 cm. Volunteers were asked to adjust the screen to an angle at which they could see the screen clearly.

3.2 Scales to assess the dependent variables

In addition to pain ratings judges were also asked to rate the confidence with which they had made their pain intensity ratings. The reason for asking for participants' confidence was to examine whether the amount of agreement between ratings of patients and judges varied with the amount of confidence judges had in their ratings.

3.2.1 Scale for pain judgements

For ease of comparison judges were given the same scale to rate the patient's pain that patients who provided the videos had used: the sensory scale by Heft and colleagues (1980). In order to rate each patient's pain intensity participants were asked to 'choose the one word which best describes the maximum pain intensity shown in the video'.

The 13 adjectives and a 'no pain' option added Prkachin and Mercer (1989) were provided ordered by the increasing numerical equivalents obtained for the adjectives by Heft and colleagues (1980): no pain, extremely weak, faint, very weak, weak, very mild, mild, slightly moderate, moderate, barely strong, clear-cut, slightly intense, strong, intense, very intense and extremely intense.

For ease of use numbers from 0 to 15 were used rather than the original numerical equivalents. Higher numbers were assigned to adjectives expressing more pain (i.e. 0 to 'no pain', 1 to 'extremely weak', 2 to 'faint' and so on), thereby keeping the original order of the adjectives obtained by Heft and colleagues (1980).

3.2.2 Scale for confidence ratings

A 7-point Likert scale was used for the confidence ratings. The scale's left hand side was labelled 'not at all', its right hand side 'extremely'.

3.3 Background variables

Several background variables were assessed since they are discussed as important in the literature: base rate estimates of patients faking, exaggerating, minimising and hiding pain; profession; specialty; years participants had been working in their profession and specialty; participants' age and sex.

4. Procedure

Volunteers were approached before the beginning of, during or after the end of their shift. Aim and content of the study were explained: I said that I was interested in how people judge pain, that I had brought with me eight videotaped faces of patients and that I would like them to give me their impressions of these faces.

After the demographic variables and the variables regarding their profession were recorded, volunteers were asked to read the following introduction (text in italics for groups 2 and 3 only; text in italics and brackets group 3 only):

In the following you will see eight videotaped faces. I would like you to imagine that all of them are your patients.

You will also be told what each of them said about how they felt.

To rate the intensity of pain you are asked to choose the one word which best describes the maximum intensity shown in the video.

*As this is **your** opinion based on all available information, your rating may or may not agree with their report.*

(Be aware that when videotaped some people were faking pain to obtain opioid drugs.)

Do you have any questions before we start?

Reading the introduction was followed by watching the videotapes and rating the patients' pain intensity. Participants' two last tasks were to state whether they had had the impression that any of the patients had faked, exaggerated, minimised or hidden the pain and what they thought the base rate estimates for these four conditions were in percentage terms.

Participants had to rate the patient's pain after each video. Each self-report was offered on a single A4 sheet of paper stating that 'The patient reports the pain to be ... *(adjective from scale was inserted)*.'

After completion of the last task, participants were offered the chance to ask questions concerning content and aim of the study.

Overall, completing the study took participants approximately ten minutes.

5. Statistical analyses

The effect of the three groups (as independent variable) on quantitative variables (as dependent variables) was tested by applying univariate analyses of variance, on categorical variables (as dependent variables) by applying a Kruskal-Wallis test. To investigate the associations of all included variables with pain judgements a chain graph was chosen as it allows the display of complex associations between a number of variables. Chain graphs are expansions of path analyses which were developed by Cox and Wermuth (Cox & Wermuth, 1996; Wermuth, 1998) and are increasingly used recently (e.g. Hardt, Petrak, Filipas & Egle, 2004). Advantages of chain graphs, for instance, compared to structural equation modelling, are that they include not only linear but also non-linear associations such as quadratic terms and interactions between variables. Further, categorical and numerical variables can be modelled. And in chain graphs every missing connection between a pair of variables corresponds to an independence statement, as every present connection corresponds to a specific conditional or marginal association. This does not hold in general for structural equation models (Wermuth, 2003).

Instead of single pain ratings, differences between the estimated and self-reported pain were used as dependent variable in the chain graph.

For statistical analyses a significance level of $p \leq 0.05$ was set with the exemption of the chain graph where the level was lowered to $p \leq 0.01$ in order to enhance the probability of replicable findings.

6. Feedback

After the data collection and the statistical analyses were completed, a feedback letter with the main results was sent to the Medical Directors, Senior Nurse Managers, Ward Managers and the doctors who had especially asked to receive feedback on the results of the study. A feedback sheet is included in Appendix III.3.

Results

The result section consists of three parts: First, the sample is described. Second, results of the univariate analysis of variance are presented. Finally, the chain graph is introduced.

1. Description of the sample

One hundred and twenty doctors and nurses took part in the study, all working for one or both of two big teaching hospitals in London. Sixty of them were working for A & E and 60 for Oncology (including Haematology and Palliative Care). Although there was a difference between specialties' mean pain ratings, this difference did not reach significance (A & E: $M = 8.85$ ($SD = 1.63$); Oncology: $M = 9.43$ ($SD = 2.00$); $t = -1.74$, $p = 0.084$). Therefore, both specialties were combined for analysis. Among the 120 participants were seven medical or nursing students, five on placement in A & E and two on placement in Oncology.

Volunteers' age ranged from 21 to 59 with a mean of 32 years. As in the previous study there was an association between profession and sex: More doctors were male (58.3 %) and more nurses female (87.5 %) which is a result typical for hospitals of the National Health Service. Participants had been working in their profession between one week and 43 years with a mean of 8 years and in their specialty between one week and 28 years with a mean of 4 ½ years.

1.1 Effect of the faking instruction (group 3)

In group 3, participants were told that some patients would fake pain in order to obtain opioids. Since I wished to check whether the instruction had the desired effect, I asked all volunteers whether they thought that one or several patients had faked pain (possible answers: yes or no).

The Kruskal-Wallis test revealed significant differences between groups for pain faking ($\chi^2 = 15.16$, $p = 0.001$). Post-hoc tests (Mann-Whitney with adjusted $p = 0.01/3 = 0.0033$) disclosed that participants in group 3 were more likely to judge patients as faking than participants in group 1 (faces only; $z = -3.81$, $p < 0.001$) and in group 2 (faces and verbal report; $z = -2.445$, $p = 0.014$). Apparently, the faking instruction of group 3 had the desired effect on participants.

2. Results of comparisons between groups

2.1 Descriptive statistics

Table 2 summarises the means and standard deviations for all quantitative variables included in the study across the three groups as well as for each group separately. Furthermore, it shows the results of the oneway ANOVA applied to test for differences between groups.

2.2 Results of the ANOVA

The only quantitative variable for which significant differences between groups occurred was the difference in pain ratings (see Figure 1). A post-hoc Tukey-HSD disclosed that there was a significant difference between group 1 and 2 (mean difference $M = -1.51$; $p = 0.001$) as well as between group 2 and 3 ($M = 1.08$; $p = 0.018$), but that the difference between group 1 and 3 was not statistically significant ($M = -0.43$; $p = 0.515$).

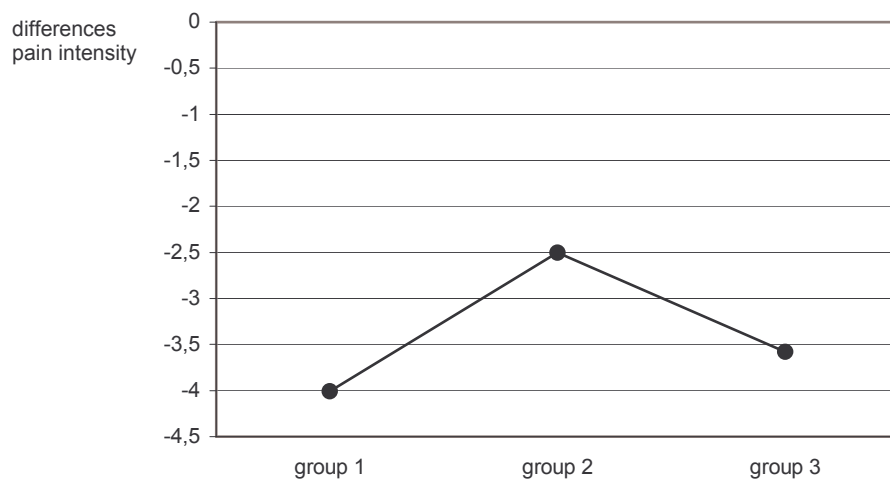


Figure 1. Differences in pain intensity ratings between groups.

Table 2. Means and standard deviations (*SD*) for all quantitative variables.

variable	overall	group 1 (face)	group 2 (face, verbal)	group 3 (face, verbal, opioid)	oneway ANOVA
	mean (<i>SD</i>)	mean (<i>SD</i>)	mean (<i>SD</i>)	mean (<i>SD</i>)	<i>F</i> -value (<i>p</i> -value)
confidence	3.63 (1.01)	3.68 (1.02)	3.85 (1.02)	3.36 (0.94)	2.55 (0.082)
difference (patient report - participant rating)	- 3.36 (1.84)	- 4.01 (1.78)	- 2.50 (1.97)	- 3.58 (1.42)	7.97 (0.001)
base rate estimate for faking (in percent)	11.80 (13.60)	11.56 (12.55)	13.58 (15.84)	10.25 (12.26)	0.60 (0.549)
base rate estimate for exaggerating (in percent)	25.15 (18.85)	24.23 (18.96)	28.69 (21.18)	22.54 (16.00)	1.14 (0.324)
base rate estimate for minimising (in percent)	36.26 (20.87)	40.36 (21.75)	32.93 (20.33)	35.50 (20.34)	1.32 (0.272)
base rate estimate for hiding (in percent)	29.19 (20.70)	34.61 (23.57)	27.40 (19.56)	25.55 (17.97)	2.18 (0.117)
age (in years)	32.21 (7.95)	32.40 (7.63)	31.33 (7.82)	32.90 (8.50)	0.41 (0.668)
years working in profession	7.73 (7.59)	8.39 (8.29)	7.39 (8.00)	7.41 (6.50)	0.22 (0.802)
years working in specialty	4.42 (4.67)	4.62 (5.14)	4.37 (4.64)	4.26 (4.33)	0.06 (0.942)

2.3 Results of the Kruskal-Wallis test

Of the three categorical variables numbers for two of them (profession and specialty) were held constant (20 % doctors; 50 % A & E in each group). However, it was not possible to control for the third qualitative variable, sex. Overall, 26 participants (21.7 %) were male, with 12 men (30 %) being in the first, 7 (17.5 %) in the second as well as in the third group. However, results of a Kruskal-Wallis test revealed that the number of men and women did not differ significantly across conditions ($\chi^2 = 2.44$, $p = 0.296$).

3. The chain graph

The chain graph was used to examine associations between all variables included in the present study. First, basic construction principles of chain graphs are described. This is followed by a description of the order of the variables included in the chain graph and the test for non-linearities. Finally, the overall model is summarised and the chain graph displayed.

Table 3. Construction principles of chain graphs.

Variables

- all variables are ordered according to logical, temporal or causal reasons so that variable(s) that are primary response variable(s) are placed on the left hand side, purely explanatory variables on the right hand side; variables between both are called intermediate variables and are response variables for all variables to their right and explanatory variables for all variables to their left
- qualitative variables are represented by dots, quantitative variables by circles

Boxes

- boxes enclose single variables or groups of variables (if variables within this group cannot be divided into response and explanatory variables)
- double framed boxes indicate that associations of variables within this box are not specified by the model

Dependencies between variables

- dependencies between variables are examined using regressions; the appropriate type of regression depends on the type of response (multiple regressions for quantitative variables; logistic regressions for categorical variables) as well as on assumptions about variables in the same boxes (see next paragraph)
- dependencies between variables are graphically represented as edges and there is at most one edge between any pair of nodes in the graph
 - edges can be directed (arrows) or undirected (lines);
 - edges can be dashed (association resulting from a multivariate regression) or full (association resulting from a block regression)

Information provided in chain graphs

- final chain graphs only provide information about direct and indirect associations between different types of variables; information about direction, strength and linearity of associations can only be obtained from the regression equation for each response variables which must be provided in addition to the chain graph
-

3.1 Construction principles of chain graphs

For readers unfamiliar with chain graphs their basic construction principles are briefly described in Table 3. (For more detailed descriptions and applications see Cox & Wermuth (1996), Wermuth (1998) or the website <http://psystat.sowi.uni-mainz.de>.)

3.2 Order of variables

Figure 2 displays the order of variables for the present study. All variables with the exception of confidence were included in the chain graph. The association between agreement and confidence is described in an additional section following the description of the chain graph.

Primary response variable was the difference between participants' and patients' pain ratings. I considered the base rate estimates as intermediate variables since they may be shaped by one or several of the background variables (age, time in profession, time in specialty, sex, profession and specialty). The background variables as purely explanatory variables were placed in the box at the right end side of the chain graph. Because the group participants were assigned to is an experimental manipulation, it is not supposed to act as a response variable for any of the other explanatory variables. Therefore, I considered it as a purely explanatory variable and placed it in the same box as the background variables. In order not to complicate analyses unnecessarily I decided not to examine associations between variables placed within one box.



Figure 2. Order of variables.

3.3 Test for non-linearities

Before commencing with examinations of dependency or independency of variables in the chain graph, data are usually screened for quadratic effects and interactions.

Cox and Wermuth (1994) developed a test in which the Student t-statistics of all quadratic effects and trivariate distributions is calculated and plotted against the expected value. Large deviations from the diagonal indicate substantial effects and should be taken into account in further analyses. The programme that runs the test is in the meantime available for free use on the web (<http://psystat.sowi.uni-mainz.de>).

No non-linear effects had to be taken into account in the search process since values deviating from the diagonal either did not fit the structure determined by the order of variables in the chain graph (e.g. years in profession = age x baserate faking) or they concerned variables within one box the associations of which were not of interest for the present study (e.g. profession = sex x time in specialty). Similarly, no quadratic effects had to be taken into account in the chain graph.

3.4 Summary of the final model

Table 4. Regression equations for all response variables in the chain graph.

difference in	R^2	b	s.e.	t	p
pain ratings					
(constant)		- 2.590			
group (dummy 1)	0.218	- 0.673	0.213	- 3.166	0.002
group (dummy 2)		0.970	0.214	4.525	< 0.001
exaggerating		- 0.031	0.008	- 3.809	< 0.001
faking					
(constant)		15.166			
time in profession	0.059	- 0.436	0.160	- 2.723	0.007
exaggerating					
(constant)		47.520			
specialty	0.180	- 11.945	3.190	- 3.744	< 0.001
time in profession		- 0.576	0.211	- 2.728	0.007
minimising					
(constant)		- 9.359			
profession	0.174	14.755	4.366	3.380	0.001
specialty		12.708	3.493	3.639	< 0.001
hiding					
(constant)		13.350			
specialty	0.066	10.558	3.669	2.878	0.005

Multiple regressions were used to search for associations between variables [for more detailed information on which type of regression to choose and a description of the underlying concepts of marginal and conditional independence see Cox & Wermuth (1996), Wermuth & Cox (1998)]. When searching for associations a significance level of $p \leq 0.01$ for main effects and of $p \leq 0.005$ for non-linear effects was chosen. These conservative levels were set to minimise the risk of putting undue weight on chance. Explanatory variables not meeting these criteria were stepwise excluded in the search process. Table 4 provides the final regression equations for each response variable. The resulting chain model is displayed in Figure 3. In Appendix III.4 correlation coefficients for all variables are given and in Appendix III.5 the regression parameters for all excluded variables for each response variable are listed.

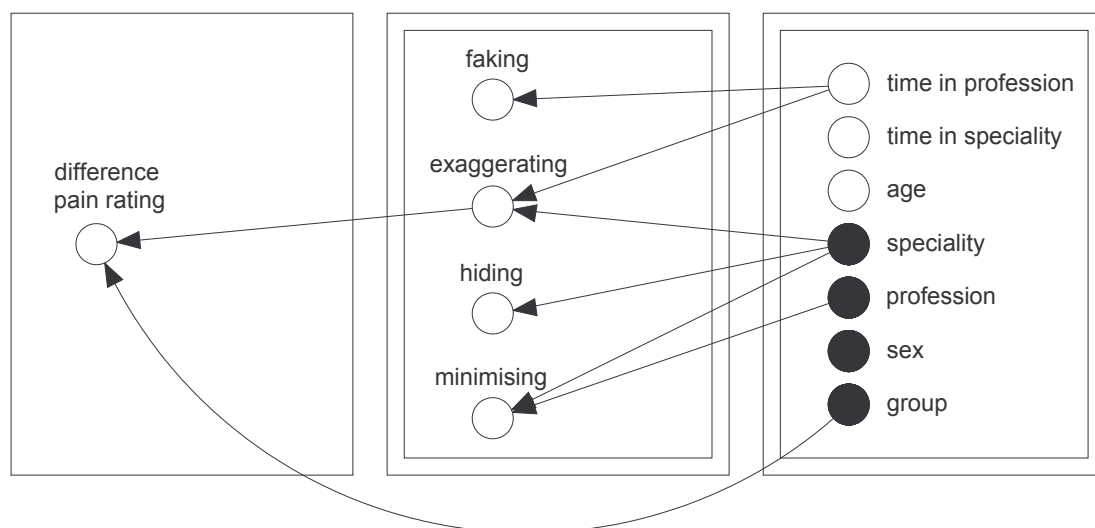


Figure 3. The resulting chain graph.

What does the chain graph tell us? It reveals that there are two explanatory variables accounting significantly for differences in pain intensity ratings: group and base rate estimates of exaggeration. Although pain was underestimated by participants in all three conditions, underestimation was greater if they had no information beyond patients' faces and greater if a context cue was present suggesting a possible secondary gain. Further, the higher the base rate estimate of exaggeration the greater the underestimation. Figure 3 displays these effects: With each percentage point increase in the base rate estimate of exaggeration, underestimation became slightly

greater. However, when comparing the three groups, underestimation was less in the group who saw the face and was given verbal report as cues.

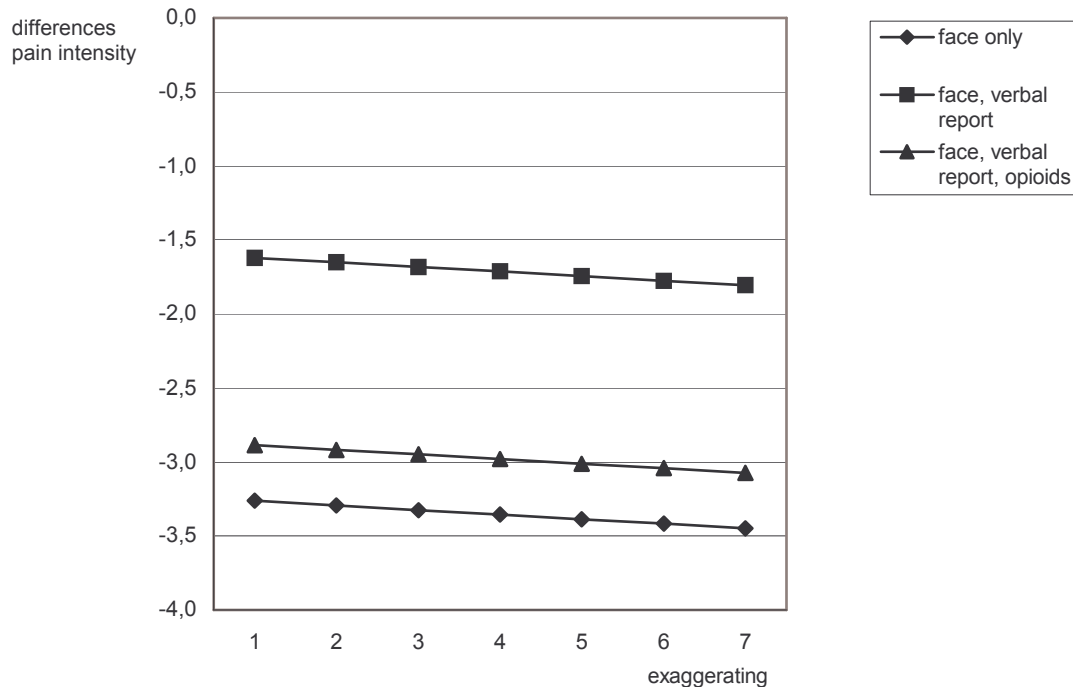


Figure 4. Effect of explanatory variables on the difference in pain ratings (judges - patients).

The chain graph further reveals that two of the background variables had an indirect effect (via the exaggeration base rate) on the difference in pain ratings: the time health care professionals had been working in their profession and the specialty they were working in. The longer health care professionals had been working in their profession, the smaller their estimate of patients exaggerating pain. And health care professionals working in A & E estimated a higher number of patients to exaggerate their pain when compared to health care professionals working for Oncology.

3.5 Association between difference in pain intensity and confidence ratings

Confidence ratings were included in the study to allow judges a statement about the certainty with which they had made their pain ratings. It was interesting that mean confidence varied considerably between groups (highest for group 2 and lowest for group 3). However, differences did not reach significance (see Table 2).

Of further interest was to investigate which variables could predict participants' confidence ratings in general. Multiple regressions were used for these investigations after screening for non-linear effects. Appendix III.6 gives the correlations between all variables and Appendix III.7 lists regression parameters for all excluded variables.

Table 5. Regression equation for confidence as response variable.

confidence	R^2	b	s.e.	t	p
(constant)		5.126			
difference	0.161	0.451	0.132	3.420	0.001
difference ²		0.064	0.018	3.605	< 0.001
specialty		-0.614	0.175	-3.517	0.001

When examining the non-linear effects, one value was deviating indicating a quadratic effect of pain difference on confidence. Consequently, this quadratic effect was included in the search process of explanatory variables for confidence. The result of the search process is summarised in Table 5.

Two significant predictors for confidence emerged: differences in pain intensity ratings and specialty. Health care professionals working in A & E felt more confident with their pain intensity ratings than those working in Oncology. The association between differences in pain ratings and confidence, displayed in Figure 5, is quadratic. The quadratic effect, however, could not be accounted for by specialty, group or base rate estimates of faking or exaggerating.

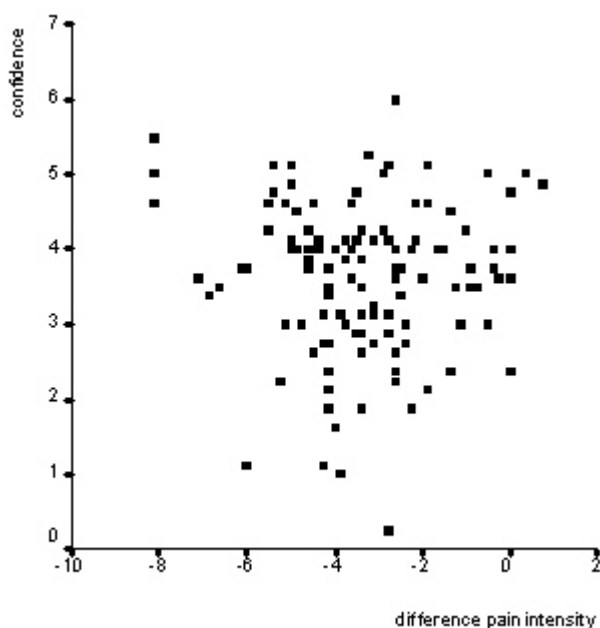


Figure 5. Quadratic effect of difference in pain intensity ratings on confidence.

Discussion

1. Discussion of results

Aim of this study was to test the two accounts for underestimation: whether underestimation can be accounted for by verbal report as missing cue (Objective 1) and/or a context cue making judges suspicious (Objective 2). Additionally, the impact of several background variables on underestimation was investigated.

1.1 Accounts of underestimation

Discrepancies between health care professionals' and patients' pain ratings occurred in all three groups with health care professionals rating patients' pain intensity as lower than patients themselves did. The mean difference between health care professionals and patients was -3.36 units ($SD = 1.84$) with a range from -8.13 to 0.75 . This is a considerable difference far exceeding the lafrati criterion and the clinically significant changes as found by Bird & Dickson (2001), Gallagher et al. (2001), Kelly (1998) and Todd et al. (1996).

Significant differences emerged between groups depending on the number and kind of cues present in the judgement situation. The least discrepancy between judges' and patients' ratings occurred when judges saw patients' faces and were given their verbal reports. The discrepancy between judges' and patients' ratings was greater when a context factor, possibly alerting a cheating detection device, was present in addition to facial expressions and verbal report. Greatest discrepancies occurred when judges had to base their ratings on facial expressions only. Consequently, results lent support to both accounts for underestimation: the unavailability of verbal report as cue led to greater discrepancy between ratings as did the presence of a context cue alerting the cheating detection device.

1.2 Effect of additional variables on underestimation

The chain graph showed that three additional variables, base rate estimate of exaggeration, working time in profession and specialty, impacted on the difference between patients' and judges' ratings. Exaggeration base rate was directly associated with the difference in pain ratings: Health care professionals who believed more patients to exaggerate their pain differed to a greater extent from patients' ratings than health care professionals who believed few patients to exaggerate their pain. Since

exaggeration can be considered to be part of deception (as defined by the American Psychiatric Association (1994, p. 683), results of this study confirm that judges' base rate estimates of deception impact on their pain ratings. At first glance, it is interesting that this finding concerns only the base rate estimate of exaggeration, not the ones of faking, minimising or hiding. However, when looking at the correlation coefficients as well as at the regression coefficients, it becomes apparent that the base rate estimate of minimising has a considerable impact on the difference in pain ratings too ($p = 0.044$; health care professionals who thought more patients to minimise their pain differed to a lesser extent from patients' ratings). Further, base rate estimates for faking and exaggeration are highly correlated ($r = 0.49$, $p \leq 0.01$) as are base rate estimates for hiding and minimising ($r = 0.76$, $p \leq 0.01$).

The chain graph also revealed that base rate estimates were shaped by other variables. Exaggeration base rate was affected by time doctors and nurses had been working in their professions as well as the specialty they were working in. Therefore, these two variables had an indirect effect on differences in pain ratings. The longer participants had been working in their profession the less they expected patients to exaggerate their pain. As a result, they were also less likely to underestimate patients' pain intensity. In contrast to the hypothesis of developed insensitivity (according to which more experienced judges should differ more greatly in their ratings due to their developed insensitivity) the present study found that less experienced judges estimated more patients to exaggerate pain and, consequently, differ from patients' ratings to a greater extent than more experienced judges whose estimates of exaggeration were smaller. Possibly the effect of experience on pain judgements and pain underestimation that has been found repeatedly in the literature is due to moderating variables such as expectations that have been shaped over time by the context in which health care professionals work.

Specialty had an indirect effect on differences in pain intensity too. Health care professionals working in A & E estimated more patients to exaggerate their pain and, as a consequence, were more likely to underestimate patients' pain than health care professionals working in Oncology. Possibly, these differences are due to health care professionals working for A & E seeing their patients mostly on one occasion only so they can hardly ever confirm or disconfirm their judgements. Doctors and nurses

working for oncology, on the other hand, mostly see patients over a series of encounters and, therefore, have the chance to get to know their patients. Another possible account for the differences between specialties is that doctors and nurses in A & E have to deal with finding medical causes for the pain patients are reporting. For this reason they might consider the chances of patients exaggerating pain as higher. Health care professionals in Oncology, on the contrary, have to deal with treating the pain caused by diagnosed cancer and might, therefore, estimate the number of patients exaggerating their pain as smaller.

Using such a statistical model as the chain graph is helpful not only to discover direct and indirect associations but in the case of indirect associations also to explain **how** (via what paths) explanatory variables impact on response variables of interest.

1.3 Association between agreement and confidence

Although mean confidence varied considerably across groups (participants in the faking instruction felt least confident with and participants who saw the faces and were given verbal report felt most confident with their pain ratings), these differences did not reach significance ($p = 0.082$), possibly due to the restricted sample size.

Two predictors, specialty and difference in pain intensity ratings, accounted for confidence. Doctors and nurses working in A & E felt more confident with their pain ratings than health care professionals working in Oncology. The quadratic association between agreement and confidence is puzzling as it could not be explained by either specialty, group or base rate estimates of faking or exaggeration. Apparently, there was one group of judges in the sample who felt confident with their ratings no matter how intense they rated the patients' pain to be. Another group of judges felt more confident the smaller the difference between their own and the patients' ratings was. Possibly this is a trait inherent in judges. However, the effect needs replication before too much emphasis is put on its interpretation.

2. Criticism

There are, of course, several limitations affecting interpretation of results. It is uncertain to what an extent the results can be generalised with regard to the sample and the videotaped faces used. However, despite the sample being limited to employees

working in one or both of two big London hospitals, the sample size is fairly robust for statistical analyses (as also indicated by the a-priori power analysis) and interpretation of results. Further, there is no reason why London doctors and nurses should differ from doctors and nurses working in other hospitals. Yet, it is questionable whether results from this study could be generalised to different settings such as health care professionals working in community settings. Although the number of faces shown to participants was limited to eight, this number is substantial and more faces would have taken too much of participants' time. Further, faces had been FACS analysed to ensure that they showed facial muscle movements prototypical for pain.

With regard to cues present in the judgement situation, results stress the necessity to control for cues. As this is nearly impossible in clinical situations I used videotaped facial expressions of patients undergoing a painful procedure as non-verbal cues and restricted the number of cues present in the experimental judgement situation. The limited number of cues is the reason for some participants' comments that the situation was not very realistic. The use of videos and the restricted number of cues should be considered when generalising results to natural settings. On the other hand, in order to start investigating the effect cues have on pain judgements it is necessary to limit information. Furthermore, the use of videos rather than, for example, photographs definitely makes the situation more realistic and is a step towards increasing the ecological validity of the experimental situation.

The alerting of a cheating detection device is difficult to test empirically. Results showed that the presence of a context cue which made health care professionals more cautious when judging pain (previous study) led to a greater amount of underestimation. However, it is only still a theoretical assumption that this happens because of an alerted cheating detection device. Yet, this limitation is true not only for the present study but - to my knowledge - also for all other studies investigating the cheating detection device since the majority of them used the Wason selection task as means to investigate the cheating detection device and deduced its presence from the effects different instructions had on whether participants solved the tasks appropriately.

The sensory verbal descriptor scale was used to let health care professionals rate patients' pain intensity. The order of the adjectives was determined by numerical equivalents identified by cross-modality matching. Cross-modality matching requires each person to indicate the severity represented by each adjective in reference to one or several other modalities (e.g. hearing – loudness of a tone, visual – length of a line). The sum of these ratings creates the adjective's numerical equivalent. Whether these numerical equivalents generated in the original sample can be generalised to other populations is arguable. However, reliability and validity of these scales were proven. Further, to enable comparisons between judges' and patients' pain intensity ratings, judges had to be given the same scale as patients who provided the videos.

The present data also allowed analysing effects on judges' pain ratings (e.g. patients' facial expression, patients' verbal report, patients' sex). However, I decided not to report on them in this context since the main focus of this study was to investigate influences on the difference between patients' and judges' ratings rather than influences on judges' ratings.

3. Outlook

Despite these limitations results of the study showed that the kind of cues present or absent in the judgement situation, expectations judges have about the likelihood of patients exaggerating pain, the length of time since judges have been working in their profession as well as the specialty they work in affect pain underestimation. Whereas the experience of judges (time working in their profession) and base rate estimates have already been discussed as important explanatory variables in the pain or judgement literature, the results of the present study show that also the kind of cues present and the situation judges work in (specialty) have an effect on the pain ratings. Further studies are needed to find out more about which cues and in what way they affect pain ratings, an approach very commonly used in the research on judgement and decision making but not yet in the pain field. Likewise more studies are needed to find out how situations judges work in shape their impressions of patients and how this affects their pain ratings.

Study 3

Which cues affect pain judgements of patients' relatives?

Aims and objectives

Factors impacting on pain judgements of health care professionals have often been investigated (e.g. Chibnall & Tait, 1999; Chibnall et al., 1997; Tait & Chibnall, 1997). However, not very much is known about factors that might have an effect on pain judgements of lay people who are close to chronic pain patients. In the four studies (Cremeans-Smith et al., 2003; Madison & Wilkie, 1995; Miaskowski et al., 1997; Riemsma et al., 2000) which compare judgements of relatives with that of patients, relatives were found to be more likely to over- rather than to underestimate or agree with patients' pain. Therefore, the aim of this study is to investigate what factors possibly impact on pain judgements of chronic pain patients' relatives and in what way they do so.

1. Choice of cues

In previous studies two factors have been found consistently to affect health care professionals' pain judgements: presence or absence of medical findings related to the pain and the intensity of pain as experienced by the sufferer. No study so far has investigated whether these two variables have similar effects on pain as rated by relatives of chronic pain patients.

Two additional independent variables were considered due to the extensive literature about suspicion of others' pain (e.g. Craig et al., 1999; Dworkin, Handlin, Richlin, Brand & Vannucci, 1985; Fishbain et al., 1999; Fishbain et al., 1995; Main & Spanswick, 1995; Mendelson, 1992b; Pilowsky, 1985): 'continuation vs. discontinuation of tasks pleasant to the pain patient (pleasant tasks)' in combination with 'continuation vs. discontinuation of tasks unpleasant to the pain patient (unpleasant tasks)'.

Ratings of how fair the patient's behaviour is by reference to the patient's spouse were included as an additional dependent variable. The reason for its inclusion was to examine whether the manipulation of the variables 'pleasant tasks' and 'unpleasant tasks' had the intended effect. From the literature about suspicion of other's pain

and from social contract theory I assumed that behaviours seeming unfair in the sense of taking advantage of one's pain could lead to low pain judgements according to the concept of an alerted cheating detection mechanism. More precisely, the combination 'continuation of pleasant and discontinuation of unpleasant tasks' was assumed to be rated as most unfair when compared to the other three possible combinations.

To summarise, the objectives addressed by the present study were as follows:

1. Does the combination of continuation of tasks the pain sufferer likes with discontinuation of tasks the pain sufferer dislikes affect the judgements of fairness of that behaviour and of pain intensity by people who are close to a chronic pain sufferer? (Objective 1)

The null hypothesis with regard to fairness of behaviour ratings would be that neither of the two variables 'pleasant tasks' and 'unpleasant tasks' nor their combination affects fairness ratings. Alternatively, the combination 'continuation of pleasant tasks' with 'discontinuation of unpleasant tasks' leads to lower fairness ratings.

The null hypothesis with regard to pain intensity ratings would be that neither the variable 'continuation of pleasant tasks' nor the variable 'continuation of unpleasant tasks' nor their combination has an effect on judges' pain rating. The alternative hypothesis would be that the combination 'continuation of pleasant tasks' with 'discontinuation of unpleasant tasks' is a behaviour that seems unfair in the sense of taking advantage of one's pain (tested by fairness of behaviour ratings) and, therefore, leads to lower pain judgements.

2. Do the results of medical investigations (cause for pain present vs. no apparent cause) have an impact on pain severity as rated by people who are close to a chronic pain sufferer? (Objective 2)

The null hypothesis would be that positive results do not have an effect on judges' ratings. The alternative hypothesis proposed here is that positive results do have an effect on judgements by lay people. Similar to the finding in students, medical students, nurses and physicians, I assume that the absence of positive medical findings leads relatives to underestimate pain.

3. Does the pain intensity of a sufferer have an impact on the pain intensity as rated by people who are close to a chronic pain sufferer? (Objective 3)

The null hypothesis would be that the pain intensity of the sufferer would have no impact on the rating of the judge. Extrapolating from the results of studies with health care professionals suggests that, alternatively, higher pain intensity of the sufferer would lead to underestimation. However, since relatives seemed to differ from health care professionals with regard to over- and underestimation of pain, I argue that the pain intensity of the sufferer has a positive effect: the higher the pain intensity of the sufferer, the higher the rating of the relative.

Method

1. Design

The study was set up in a 2 x 2 x (2 x 2) mixed factorial design. Two of the independent variables were between-factors (positive findings of medical investigations with two levels: present vs. absent; and self-reported pain intensity of patient with two levels: low vs. high), the other two independent variables were within-factors (pleasant tasks with two levels: continuation vs. discontinuation; and unpleasant tasks with two levels: continuation vs. discontinuation). Dependent variables were ratings of pain intensity and fairness of behaviour.

1.1 Independent variables

Independent variables were presented in vignettes, i.e. brief written case histories. Each volunteer was asked to read four vignettes altogether, each of which contained one of four possible combinations of the two within-factors. As between-factors, self-reported pain level and supporting results of medical investigations were kept constant across all four vignettes for each volunteer.

In an attempt to verify that volunteers had accurate grasp of the independent variables, a short test of participants' knowledge of information given in each vignette was added. After reading each vignette, volunteers were asked to answer four true-false questions two of which referred to the two manipulated independent variables. The other two questions referred to information that was held constant across the

vignettes. Participants were asked to re-read the vignette if they were unsure about a question.

1.2 The use of vignettes

Using vignettes has advantages as well as disadvantages. The most essential disadvantage of vignettes probably is the question of their ecological validity, i.e., their power to represent real situations and to predict behaviour. Among their advantages are easy administration, control for case mix, possibility to use them universally and their comparatively low costs. Further, in a recent study Peabody and colleagues (Peabody, Luck, Glassman, Dresselhaus & Lee, 2000) compared the validity of vignettes, chart abstraction and standardised patient visits (where one of the investigators acts as a patient, presenting the same medical complaints in the same way to different providers and directly observes providers' reactions; methodologically, standardised patient visits are commonly considered the gold standard). They showed that vignettes were better in predicting the process of care than chart reviews.

Despite the question of ecological validity and for reasons of practicality I decided to choose vignettes as a means to investigate the questions of interest. There were two reasons of practicality: First, my target group was relatives and friends of chronic pain patients who had been treated in a pain management programme. As these patients come from a large area (some of them even from Northern England and Scotland), the costs of making personal contact with relatives and friends of these patients and bringing them together with patients for observation was prohibitive. In contrast, inviting relatives and friends via mail to fill in a questionnaire containing vignettes did seem feasible and possible. Further, the aim of this study was to investigate what factors possibly impact on pain judgements in chronic pain patients' relatives. For this purpose a vignette study seemed to be an appropriate way.

1.3 Background variables

Participants were asked to

- describe themselves

(age, sex, current health status, how much time on average they spend with the pain sufferer, base rate estimate of chronic pain sufferers in the population,

whether pain sufferers in general tend to hide or express their pain, whether they took over doing some tasks for the pain sufferer and if so how fair this is),

- describe the pain sufferer they are close to
(age, sex, onset of pain, whether the sufferer tends to hide or express his/her pain, whether the sufferer has stopped tasks he/she likes or dislikes because of the pain) and
- describe their relationship to the pain sufferer
(how they are related to the sufferer, whether they share home with the sufferer, how satisfied they are with the relation to the sufferer, whether the relationship to the sufferer has changed because of the pain, how long they have known the sufferer, whether the sufferer tries to make up for what the volunteer does for him/her and whether the sufferer would do the same for the volunteer if he/she had chronic pain).

All background variables were controlled for by including them as covariates in the statistical analyses.

2. Participants

Relatives and friends of chronic pain patients were targeted as participants for this study. They were approached via patients who had been treated for their chronic pain in a cognitive-behavioural pain management programme in a big teaching hospital in London.

After initial enquiry, the Ethics Committee established that full application was required and granted approval. The application form containing the consent form for participants included in Appendix IV.1, succeeding correspondence with the Ethics Committee in Appendix IV.2.

Effects reported in the vignette study by Chibnall and Tait (1995) were used for power calculation. As all significant effects in this study were of large size, it was calculated that a minimum of 30 participants in each cell was needed for equivalent effect sizes with power approximating 0.70. Therefore, I aimed for a total of at least 120 participants (4 x 30).

Possible volunteers were always contacted via chronic pain patients who because of their pain had been treated in the pain management programme. At the beginning of data collection the addresses of all chronic pain patients who had been treated in the programme up to nine months ago were collected (this included patients discharged from the programme between 19th of October 2001 until 12th of July 2002). The nine months criterion was chosen as former patients after their discharge were contacted at regular time intervals for follow-ups at which treatment outcome was assessed. The last time patients are contacted is nine months after their discharge of their programme. Altogether, this process resulted in addresses of 282 former patients. Of those, one had died in the mean time, three had never replied to any of the follow-up invitations and 13 had dropped out of the programme. With the exception of these 17 patients a letter explaining aims and contents of the study was sent to the remaining 265 patients.

The letter, of which there is a copy in Appendix IV.3, stated that I needed the help of former patients with a study aiming to improve our understanding of how relatives and friends of pain patients judge how bad the pain is. I asked patients to give the questionnaire with the vignettes I had attached to the letter to the person who was closest to them and who they felt knew most about their pain. I further gave them my phone number in the hospital to allow them to contact me in case of any questions.

Of these questionnaires not enough were returned to meet the total of 120 participants set by the power calculation. Therefore, I additionally started to hand out letters and questionnaires to patients volunteering to take them with them at home. The letters and questionnaires were handed out in the last psychology teaching session before patients were discharged. Some patients volunteered to take more than one letter and questionnaire with them as they had more than one close relative or friend.

Altogether, I gave and sent out 552 questionnaires of which 181 (33 %) were completed and returned by the beginning of September 2003. The period of patients' discharges spanned 22 months.

3. Materials

A questionnaire including all materials discussed below is enclosed in Appendix IV.4.

3.1 Vignettes

The following vignette was created and discussed with experts in the pain management unit (text in italics is the manipulation of the independent variables).

[A/B/C/D] is the spouse of a friend. You have known both of them very well for years. About three years ago [A/B/C/D] started complaining about low back pain. Because of the pain [A/B/C/D] underwent several medical investigations including a CT scan,

- *all of which showed essentially a normal back. None of the doctors could suggest a definite cause for [A/B/C/D]'s pain.*
- *which together showed a narrowing of the spine with pressure on the nerve. The doctors said that this spinal stenosis was the cause for [A/B/C/D]'s pain.*

When you talked to your friend recently about the consequences of [A/B/C/D]'s pain on the couple's life, you learned that since it started

- *things have not really changed.*
- *some but not all things have changed.*
- *things have changed quite a lot.*

The couple had originally agreed to share the household chores between them. Among [A/B/C/D]'s jobs were things like mowing the lawn, ironing, decorating and cooking. Now

- *because of the pain [A/B/C/D] tends not to do the things which you know [A/B/C/D] dislikes (such as mowing the lawn and ironing). However, despite the pain [A/B/C/D] keeps doing the things which you know [A/B/C/D] likes (such as decorating and cooking).*
- *because of the pain [A/B/C/D] tends not to do things which you know [A/B/C/D] dislikes (such as mowing the lawn and ironing). Also because of the pain [A/B/C/D] has stopped doing things which you know [A/B/C/D] likes (such as decorating and cooking).*
- *despite the pain [A/B/C/D] keeps doing things which you know [A/B/C/D] dislikes (such as mowing the lawn and ironing). Also despite the pain [A/B/C/D] keeps doing things which you know [A/B/C/D] likes (such as decorating and cooking).*

- *despite the pain [A/B/C/D] keeps doing things which you know [A/B/C/D] dislikes (such as mowing the lawn and ironing). However, because of the pain [A/B/C/D] has stopped doing things which you know [A/B/C/D] likes (such as decorating and cooking).*

On a scale from 'no pain' (0) to 'pain as bad as it can be' (10) [A/B/C/D] rates the pain as 3 / 7 on average.

3.2 Scales to assess the dependent variables

3.2.1 Scale for pain intensity ratings

As in the vignette, volunteers were given a numerical rating scale from 0 to 10 with 0 meaning 'no pain' and 10 'pain as bad as it can be' to rate the pain intensity. Numerical rating scales are commonly used to assess pain intensity and have been shown to have satisfactory properties (Jensen & Karoly, 2001).

3.2.2 Scale for fairness ratings

Fairness was assessed on a 7-point numerical rating scale with 0 being 'very unfair' and 6 being 'very fair'.

4. Procedure

On the first page of the questionnaire its contents were explained. On the second page demographic background variables regarding the relative/friend and the patient in pain were assessed. Pages 3-6 contained the vignettes, the true-false questions and the rating scales assessing the dependent variables. On page 7 volunteers were asked questions concerned with their experience of being close to somebody who has chronic pain. On the very last page of the questionnaire there was space for comments on the questionnaire or other comments participants might have.

For the background variables volunteers were asked to answer the following questions or to tick the box for the answer that seemed most suitable.

The vignettes were introduced with the following words:

'I would like you to read four short stories now, each describing a person with chronic pain. For each story please assume that the facts and circum-

stances described in the story have actually occurred. After you read each story, you will be asked to answer some questions about the story you read and make judgements about the person's behaviour and pain. I would like you to tell me how much pain **you** think the person in the story is experiencing, given what you know from the story. Please read each of the following stories carefully first before answering the questions below.'

Since there was a self-reported pain intensity in the vignette, it was emphasised in the instruction that the volunteer was asked for his or her personal opinion of the patient's pain, based on everything described in the story and that, therefore, volunteers' ratings may or may not agree with the self-reported pain rating.

5. Statistical analyses

In addition to descriptive statistics, analyses of variance were chosen to examine the impact of the four independent (medical evidence, pleasant tasks, unpleasant tasks, patient's pain rating) on the two dependent variables (fairness of patient's behaviour, pain intensity). Background variables discussed as relevant in the literature were included as covariates in the analyses of variances. Further, one-sample t-tests were applied to investigate differences between pain intensity as reported in the vignettes and as rated by participants.

Independent of the kind of analyses, effects were considered as significant if $\alpha \leq 0.01$. This comparatively strict level of significance was set to increase the likelihood to replicate significant effects found in this study.

6. Feedback

In the space left for comments, one participant expressed her wish to be informed about the results of this study. She was sent a feedback sheet which was also sent via email to staff at the pain management unit who very kindly supported me while I was collecting data. A copy of this feedback sheet is provided in Appendix IV.5.

Results

After the sample and the results of the questionnaires are described, results of the analyses of variance and the one-sample t-tests are outlined.

1. Description of the sample

Of the returned questionnaires only those for which all control questions were answered correctly and for which there was a pain intensity rating for each vignette were included in the analyses. This resulted in a final sample size of 127.

Participants were on average 44 years old ($SD = 13$) with the youngest being 18 and the eldest being 73 years old. Nearly half of volunteers were male (46.5 %) and about two thirds (67 %) were spouses or partners of chronic pain patients. Of the remaining third 12 % were friends, 9 % parents, 7 % children and 4 % siblings. More than 75 % of relatives and friends shared their home with the chronic pain patient. Of the participants 16 % suffered from chronic pain themselves and 9 % from other chronic illnesses. Volunteers had known the chronic pain patient for an average of 21 years ($SD = 13$, $range = 1-54$). About one third of these chronic pain patients were male, they were on average 45 years old ($SD = 11$) with a range of 20 to 68 years and had been suffering from chronic pain for about 9 years ($SD = 6$).

2. Descriptive statistics of quantitative and qualitative variables

Table 1 describes the qualitative, Table 2 the quantitative variables assessed in the study.

Table 1. Description of qualitative variables in the present study: percentage, number of participants who answered and possible range on scale.

background variables	percentage	<i>n</i>	scale
things patient likes	47.2 % continues 47.2 % has stopped	123	continues, has stopped, both
things patient dislikes	59.8 % continues 35.4 % has stopped	123	continues, has stopped, both
did relative take on things?	64.4 % yes	111	yes, no

Table 2. Description of quantitative variables in the present study: means, standard deviations (SD), ranges, number of participants who answered and possible range on scale.

variable	mean (SD)	range	n	scale	extremes of scale
dependent variables in vignette					
mean pain intensity	5.27 (1.39)	2.75 - 8	127	0 - 10	0: no pain 10: pain as bad as it can be
mean fairness rating	3.84 (0.95)	1 - 6	123	0 - 6	0: very unfair 6: very fair
background variables					
time spent with patient on weekday	2.96 (1.06)	1 - 5	124	1 - 5	1: no time 5: all day
time spent with patient on weekend	3.77 (1.05)	1 - 5	124	1 - 5	1: no time 5: all day
estimate of general occurrence of chronic pain (in %)	21.99 (17.30)	1 - 80	118	0 - 100	
change in relationship to patient due to pain?	- 0.22 (0.83)	-2 - +2	126	-2 - +2	-2: has worsened a lot +2: has improved a lot
do patients try to make up for things participants do for them?	3.88 (1.90)	0 - 6	120	0 - 6	0: not at all 6: very much
would patients do same if participants had chronic pain?	5.08 (1.53)	0 - 6	123	0 - 6	0: very likely 6: very unlikely
satisfaction with relationship to patient	4.63 (1.53)	0 - 6	126	0 - 6	0: not at all satisfied 6: highly satisfied
expression of pain in patient: hiding vs. complaining	2.36 (1.61)	0 - 6	126	0 - 6	0: hide pain 6: make big deal/complain strongly about pain
expression of pain in chronic pain patients: hiding vs. complaining	2.89 (1.14)	0 - 6	119	0 - 6	0: hide pain 6: make big deal/complain strongly about pain

3. Analyses of variance

3.1 Fairness of behaviour as dependent variable

Fairness of behaviour was included as dependent variable to examine whether the manipulation of the two within-factors, pleasant and unpleasant tasks, had the intended effect on fairness.

A significant main effect for unpleasant tasks ($F(1, 122) = 47.132; p < 0.001; \eta^2 = 0.279$) and a significant interaction of pleasant and unpleasant tasks ($F(1, 119) = 51.085; p < 0.001; \eta^2 = 0.295$) emerged for fairness. As shown in Figure 1 behaviour of patients was generally rated as less fair had they stopped rather than continued with unpleasant tasks. Additionally, behaviour of patients was rated as fairest when they continued with pleasant as well as unpleasant tasks. Behaviour of patients continuing pleasant but stopping unpleasant tasks, however, was rated as least fair. Fairness ratings for patients who stopped pleasant tasks no matter whether they stopped or continued with unpleasant ones were fairly similar and placed in between the two extremes.

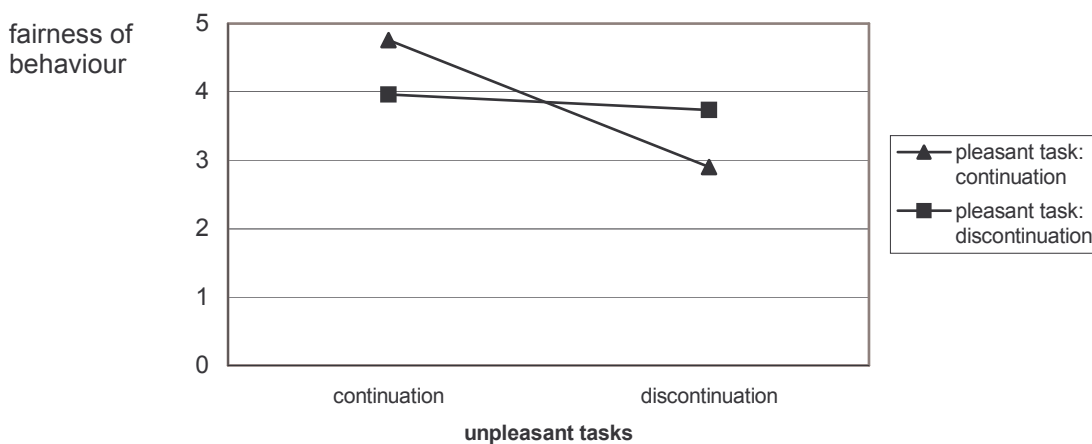


Figure 1. Interaction between pleasant and unpleasant tasks on fairness of behaviour ratings.

Since the combination 'continuation of pleasant tasks' with 'discontinuation of unpleasant tasks' did - as hypothesised - lead to lower fairness ratings, both independent variables were included in the following analysis of variance for pain intensity as dependent variables.

3.2 Pain intensity ratings as dependent variable

3.2.1 Main effects

As shown in Table 3, there were significant main effects on pain intensity ratings of pain intensity (as reported by the patient) and of pleasant tasks. Relatives rated patients' pain intensities higher when patients reported high rather than low pain intensities. Further, pain of patients continuing with pleasant tasks was rated as less intense than pain of patients stopping pleasant tasks. The effect for medical evidence, although considerable ($p = 0.053$), did not reach significance, neither did the effect of unpleasant tasks.

Table 3. Main effects on ratings of pain intensity.

independent variable	dependent variable			
	pain intensity ¹			
	mean (SD)	F	p-value ²	η^2
reported pain intensity				
low	4.44 (0.14)	69.74	< 0.001	0.36
high	6.10 (0.14)			
medical evidence				
absent	5.07 (0.14)	3.81	0.053	0.03
present	5.46 (0.14)			
pleasant tasks				
continuation	4.45 (0.12)	110.27	< 0.001	0.47
discontinuation	6.09 (0.13)			
unpleasant tasks				
continuation	5.20 (0.12)	1.02	0.314	0.01
discontinuation	5.33 (0.12)			

¹ range: 0-10; degrees of freedom: 1;123

² bold if significant

3.2.2 Interactions

No significant three- or four-way interactions were found. However, a significant two-way interaction of the two within-factors (pleasant x unpleasant tasks) emerged ($F(1, 123) = 14.51$; $p < 0.001$; $\eta^2 = 0.11$). The interaction is displayed in Figure 2 and shows that in general pain intensity ratings were higher for patients who stopped doing pleasant tasks. Pain of patients who stopped unpleasant tasks was rated as higher if they also stopped rather than continued with unpleasant tasks. Pain of patients who continued with pleasant tasks was rated as more intense if they also continued (rather than stopped) with unpleasant tasks.

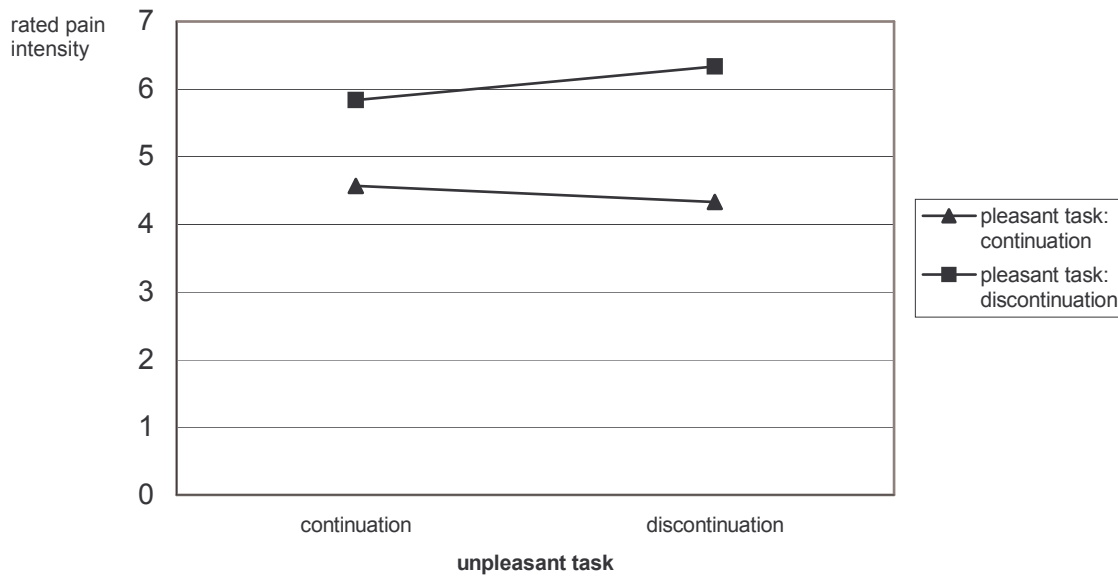


Figure 2. Interaction of pleasant and unpleasant tasks on pain intensity ratings.

3.3 Differences between reported and rated pain intensities

Overall the mean difference between reported and rated pain intensity was 0.29 ($SD = 1.64$, range = -3.5 - 4). The distribution of values is displayed in Figure 3.

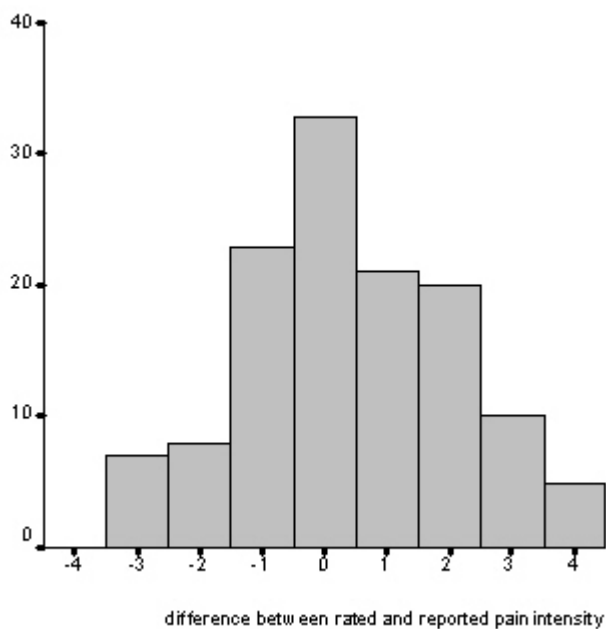


Figure 4. Distribution of differences between rated and reported pain intensity (differences calculated by subtracting reported from rated pain intensity for each participant).

One-sample t-tests revealed that despite the main effect of reported on rated pain intensity (the higher the reported the higher the rated pain intensity) in the ANOVA,

rated pain intensities differed significantly from the reported pain intensity: When the reported pain intensity was low (3), participants rated the patient's pain significantly higher ($mean = 4.46$, $SD = 1.15$; $t = 10.13$, $p < 0.001$). When the reported pain intensity was high (7), participants rated the patient's pain significantly lower ($mean = 6.10$, $SD = 1.11$; $t = -6.48$, $p < 0.001$).

Table 4 shows the mean rated pain intensities sorted by reported pain intensity, pleasant and unpleasant tasks since these were the predictor variables identified as important in the analysis of variance. One-sample t-tests revealed that three out of four means for low reported pain intensity were significantly higher than the one reported in the vignette, the fourth one (continuation of unpleasant, discontinuation of pleasant tasks) approaching significance. In the high reported pain intensity group only two out of four means differed significantly from the pain intensity as reported in the vignette. Pain was rated as significantly less intense than reported in the vignette for patients who continued with unpleasant tasks.

Table 4 One-sample t-test results for reported pain intensity, pleasant and unpleasant tasks.

	low pain intensity (3) ¹				high pain intensity (7) ²			
	pleasant tasks				pleasant tasks			
	continuation		discontinuation		continuation		discontinuation	
unpleasant tasks	mean (SD)	t-value p-value ³	mean (SD)	t-value p-value ³	mean (SD)	t-value p-value ³	mean (SD)	t-value p-value ³
continuation	3.73 (1.46)	3.98 < 0.001	3.45 (1.47)	2.47 0.016	5.43 (1.97)	-6.35 < 0.001	5.22 (1.79)	-7.88 < 0.001
discontinuation	5.02 (1.64)	9.85 < 0.001	5.64 (2.07)	10.23 < 0.001	6.68 (1.43)	-1.76 0.084	7.05 (1.59)	0.24 0.813

¹ $n = 64$

² $n = 63$

³ bold if significant

3.4 Examination of covariation

3.4.1 Impact of one dependent variable on the other

Since mean pain intensity and mean fairness ratings over conditions were moderately correlated (Pearson's $r = 0.243$, $p = 0.007$), the effect of fairness on pain intensity was examined by including fairness as covariate in the analysis of variance.

However, none of the possible interactions nor any of the main effects reached significance.

3.4.2 Impact of the background variables

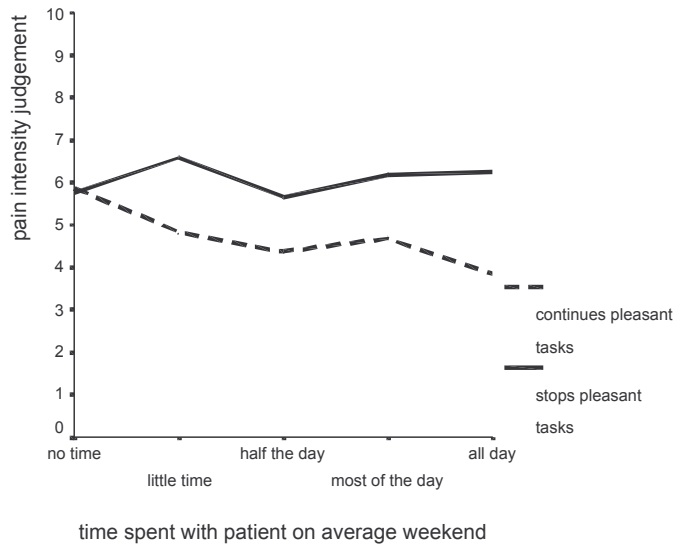


Figure 4. Interaction between pleasant tasks and time spent with pain patient at an average weekend.

Also the effects of background variables on pain judgements were examined by including each background variable as covariate in the analyses of variance. Of all background variables only the time participants spent with their pain patients on an average weekend interacts significantly with pleasant tasks ($F(1, 119) = 6.76$; $p = 0.01$; $\eta^2 = 0.05$) as shown in Figure 4. The more time volunteers spent with the patients they know at weekends, the lower their pain judgements for patients in the vignette who continued with pleasant tasks. However, pain judgements are comparatively consistent across the time on average weekends for patients in the vignette who stopped with pleasant tasks. In addition to this interaction there appeared a main effect of reported pain intensity ($F(1, 119) = 68.54$; $p < 0.001$; $\eta^2 = 0.37$). The higher the reported pain intensity, the higher the pain intensity judgements.

Discussion

1. Discussion of results

The overall aim of the present study was to examine which factors affect pain judgements made by lay people close to chronic pain patients. Four independent variables were considered (pleasant tasks were continued or stopped, unpleasant tasks were continued or stopped, medical evidence present or absent, reported pain intensity high or low). Further, fairness of the patients' behaviour in the vignette was included as dependent variable to examine whether the manipulation of pleasant and unpleasant tasks had the intended effect on this variable.

1.1 Pleasant and unpleasant tasks

Continuation of pleasant in combination with discontinuation of unpleasant tasks was hypothesised to lead to lowest fairness ratings as this behaviour can be interpreted as taking advantage of one's pain. This hypothesis was supported by an interaction between pleasant and unpleasant tasks in which behaviour of patients who continued pleasant but stopped unpleasant tasks was rated as least fair. Since the two variables, pleasant and unpleasant tasks, had the intended effect on fairness ratings, they were included in further analyses of pain judgements.

As for fairness it was hypothesised also for pain intensity that the continuation of pleasant in combination with the discontinuation of unpleasant tasks leads to lower pain ratings since it may hint at taking advantage of one's pain and, therefore, lead to caution in judgements. Again, this hypothesis was supported by an interaction between pleasant and unpleasant tasks. Results showed that pain was rated as less intense if patients continued with pleasant tasks generally and was rated as least intense if they continued doing pleasant but stopped doing unpleasant tasks.

Despite the predicted effect of continuation of pleasant in combination with discontinuation of unpleasant tasks on fairness as well as pain intensity ratings, results also show that both independent variables do not affect both dependent variables in the same way. Whereas there was a main effect of unpleasant task on fairness ratings, there was a main effect of pleasant task on pain intensity ratings. Behaviour was rated as fairest if patients continued with unpleasant tasks whereas pain was rated as most intense for patients who stopped pleasant tasks. Possibly participants think

that 'if patients give up even the tasks they like, the pain must be really bad'. This finding might also explain why fairness was not significant when included as covariate in the analysis of variance for pain intensity. Nevertheless, the results emphasise the importance of looking at mediating variables such as fairness that could help to explain when and why pain is underestimated.

1.2 Results of medical investigations

Results of medical investigations have been consistently found to impact on pain judgements in previous studies with students, medical students, nurses and physicians as judges. Accordingly, the alternative hypothesis was that relatives discount pain when medical findings are negative.

There was a considerable effect of medical evidence on pain intensity ratings insofar as judges in whose vignettes medical evidence was absent tended to rate pain lower than judges in whose vignettes medical evidence was present. However, this effect did not reach significance ($p = 0.053$) and, therefore, the null hypothesis that medical evidence does not affect pain intensity ratings cannot be rejected. There are several possible reasons for this finding. It is possible that the test I designed was not good enough to show a notable effect. Yet, other studies that used very similar vignettes found significant effects for medical evidence. Further, due to the control questions only the participants who knew the content of the vignette were included in the analyses. Therefore, it seems unlikely that the design of the test was not good enough to show effects. Alternatively, there was no effect because I was wrong to take health care professionals as reference group for setting the alternative hypothesis about lay people's judgements. However, medical evidence was found to affect pain intensity judgements also in studies using students as samples. Most likely, the study had not enough power since the sample size for the present study was calculated to detect large effects only. Therefore, a replication with a sample large enough to detect at least medium effects is needed to allow more definite conclusions.

1.3 Reported pain intensity

Like the results of medical investigations also the pain intensity as reported by patients has consistently been found to impact on pain judgements of health care professionals. Results of previous studies suggest that higher pain intensity of the suf-

ferer leads to underestimation. However, due to the differences in judgements of health care professionals and relatives found in studies in which both groups had to judge pain, the alternative hypothesis stated here was that the higher the pain intensity of the sufferer is, the higher is the rating of the relative.

Support is lent to this alternative hypothesis by the results of the analysis of variance. Relatives and friends rated pain as lower when the reported pain was low and rated pain as higher when the reported pain was high. However, when reported and rated pain were tested directly against each other, results do not support the alternative hypothesis. They showed that when pain patients reported their pain intensity to be low, their pain was rated as more intense than their report. However, when pain patients reported their pain intensity to be high, then their pain was rated as less intense. Interestingly, both results are consistent with findings of previous studies with health care professionals as judges. Apparently, lay people and health care professionals are affected by the reported pain intensity in a similar way which is puzzling, given the differences in proportions of agreement and disagreement found in the literature. Without doubt, however, it is different to judge pain of a hypothetical patient from very limited information rather than pain of a family member or friend suffering for example from cancer. Apparently, I was wrong to take the findings of the studies investigating agreement in clinical context as reference for setting my alternative hypothesis that relatives would not underestimate high pain intensity.

1.4 Background variables

The only background variable that had an effect on pain judgements was the time relatives and friends spent with the pain patients at weekends although the partial effect size of this interaction was very small ($\eta^2 = 0.05$). Time had no appreciable impact on judgements for patients who stopped carrying out pleasant tasks. Relatives judged patients' pain gradually lower, however, the more time they spent with patients at an average weekend who continued carrying out pleasant tasks. What does this finding suggest? I find it difficult to come up with an interpretation. But apparently relatives who spend more time with the person in pain at weekends think that the pain is less intense since the person in pain is still able to carry out these pleasant tasks.

Given the findings in the literature it is surprising that none of the other background variables such as satisfaction with the relationship had an effect on how relatives judged pain. It is further surprising that time spent with patients at an average weekend had an effect while time spent with patients during an average week did not. Considering the small effect size ($\eta^2 = 0.05$), further examination and replication of this result seems necessary before placing too much emphasis on it.

2. Criticism

Some methodological considerations are important to bear in mind when interpreting the results. The two main methodological problems are the response rate and the ecological validity of vignettes. Participants in this study were a highly selected group in that they were contacted via chronic pain patients who had been treated in a cognitive-behavioural inpatient unit for their chronic pain. This way of approaching participants can lead to biases such as that only relatives of patients who were satisfied with their treatment or relatives who were maritally satisfied returned the questionnaire.

Further, the response rate of 33 % is low, even given that relatives and friends were not contacted directly but via the patients. However, when looking at reasons former patients gave for not passing the questionnaire on to relatives but for returning it to me, it becomes clear that sending out questionnaires to patients was a fairly non-specific way of contacting relatives and friends. About 10 % of patients returned the questionnaire mailed to them stating that they were alone and that nobody would know them/their pain well enough or that the person they chose could not take part. When contacting patients personally, only those who knew someone they felt would be willing to take part took a questionnaire home with them. Therefore, the low response rate is possibly due to a very low response rate when questionnaires were mailed to patients rather than given out in person. Unfortunately, no separate response rates for the groups could be calculated since returned questionnaires did not allow me to distinguish between both groups. As a result, the non-random nature of the sample prevents generalisation of the findings to the larger population of chronic pain patients' relatives.

Another 'response rate' to bear in mind when interpreting the results is that of the 181 questionnaires returned only 127 (70 %) had correctly answered the control questions. Given that participants were explicitly asked to reread the vignette should they be unsure about a question, this response rate seems alarming. It is possible that the wording of the vignette text was so difficult that participants did not understand the texts correctly. However, if participants misunderstood the texts, all of them should have made more rather than less mistakes. Yet, the distribution of mistakes seems to show the opposite: Out of eight possible mistakes participants could make, 32 participants made one mistake, 13 two, 13 between three and five and two participants six mistakes. No participant made seven or eight mistakes. And one participant refused to answer any of the control questions. This points towards carelessness rather than misunderstanding of the texts. Further, to my knowledge nothing is known about how frequently readers misunderstand vignettes. Therefore, despite potential exclusion of volunteers including manipulation checks seems an improvement: interpretation of results is stronger when only answers of those participants are included who understood the vignette's content accurately.

Since vignettes are brief descriptions of hypothetical situations it is questionable to what extent they are able to represent real situations and, therefore, to what extent they are able to predict actual behaviour. However, the aim of the present study was to investigate relatives' pain judgements rather than actual behaviour in real situations. The study was thought of as a start to examine what factors possibly affect relatives when they have to judge pain. If reasonably strong effects exist, they should show up in a vignette. Now, based on the information from this vignette study, further approaches such as the use of videos or interviews as well as the use of more elaborate designs allowing the analysis and display of complex associations between variables should be used.

With regard to content validity some participants remarked on their questionnaire that they did not know whether fair behaviour meant fair of the pain patients or fair to the pain patient. Apparently the question needs more careful wording or explanation should it be used again. Nevertheless, the number of people who made this comment was small ($n = 6$) and including fairness to control whether manipulation of in-

dependent variables had the intended effect was important and useful for this study. In future studies more potentially mediating variables should be included.

3. Outlook

This work has several clinical implications, the more so because results were generated from a sample of relatives and friends of chronic pain patients. Results show similarities to judgements made by health care professionals and give new insights. With regard to pain intensity relatives and friends were affected in similar ways as health care professionals were. Both generally judged the pain to be higher when the pain reported by the patient was high (results of the ANOVA). When directly comparing rated and reported pain intensity (results of the t-test) relatives and friends as health care professionals underestimate high and overestimate low pain intensity.

No effect could be found in this study for presence or absence of medical evidence. Whereas health care professionals have been found to discount pain when medical evidence is absent, relatives in this study, were not affected by medical results since the effect, although considerable ($p = 0.053$), did not reach significance. Most likely, the present study was underpowered to detect the effect since sample size calculations were based on large effect sizes found in studies with health care professionals. Therefore, replications are needed before more definite conclusions can be drawn.

Moreover, some new variables impacting on pain judgements could be added to existing findings. Certain behaviours of pain patients such as continuation with pleasant but discontinuation with unpleasant tasks were rated as less fair. Pain of patients showing these behaviours was in turn rated as less intense. One could speculate that behaviours perceived to be unfair are able to alert the cheating detection mechanism which in turn leads to more conservative pain intensity ratings. However, this hypothesis is only speculative and needs further investigation.

Discussion

1. Discussion of results

When putting together findings on cues impacting on pain judgements and findings on pain underestimation, a variety of questions appeared of which I set out to explore two in the present piece of work:

1. Can underestimation of pain be accounted for by one or both of two explanations:
 - verbal report as important but missing cue and / or
 - an alerted cheating detection mechanism?
2. Which cues do affect relatives' rather than health care professionals' pain judgements?

1.1 Accounts for pain underestimation

In my first study I looked at cue utilisation of health care professionals in situations in which they have to judge pain. Results showed that health care professionals across specialties considered verbal report as an important cue. Therefore, I designed a second study in which the two proposed alternative accounts for underestimation were compared. Results of this second study indicated that health care professionals underestimated patients' pain, that the extent of underestimation, however, varied depending on the cues given. Doctors and nurses who had to make their judgement on the basis of patients' facial expressions only underestimated patients' pain intensity to a greater extent than doctors and nurses who were given patients' own pain intensity ratings in addition to the facial expressions. Doctors and nurses who in addition to both cues were primed to expect some cheating were also more conservative in their judgements and underestimated pain as much as those did who only saw patients' faces. Therefore, both alternative explanations, verbal report as important but missing cue as well as an alerted cheating detection device, could account for underestimation.

In addition, results indicated that not only the presence and kind of cues impact on pain underestimation but also judges' expectations and experience. Judges who assumed a higher rate of pain exaggeration in patients underestimated pain to a greater extent. This expectation in turn was shaped by the time doctors and nurses

had been working in their profession and by the specialty they were working in. The longer health care professionals had been working in their profession, the fewer patients they presumed to exaggerate their pain. Further, doctors and nurses working in Oncology expected fewer patients to exaggerate their pain.

Results of the second study, like those of a previous one (Kappesser & Williams, 2002), suggest once more that pain underestimation occurs not because of difficulties in pain recognition but because of the challenge observers face when interpreting pain expressions. Controlling for cues in the judgement situation revealed that the presence or absence of an important pain behaviour, verbal report, led to a greater extent of underestimation as did a cue that placed the available pain expression in a context enhancing the possibility for cheating. Further, the likelihood for pain underestimation is increased by expectations inherent in judges that in turn are shaped by the context judges work in.

Given that underestimation has adverse consequences for patients and judges, what possibilities do results of the second study suggest to enhance agreement? With regard to verbal report as missing cue, the first possibility certainly is that we should ask patients about their pain whenever this is possible. Verbal report can be elicited by simply talking to patients. A more standardised approach is to choose one or several of the reliable and valid pain scales available to assess the pain experience (for an overview see Appendix I.2).

With regard to an alerted cheating detection mechanism it is more difficult to reach a conclusion, yet, it is important that the existence of such a device is not used to excuse pain underestimation (e.g. 'Because the cheating detection device is a universal cognitive mechanism, there is nothing that can be done to counter it.'). When asked about evolutionary psychologists' view of genetic determinism, Cosmides stated that 'behavior is a joint product of the information in our environments and the programs in our heads. These programs, in turn, were created during our lifetime through a dynamic interaction between our genes and the environment. Which genes we have is a function of past environments, which, over deep time, selected for some genes over others. ... Is this 'genetic determinism'? One might as easily call it 'environmental determinism'.' (Cosmides, 2002, p. 3-4). Consequently, the cheating detection

mechanism is no excuse for underestimating pain. Rather, the awareness of its existence should help not to become more conservative when judging pain and, therefore, counteract its consequences.

Results further suggest that it might be helpful to reflect on the setting we work in as well as on our prior experiences with pain patients as they shape our expectations about faking or exaggerating in patients.

1.2 Impact of selected cues on pain judgements of patients' relatives

In the third study the impact of four cues (reported pain intensity, medical evidence, behaviour regarding pleasant and regarding unpleasant tasks) on pain intensity ratings of chronic pain patients' relatives was investigated. With regard to reported pain intensity, results resembled those of health care professionals. Generally, both groups rated pain to be intense when the patient had reported intense pain and low when the patient had reported low pain. However, when comparing the exact intensity ratings, relatives - like health care professionals - underestimated high reported pain intensities and overestimated low reported pain intensities. How can the discrepancy between these two findings be explained? In an unpublished data simulation study Idvall and Brudin (2004) showed that the apparent tendency for judges to underestimate severe but not mild pain results from judges using a narrower distribution of their ratings than patients. However, more research is needed to investigate why judges have this narrower distribution. One account - that health care professionals due to their encounter with a very broad range of pain intensities have a broader range than patients and, therefore, underestimate what patients consider as worst pain – is discussed below.

Relatives differed from health care professionals with regard to medical evidence, since relatives were not affected in their pain intensity ratings by the presence or absence of supporting medical findings. However, since it is uncertain whether the study was underpowered with regard to this effect, replications are needed before more definite conclusions can be drawn.

So far, nothing is known about utilisation of cues for pain intensity ratings by relatives of chronic pain patients or lay people in general. However, the hypothesis to examine

would be that medically trained and lay people resemble each other in utilisation of behavioural but differ in utilisation of contextual cues. Which contextual cues could be considered important by relatives? Results with regard to patients' behaviour could help answering this question. Behaviour of patients considered unfair such as continuing with pleasant but stopping unpleasant tasks apparently is a contextual cue important for relatives when judging pain. Possibly, this contextual cue alerts the cheating detection device in relatives since pain of patients whose behaviour was perceived to be unfair was rated as less intense.

2. Criticism

2.1 Studies on agreement between patients and judges in the clinical context

Although I was fairly successful in finding and putting together many studies on agreement between judges and patients, it is to be expected that not all available studies are reported here. Therefore, it is questionable to which extent the graphical summaries of proportions of (dis)agreement can be generalised. Given that pain underestimation has raised such concern in the literature and given that there exists a host of studies rather than only the few cited repeatedly in the literature, a systematic review and a meta-analysis are necessary next steps towards a better understanding of (dis)agreement on others' pain. In addition, by subgroup analysis both could possibly help to answer further questions that emerged from the introduction: Does and if so how does the kind of diagnosis the patients have affect proportions of (dis)agreement? Are there differences in proportions or amounts of (dis)agreement between different groups of judges (e.g. health care professionals and relatives)?

2.2 Neglect of other accounts for underestimation

Choosing to investigate the two questions outlined above also implied choosing not to investigate other available methodological and theoretical accounts for pain underestimation.

Of course underestimation and its extent are always dependent on the definition of agreement. The criterion first used by lafrati (1986) of ± 1 cm was set arbitrarily. If one, for instance, chose the range of one standard deviation as criterion to define agreement, percentages of agreement would be higher. On the other hand, differences of about 1 cm have in the meantime been shown to be judged as clinically sig-

nificant by patients when they had to evaluate changes in their pain intensity (Bird & Dickson, 2001; Gallagher et al., 2001; Kelly, 1998; Todd et al., 1996). Therefore, the criterion of ± 1 cm can be considered as a clinically meaningful criterion to define agreement.

Related to the question of agreement definition is the question of how the pain scale chosen to assess pain intensity affects agreement. When summarising the selected agreement studies for the introduction, to my surprise there appeared to be no systematic differences between agreement studies depending on the assessment of pain. However, studies by Clipp and George (1992) and Madison and Wilkie (1995) on agreement between judges and patients showed that agreement on more objective measures such as presence or absence of pain was better than agreement on more subjective measures such as intensity or quality of pain or its pattern. This finding is further indirectly supported by results of study 2. Although pain intensity of patients was generally underestimated by health care professionals, only very few doctors and nurses (27 ratings out of 960 possible ones; 0.03 %) said that patients did not experience any pain at all (despite all of the patients reporting experiencing pain). Another possibility to explore this question further comes to mind as the task of quantifying the intensity of an experience seems to be peculiar for the pain field. When, for instance, emotional experiences are researched, participants are frequently asked how they would feel (happy, sad, angry etc.). However, they are rarely asked to quantify this emotional experience. It would be interesting to investigate whether underestimation occurs too when the intensity of other emotional experiences has to be assessed.

Another theoretical account for underestimation occurred to me when I talked to participants about the difficulties they were facing when they had to assess pain. Health care professionals due to their work experience have encountered a very broad range of pain intensities. Therefore, what they consider as extremely intense pain might well differ from the worst experience a patient has ever had. As a result, they tend to rate patients' pain as less intense than patients themselves. Some participants working in A & E, for example, commented that they would compare patients complaining about not being treated immediately despite their unbearable pain to patients they had encountered before with what they would consider as unbearable

pain (one participant gave the example of 'a knife stuck in patient's stomach', others said that 'someone who smiles cannot be in pain'). This account is related to the hypothesis of developed insensitivity and in this context was proposed by Bergh and Sjöström (1999). Further support stems from the data simulation study by Idvall and Brudin (2004) where underestimation occurred whenever judges used a narrower range of ratings than patients did (e.g. 2-9 rather than 0-10).

To conclude, research on the effects of methodology on underestimation is necessary given the alarmingly poor quality of many agreement studies in the clinical context (see Appendix I.1). Furthermore, research on methodological accounts for underestimation could help establishing a standard how agreement studies could be most appropriately analysed. Last but not least the hypothesis that underestimation occurs because judges use narrower ranges of the rating scale needs further exploration.

3. Outlook

Resulting from this PhD, what recommendations are there for future research? Pain judgements in general and pain underestimation in particular have raised much concern in the pain literature since they have important consequences in day-to-day clinical work. Although the concern about pain underestimation is appropriate, care is needed when using this term since, as shown in the introduction, many studies investigating agreement suffer from methodological flaws. Further, there is a need to look more closely at what the methodologically sober studies tell us. Many of them were designed so that verbal report was unavailable to judges. Of course, this per se is no methodological flaw. However, authors need to be specific about why they make which cues available to judges. Teske and colleagues (Teske, Daut & Cleeland, 1983), for example, were interested in predicting patients' verbal report from certain non-verbal behaviours and explicitly stated this in their study aim. Many other authors, however, were not. Because of its emphasis on cues, the lens model seems a structure useful to be more specific about research questions and study aims.

Furthermore, caution is needed not to use the term underestimation in a too general way. There appear to be differences between groups of judges and diagnoses of patients. Additionally, the scales used to assess pain intensity possibly affect agree-

ment as well. A systematic review and a meta-analysis might help to clarify the impact of these variables.

Another improvement would be to switch research focus from evaluating pain judgements to explaining judgements. It is alarming that since Lander's review in 1990 no substantial progress has been made. The main part of studies investigating agreement is still atheoretical. In the present work, social contract theory with the cheating detection mechanism was used as social context in which pain judgements were investigated. And although more research is needed to further explore the appropriateness and usefulness of the concept of cheating detection mechanism, the results presented here seem at least promising. Only successful explanation of pain underestimation offers the chance to set up strategies capable of encountering its consequences.

References

- American Psychiatric Association, A. P. (Ed.). (1994). *Diagnostic and statistical manual of mental disorders: DSM IV TM*. (4th ed.). Washington, DC: American Psychiatric Association.
- Bartfield, J. M., Salluzo, R. F., Raccio-Robak, N., Funk, D. L. & Verdile, V. P. (1997). Physician and patient factors influencing the treatment of low back pain. *Pain*, 73, 209-211.
- Bergh, I. & Sjöström, B. (1999). A comparative study of nurses' and elderly patients' ratings of pain and pain tolerance. *Journal of Gerontological Nursing*, 25, 30-36.
- Bird, S. B. & Dickson, E. W. (2001). Clinically significant changes in pain along the Visual Analog Scale. *Annals of Emergency Medicine*, 38, 639-643.
- Borg, I. & Groenen, P. (1997). *Modern multidimensional scaling. Theory and applications*. New York: Springer.
- Brunelli, C., Costantini, M., Di Giulio, P., Gallucci, M., Fusco, F., Miccinesi, G., Paci, E., Peruselli, C., Morino, P., Piazza, M., Tamburini, M. & Toscani, F. (1998). Quality-of-life evaluation: when do terminal cancer patients and health-care providers agree? *Journal of Pain and Symptom Management*, 15, 151-158.
- Brunswik, E. (1952). *The conceptual framework of psychology* (Vol. 1). Chicago: University of Chicago Press.
- Carey, T. S. & Mills Garrett, J. (2003). The relation of race to outcomes and the use of health care services for acute low back pain. *Spine*, 28, 390-394.
- Chibnall, J. T. & Tait, R. C. (1995). Observer perception of low back pain: effects of pain report and other contextual factors. *Journal of Applied Social Psychology*, 25, 418-439.
- Chibnall, J. T. & Tait, R. C. (1999). Social and medical influences on attributions and evaluations of chronic pain. *Psychology and Health*, 14, 719-729.
- Chibnall, J. T., Tait, R. C. & Ross, L. R. (1997). The effects of medical evidence and pain intensity on medical student judgements of chronic pain patients. *Journal of Behavioral Medicine*, 20, 257-271.
- Choiniere, M., Melzack, R., Girard, N., Rondeau, J. & Paquin, M.-J. (1990). Comparisons between patients' and nurses' assessment of pain and medication efficacy in severe burn injuries. *Pain*, 40, 143-152.

- Cleeland, C. S., Cleeland, L. M., Dar, R. & Rinehardt, L. C. (1986). Factors influencing physician management of cancer pain. *Cancer*, 58, 796-800.
- Cleeland, C. S., Gonin, R., Baez, L., Loehrer, P. & Pandya, K. J. (1997). Pain and treatment of pain in minority patients with cancer. *Annals of Internal Medicine*, 127, 813-816.
- Clipp, E. C. & George, L. K. (1992). Patients with cancer and their spouse caregivers. Perceptions of illness experience. *Cancer*, 69, 1074-1079.
- Connolly, T., Arkes, H. R. & Hammond, K. R. (2000). General introduction. In K. R. Hammond (Ed.), *Judgement and decision making: an interdisciplinary reader*. (2nd ed., pp. 1-11). Cambridge: Cambridge University Press.
- Cosmides, L. (1989). The logic of social exchange: has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, 31, 187-276.
- Cosmides, L. (2002). *Interview for Chilean Newspaper 'El Mercurio'*. World Wide Web, <http://www.psych.ucsb.edu/research/cep/ledainterview.htm>. Retrieved February 05, 2002, from the World Wide Web:
- Cosmides, L. & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. Tooby (Ed.), *The adapted mind: evolutionary psychology and the generation of culture*. Oxford: Oxford University Press.
- Cosmides, L. & Tooby, J. (2000). *Evolutionary Psychology: a primer*. World Wide Web, <http://www.psych.ucsb.edu/research/cep/primer.html>. Retrieved October 17, 2002, from the World Wide Web:
- Cox, D. R. & Wermuth, N. (1994). Tests of linearity, multivariate normality and adequacy of linear scores. *Applied Statistics, Journal of the Royal Statistical Society, C*, 43, 347-355.
- Cox, D. R. & Wermuth, N. (1996). *Multivariate dependencies: models, analysis and interpretation*. London: Chapman & Hall.
- Craig, K. D. (1992). The facial expression of pain: better than a thousand words? *American Pain Society Journal*, 1, 153-162.
- Craig, K. D. (2002). Pain in infants and children: sociodevelopmental variations on the theme. In M. A. Giamberardino (Ed.), *Pain 2002 - an updated review. Refresher course syllabus* (pp. 305-314). Seattle: IASP Press.
- Craig, K. D., Hill, M. L. & McMurtry, B. W. (1999). Detecting deception and malingering. In E. F. Kremer (Ed.), *Handbook of pain syndromes: biopsychosocial per-*

- specives* (pp. 41-58). Mahwah, NJ: Lawrence Erlbaum.
- Craig, K. D., Hyde, S. A. & Patrick, C. J. (1991). Genuine, suppressed and faked facial behavior during exacerbation of chronic low back pain. *Pain, 46*, 161-171.
- Craig, K. D. & Patrick, C. J. (1985). Facial expression during induced pain. *Journal of Personality and Social Psychology, 48*, 1080-1091.
- Craig, K. D. & Prkachin, K. M. (1983). Nonverbal measures of pain. In R. Melzack (Ed.), *Pain measurement and assessment* (pp. 173-179). New York: Raven Press.
- Craig, K. D., Prkachin, K. M. & Grunau, R. E. (2001). The facial expression of pain. In R. Melzack (Ed.), *Handbook of pain assessment* (pp. 153-169). New York: Guilford Press.
- Creameans-Smith, J. K., Stephens, M. A. P., Franks, M. M., Martire, L. M., Druley, J. A. & Wojno, W. C. (2003). Spouses' and physicians' perceptions of pain severity in older women with osteoarthritis: dyadic agreement and patients' well-being. *Pain, 106*, 27-34.
- Dar, R., Beach, C. M., Barden, P. L. & Cleeland, C. S. (1992). Cancer pain in the marital system: a study of patients and their spouses. *Journal of Pain and Symptom Management, 7*, 87-93.
- Daut, R. L. & Cleeland, C. S. (1982). The prevalence and severity of pain in cancer. *Cancer, 50*, 1913-1918.
- Davitz, L. J. & Davitz, J. R. (1975). How nurses view patient suffering. *Registered Nurse, 38*, 69-74.
- Donovan, M., Dillon, P. & McGuire, L. (1987). Incidence and characteristics of pain in a sample of medical-surgical inpatients. *Pain, 30*, 69-78.
- Drayer, R. A., Henderson, J. & Reidenberg, M. (1999). Barriers to better pain control in hospitalized patients. *Journal of Pain and Symptom Management, 17*, 434-440.
- Dudley, S. R. & Holm, K. (1984). Assessment of the pain experience in relation to selected nurse characteristics. *Pain, 18*, 179-186.
- Dworkin, R. H., Handlin, D. S., Richlin, D. M., Brand, L. & Vannucci, C. (1985). Unraveling the effects of compensation, litigation, and employment on treatment response in chronic pain. *Pain, 23*, 49-59.
- Eccleston, C. & Crombez, G. (1999). Pain demands attention: a cognitive-affective model of the interruptive function of pain. *Psychological Bulletin, 125*, 356-366.

- Ekman, P. & Friesen, W. V. (1978). *Facial Action Coding System*. Palo Alto: Consulting Psychologists Press.
- Ekman, P. & O'Sullivan, M. (1991). Who can catch a liar? *American Psychologist*, *46*, 913-920.
- Everett, J. J., Patterson, D. R., Marvin, J. A., Montgomery, B., Ordonez, N. & Campbell, K. (1994). Pain assessment from patients with burns and their nurses. *Journal of Burn Care & Rehabilitation*, *15*, 194-198.
- Faul, F. & Erdfelder, E. (1992). GPOWER: A priori, post-hoc, and compromise power analyses for MS-DOS. Bonn: Bonn University, Department of Psychology.
- Faust, D., Hart, K., Guilmette, T. J. & Arkes, H. R. (1988). Neuropsychologists' capacity to detect adolescent malingerers. *Professional Psychology: Research and Practice*, *19*, 508-515.
- Ferrell, B. R., Eberts, M. T., McCaffery, M. & Grant, M. (1991). Clinical decision making and pain. *Cancer Nursing*, *14*, 289-297.
- Fishbain, D. A. (1994). Secondary gain concept. Definition problems and its abuse in medical practice. *Journal of the American Pain Society*, *3*, 264-273.
- Fishbain, D. A., Cutler, R., Rosomoff, H. L. & Rosomoff, R. S. (1999). Chronic pain disability exaggeration/malingering and submaximal effort research. *The Clinical Journal of Pain*, *15*, 244-274.
- Fishbain, D. A., Rosomoff, H. L., Cutler, R. B. & Rosomoff, R. S. (1995). Secondary gain concept: a review of the scientific evidence. *Clinical Journal of Pain*, *11*, 6-21.
- Flor, H., Turk, D. C. & Rudy, T. E. (1989). Relationship of pain impact and significant other reinforcement of pain behaviors: the mediating role of gender, marital status and marital satisfaction. *Pain*, *38*, 45-50.
- Fordyce, W. E. (1976). *Behavioral methods for chronic pain and illness*. St. Louis, MO: Mosby.
- Galin, K. E. & Thorn, B. E. (1993). Unmasking pain: detection of deception in facial expressions. *Journal of Social and Clinical Psychology*, *12*, 182-197.
- Gallagher, E. J., Liebman, M. & Bijur, P. E. (2001). Prospective validation of clinically important changes in pain severity measured on a Visual Analog Scale. *Annals of Emergency Medicine*, *38*, 633-638.
- Gauthier, J. C., Finley, G. A. & McGrath, P. J. (1998). Children's self-report of post-operative pain intensity and treatment threshold: determining the adequacy of

- medication. *Clinical Journal of Pain*, 14, 116-120.
- Gigerenzer, G. (1996). On narrow norms and vague heuristics: a reply to Kahneman and Tversky. *Psychological Review*, 103, 592-596.
- Gigerenzer, G. (2000). *Adaptive thinking. Rationality in the real world*. New York: Oxford University Press.
- Gigerenzer, G. & Hug, K. (1992). Domain-specific reasoning: social contracts, cheating, and perspective change. *Cognition*, 43, 127-171.
- Gigerenzer, G., Todd, P. M. & Gerd Gigerenzer, A. R. (Eds.). (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Hadjistavropoulos, H. D. & Craig, K. D. (1994). Acute and chronic low back pain: cognitive, affective, and behavioral dimensions. *Journal of Consulting and Clinical Psychology*, 62, 341-349.
- Hadjistavropoulos, H. D., Craig, K. D., Grunau, R. V. E. & Johnston, C. C. (1994). Judging pain in newborns: facial and cry determinants. *Journal of Pediatric Psychology*, 19, 485-491.
- Hadjistavropoulos, T. & Craig, K. D. (2002). A theoretical framework for understanding self-report and observational measures of pain: a communications model. *Behaviour Research and Therapy*, 40, 551-570.
- Hadjistavropoulos, T., McMurtry, B. & Craig, K. D. (1996). Beautiful faces in pain: biases and accuracy in the perception of pain. *Psychology and Health*, 11, 411-420.
- Hadjistavropoulos, T., von Baeyer, C. L. & Craig, K. D. (2001). Pain assessment in persons with limited ability to communicate. In R. Melzack (Ed.), *Handbook of pain assessment* (pp. 134-149). New York: Guilford Press.
- Halfens, R., Evers, G. & Abu-Saad, H. (1990). Determinants of pain assessment by nurses. *International Journal of Nursing Studies*, 27, 43-49.
- Hamers, J. P. H., van den Hout, M. A., Halfens, R. J. G., Huijjer Abu-Saad, H. & Heijltjes, A. E. G. (1997). Differences in pain assessment and decisions regarding the administration of analgesics between novices, intermediates and experts in pediatric nursing. *International Journal of Nursing Studies*, 34, 325-334.
- Hammond, K. R. (1955). Probabilistic functionalism and the clinical method. *Psychological Review*, 62, 255-262.
- Hammond, K. R. (1964). Part II. Clinical inference in nursing: a methodological ap-

- proach. *Nursing Research*, 13, 315-319.
- Hardt, J., Petrak, F., Filipas, D. & Egle, U. T. (2004). Adaptation to life after surgical removal of the bladder - an application of graphical Markov models for analysing longitudinal data. *Statistics in Medicine*, 23, 649-666.
- Harrison, A. (1993). Comparing nurses' and patients' pain evaluations: a study of hospitalised patients in Kuwait. *Social Science and Medicine*, 36, 683-692.
- Heft, M. W., Gracely, R. H., Dubner, R. & McGrath, P. A. (1980). A validation model for verbal descriptor scaling of human clinical pain. *Pain*, 9, 363-373.
- Iafrafi, N. S. (1986). Pain on the burn unit: patient vs. nurse perceptions. *Journal of Burn Care & Rehabilitation*, 7, 413-416.
- Jacox, A. K. (1979). Assessing pain. *American Journal of Nursing*, 79, 895-900.
- Jacox, A. K. (1980). The assessment of pain. In S. Grass (Ed.), *Pain: meaning and management* (pp. 75-88). New York: Spectrum.
- Jensen, M. P. & Karoly, P. (2001). Self-report scales and procedures for assessing pain in adults. In R. Melzack (Ed.), *Handbook of pain assessment* (pp. 15-34). New York: Guilford Press.
- Johnson, M. (1977). Assessment of clinical pain. In A. K. Jacox (Ed.), *Pain: a source-book for nurses and other health professionals* (pp. 139-166). Boston: Little, Brown.
- Johnston, C. C., Gagnon, A. J., Fullerton, L., Common, C., Ladores, M. & Forlini, S. (1998). One-week survey of pain intensity on admission to and discharge from the emergency department: a pilot study. *Journal of Emergency Medicine*, 16, 377-382.
- Kahneman, D. & Tversky, A. (1982). Variants of uncertainty. *Cognition*, 11, 143-157.
- Kappesser, J. & Williams, A. C. d. C. (2002). Pain and negative emotions in the face: judgements of health care professionals. *Pain*, 99, 197-206.
- Keefe, F. J. & Dunsmore, J. (1992). Pain behaviour. Concepts and controversies. *American Pain Society Journal*, 1, 92-100.
- Kelly, A.-M. (1998). Does the clinically significant difference in Visual Analog Scale pain scores vary with gender, age, or cause of pain? *Academic Emergency Medicine*, 5, 1086-1090.
- Kelly, K. J. (1964a). Part I. An approach to the study of clinical inferences in nursing. *Nursing Research*, 13, 314-315.
- Kelly, K. J. (1964b). Part III. Utilization of the 'lens model' method to study the infer-

- ential process of the nurse. *Nursing Research*, 13, 319-322.
- Krivo, S. & Reidenberg, M. M. (1996). Assessment of patients' pain. *New England Journal of Medicine*, 334, 59.
- Krokosky, N. J. & Reardon, R. C. (1989). The accuracy of nurses' and doctors' perceptions of patient pain. In R. A. Wiese (Ed.), *Key aspects of comfort: management of pain, fatigue and nausea* (pp. 127-134). New York: Springer.
- Lander, J. (1990a). Clinical judgements in pain management. *Pain*, 42, 15-22.
- Lander, J. (1990b). Fallacies and phobias about addiction and pain. *British Journal of Addiction*, 85, 803-809.
- Lavies, N., Hart, H., Rounsefell, B. & Runciman, W. (1992). Identification of patient, medical and nursing staff attitudes to postoperative opioid analgesia: stage 1 of a longitudinal study of postoperative analgesia. *Pain*, 48, 313-319.
- Leavitt, F. & Sweet, J. J. (1986). Characteristics and frequency of malingering among patients with low back pain. *Pain*, 25, 357-364.
- Lenburg, C. B., Glass, H. P. & Davitz, L. J. (1970). Inferences of physical pain and psychological distress. II. In the relation to the stage of the patient's illness and occupation of the perceiver. *Nursing Research*, 19, 392-398.
- LeResche, L. (1982). Facial expression of pain: a study of candid photographs. *Journal of Nonverbal Behaviour*, 7, 46-56.
- LeResche, L. & Dworkin, S. F. (1988). Facial expression of pain and emotions in chronic TMD patients. *Pain*, 35, 71-78.
- LeResche, L., Dworkin, S. F., Wilson, L. & Ehrlich, K. J. (1992). Effect of temporomandibular disorder pain duration on facial expressions and verbal report of pain. *Pain*, 51, 289-295.
- Lousberg, R., Schmidt, A. J. M. & Goenman, N. H. (1992). The relationship between spouse solicitousness and pain behaviour: searching for more experimental evidence. *Pain*, 51, 75-79.
- Madison, J. L. & Wilkie, D. J. (1995). Family members' perceptions of cancer pain. Comparisons with patient sensory report and by patient psychologic status. *Nursing Clinics of North America*, 30, 625-645.
- Main, C. J. & Spanswick, C. C. (1995). 'Functional overlay', and illness behaviour in chronic pain: distress or malingering? Conceptual difficulties in medico-legal assessment of personal injury claims. *Journal of Psychosomatic Research*, 39, 737-753.

- Marks, R. M. & Sachar, E. J. (1973). Undertreatment of medical inpatients with narcotic analgesics. *Annals of Internal Medicine*, 78, 173-181.
- Mason, D. J. (1981). An investigation of the influences of selected factors on nurses' inferences of patient suffering. *International Journal of Nursing Studies*, 18, 251-259.
- McCaffery, M. (1980). Understanding your patient's pain. *Nursing*, 10, 26-31.
- McKinley, S. & Botti, M. (1991). Nurses' assessment of pain in hospitalised patients. *The Australian Journal of Advanced Nursing*, 9, 8-14.
- Meenan, R. F., Mason, J. H., Anderson, J. J., Guccione, A. A. & Kazis, L. E. (1992). AIMS2. The content and properties of a revised and expanded Arthritis Impact Measurement Scales health status questionnaire. *Arthritis & Rheumatism*, 35, 1-10.
- Melzack, R., Abbott, F. V., Zackon, W., Mulder, D. S. & Davis, M. W. L. (1987). Pain on a surgical ward: a survey of the duration and intensity of pain and the effectiveness of medication. *Pain*, 29, 67-72.
- Mendelson, G. (1992a). Chronic pain and compensation. In S. P. Tyrer (Ed.), *Psychology, psychiatry and chronic pain* (pp. 67-78). Cambridge: University Press.
- Mendelson, G. (1992b). Compensation and pain. *Pain*, 48, 121-123.
- Merskey, H., Albe-Fessard, D. G., Bonica, J. J., Carmen, A., Dubner, R., Kerr, F. W. L., Lindblom, U., Mumford, J. M., Nathan, P. W., Noordenbos, W., Pagni, C. A., Renaer, M. J., Sternbach, R. A. & Sunderland, S. (1979). IASP subcommittee on taxonomy. *Pain*, 6, 249-255.
- Miaskowski, C., Zimmer, E. F., Barrett, K. M., Dibble, S. L. & Walhagen, M. (1997). Differences in patients' and family caregivers' perceptions of the pain experience influence patient and caregiver outcomes. *Pain*, 72, 217-226.
- Newton-John, T. R. O. (2002). Solicitousness and chronic pain: a critical review. *Pain Reviews*, 9, 7-27.
- O'Brien, J. & Francis, A. (1988). The use of next-of-kin to estimate pain in cancer patients. *Pain*, 35, 171-178.
- Oberst, M. T. (1978). Nurses' inferences of suffering: the effects of nurse-patient similarity and verbalisations of distress. In M. J. Nelson (Ed.), *Clinical perspectives in nursing research* (pp. 38-60). New York: Teachers College Press.
- Parsons, T. (1951). Social structure and dynamic process: the case of modern medical practice. In T. Parsons (Ed.), *The Social System* (pp. 428-479). Glencoe:

The Free Press.

- Patrick, C. J., Craig, K. D. & Prkachin, K. M. (1986). Observer judgements of acute pain: facial action determinants. *Journal of Personality and Social Psychology*, *50*, 1291-1298.
- Peabody, J. W., Luck, J., Glassman, P., Dresselhaus, T. R. & Lee, M. (2000). Comparison of vignettes, standardized patients, and chart abstraction. A prospective validation study of 3 methods for measuring quality. *Journal of the American Medical Association*, *283*, 1715-1722.
- Pilowsky, I. (1985). Malingeringophobia. *The Medical Journal of Australia*, *143*, 571-572.
- Poole, G. & Craig, K. D. (1992). Judgements of genuine, suppressed, and faked facial expressions of pain. *Journal of Personality and Social Psychology*, *5*, 797-805.
- Prkachin, K. M. (1986). Pain behaviour is not unitary. *The Behavioral and Brain Sciences*, *9*, 754-755.
- Prkachin, K. M. (1992). The consistency of facial expressions of pain: a comparison across modalities. *Pain*, *51*, 297-306.
- Prkachin, K. M. (2002). Pain behavior and the pretence of knowledge. *Behavioral and Brain Sciences*, *25*, 470.
- Prkachin, K. M., Berzins, S. & Mercer, S. R. (1994). Encoding and decoding of pain expressions: a judgement study. *Pain*, *58*, 253-259.
- Prkachin, K. M. & Craig, K. D. (1995). Expressing pain: the communication and interpretation of facial pain signals. *Journal of Nonverbal Behaviour*, *19*, 191-205.
- Prkachin, K. M., Currie, N. A. & Craig, K. D. (1983). Judging nonverbal expressions of pain. *Canadian Journal of Behavioral Sciences*, *15*, 409-421.
- Prkachin, K. M. & Mercer, S. R. (1989). Pain expression in patients with shoulder pathology: validity, properties and relationship to sickness impact. *Pain*, *39*, 257-265.
- Prkachin, K. M., Solomon, P., Hwang, T. & Mercer, S. R. (2001). Does experience influence judgements of pain behaviour? Evidence from relatives of pain patients and therapists. *Pain Research and Management*, *6*, 105-112.
- Riemsma, R. P., Taal, E. & Rasker, J. J. (2000). Perceptions about perceived functional disabilities and pain of people with rheumatoid arthritis: differences between patients and their spouses and correlates with well-being. *Arthritis Care and Research*, *13*, 255-261.

- Riemsma, R. P., Taal, E., Rasker, J. J., Houtman, P. M., van Paassen, H. C. & Wiegman, O. (1996). Evaluation of a Dutch version of the AIMS2 for patients with rheumatoid arthritis. *British Journal of Rheumatology*, *35*, 755-760.
- Romano, J. M., Turner, J. A., Friedman, L. S., Bulcroft, R. A., Jensen, M. P., Hops, H. & Wright, S. F. (1992). Sequential analysis of chronic pain behaviors and spouse responses. *Journal of Consulting and Clinical Psychology*, *60*, 777-782.
- Rosenthal, R. (1982). Conducting judgment studies. In P. Ekman (Ed.), *Handbook of methods in nonverbal behavior research*. (pp. 287-361). Cambridge: Cambridge University Press.
- Schiffman, S. S., Reynolds, M. L. & Young, F. W. (1981). *Introduction to multidimensional scaling*. New York: Academic Press.
- Schuler, M., Neuhauser, T., Hauer, K., Oster, P., Razus, D. & Hacker, M. (2001). Schmerzerkennung bei geriatrischen Patienten durch ein interdisziplinäres Team: Urteilssicherheit und Einflussfaktoren. *Zeitschrift für Gerontologie und Geriatrie*, *34*, 376-386.
- Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill.
- Simmonds, M. J., Barlow, T. & Kreth, K. (1996). *Symptom sincerity of patients with chronic low back pain (LBP). Do physical therapists believe their patients?* Paper presented at the 8th World Congress on Pain, Vancouver, Canada.
- Skevington, S. M. (1995). *Psychology of pain*. Chichester: John Wiley & Sons.
- Solomon, P. (2001). Congruence between health professionals' and patients' pain ratings: a review of the literature. *Scandinavian Journal of Caring Sciences*, *15*, 174-180.
- Sullivan, M. J. L., Adams, H. & Sullivan, M. E. (2004). Communicative dimensions of pain catastrophizing: social cueing effects on pain behaviour and coping. *Pain*, *107*, 220-226.
- Tait, R. C. & Chibnall, J. T. (1997). Physician judgement of chronic pain patients. *Social Science and Medicine*, *45*, 1199-1205.
- Tamburini, M., Rosso, S., Gamba, A., Mencaglia, E., De Conno, F. & Ventafridda, V. (1992). A therapy impact questionnaire for quality-of-life assessment in advanced cancer research. *Annals of Oncology*, *3*, 565-570.
- Taylor, A. G., Skelton, J. A. & Butcher, J. (1984). Duration of pain condition and

- physical pathology as determinants of nurses' assessments of patients in pain. *Nursing Research*, 33, 4-8.
- Teasell, R. W. & Merskey, H. (1997). Chronic pain disability in the workplace. *Pain Research and Management*, 2, 197-205.
- Teske, K., Daut, R. L. & Cleeland, C. S. (1983). Relationships between nurses' observations and patients' self-reports of pain. *Pain*, 16, 289-296.
- Todd, K. H., Funk, G. F., Fund, J. P. & Bonacci, R. (1996). Clinical significance of reported changes in pain severity. *Annals of Emergency Medicine*, 27, 485-489.
- Todd, K. H., Lee, T. & Hoffman, J. R. (1994). The effect of ethnicity on physician estimates of pain severity in patients with isolated extremity trauma. *Journal of the American Medical Association*, 271, 925-928.
- Turk, D. C. (1996). Cognitive factors in chronic pain and disability. In K. D. Craig (Ed.), *Advances in cognitive-behavioral therapy*. (pp. 83-115). Thousand Oaks, CA: Sage Publications.
- Turk, D. C. & Melzack, R. (2001). The measurement of pain and the assessment of people experiencing pain. In R. Melzack (Ed.), *Handbook of pain assessment* (pp. 3-11). New York: Guilford Press.
- Tversky, A. & Kahneman, D. (1974). Judgement under uncertainty: heuristics and biases. *Science*, 185, 1124-1131.
- Tversky, A. & Kahneman, D. (2000). Judgment under uncertainty: heuristics and biases. In K. R. Hammond (Ed.), *Judgment and decision making* (pp. 35-52). Cambridge: Cambridge University Press.
- van der Does, A. J. W. (1989). Patients' and nurses' ratings of pain and anxiety during burn wound care. *Pain*, 39, 95-101.
- Vaughan, K. B. & Lanzetta, J. T. (1980). Vicarious instigation and conditioning of facial expressive and autonomic responses to a model's expressive display of pain. *Journal of Personality and Social Psychology*, 38, 909-923.
- von Baeyer, C. L., Johnson, M. E. & McMillan, M. J. (1984). Consequences of non-verbal expression of pain: patient distress and observer concern. *Social Science and Medicine*, 19, 1319-1324.
- Waddell, G. (1991). Low back disability. A syndrome of Western civilization. *Neurosurgery Clinics of North America*, 2, 719-738.
- Wall, P. (1999). *Pain the science of suffering*. London: Phoenix.

- Wermuth, N. (1998). Graphical Markov Models. In D. Banks (Ed.), *Encyclopedia of Statistical Science* (pp. 284-300). New York: Wiley.
- Wermuth, N. (2003). Analysing social science data with graphical Markov models. In S. Richardson (Ed.), *Highly structured stochastic systems*. (pp. 47-52). Oxford: Oxford University Press.
- Wermuth, N. & Cox, D. R. (1998). Statistical dependence and independence. In T. Colton (Ed.), *Encyclopedia of Biostatistics* (pp. 4260-4267). New York: Wiley.
- Wilkie, D. J., Keefe, F. J., Dodd, M. J. & Copp, L. A. (1992). Behavior of patients with lung cancer: description and associations with oncologic and pain variables. *Pain*, 51, 231-240.
- Wilkinson, G. S. (1990). Food sharing in vampire bats. *Scientific American*, 262, 76-82.
- Williams, A. C. d. C. (2002). Facial expression of pain: an evolutionary account. *Behavioral and Brain Sciences*, 25, 439-488.
- Winefield, H. R., Katsikitis, M., Hart, L. M. & Rounsefell, B. F. (1990). Postoperative pain experiences. *Journal of Psychosomatic Research*, 34, 543-552.
- Zalon, M. L. (1993). Nurses' assessment of postoperative patients' pain. *Pain*, 54, 329-334.
- Zborowski, M. (1952). Cultural components in responses to pain. *Journal of Social Issues*, 8, 16-30.

APPENDIX I: Introduction

I.1 Excluded studies in the clinical context

Authors	Sample	Research question	Pain intensity measure ¹ and time of measure	Relevant results	Comments
(1) Au, Loprinzi, Dhodapkar, Nelson, Novotny, Ham-mack & O'Fallon, 1994)	patients: patients with malignant solid tumors judges: primary nurse, resident physician	determine whether correlation of pain perception between patients and caregivers could be improved by requiring caregivers to ask patients to rate their pain daily	measure: NRS 0-10 instruction: verbally rate average pain over past 24 hours time: <u>patient</u> : after morning round; <u>nurse/doctor</u> : on same morning	agreement on: <u>baseline</u> : 64 % <u>follow-up 1</u> : 68 % <u>follow-up 2</u> : 85 %	- agreement defined as within \pm 2 points - proportion of over-, underestimation not reported
(2) Bergh & Sjöström, 1999	patients: on geriatric hospital unit with arthritis and/or nonpathological fracture (n = 39) judges: nurses responsible for each patient (n = 39)	compare elderly patients' and nurses' ratings of pain and pain tolerance	measure: VAS 0-100 instruction: current pain, time: <u>patients</u> : interviewed; <u>nurses</u> : completed questionnaire after-wards	mean: 30 (p), 20 (n) t-test: p = 0.069 correlation: 0.39 (p = 0.02)	- no definition of agreement - proportion of over-, underestimation not reported
(3) Bondestam, Hovgren, Johansson, Jern, Herlitz & Holmberg, 1987	patients: acute myocardial infarction (n = 47) judges: nurses on duty judgement pairs: 127	compare patient's pain assessment with that made by nurse	measure: NRS 0-10 instruction: maximal pain time: <u>patient</u> : at home (retrospectively), every 15 minutes during first hour, then each hour during first 24 hours after admission	correlation: 0.76 (p<0.001) underest.: 23 % overest.: 20 %	- no definition of agreement

Authors	Sample	Research question	Pain intensity measure ¹ and time of measure	Relevant results	Comments
(3) continued			<i>nurse</i> : each round, each time before giving pain relieving treatment; ratings separately from each other		
(4) Clipp & George, 1992	<u>patients</u> : patients with metastatic lung or colon cancer (n = 30) <u>judges</u> : spouse primary caregiver, i.e. husband or wife (n = 30) <u>judgement pairs</u> : 29 (aver.), 28 (worst)	focus on reciprocal influences operating in caregiver-patient dyads to examine reliability of spouse informants	<u>measure</u> : ? <u>instruction</u> : worst pain, average daily pain during last month <u>time</u> : separate interviews with patients and spouses at same time	<u>underestimation</u> : <u>average</u> : 24 %, <u>worst</u> : 21 % <u>overestimation</u> : <u>average</u> : 28 %, <u>worst</u> : 39 % <u>correlation</u> : <u>average</u> : 0.55 <u>worst</u> : 0.36	- no definition of agreement
(5) Cohen-Mansfield, 2002	<u>patients</u> : cognitively impaired nursing home residents <u>judges</u> : relatives (children, spouses, siblings, other relatives)	examine relatives' evaluations of pain in nursing home residents suffering from dementia	<u>measure</u> : VRS 0-5 <u>instruction</u> : pain during past two weeks <u>time</u> : patients were interviewed, relatives were sent questionnaire	<u>correlation</u> : 0.49 (p<0.05)	- agreement defined by sig. correlation coefficients
(6) Curtis & Fernsler, 1989	<u>patients</u> : oncology patients in hospice (n = 23) <u>judges</u> : primary caregiver (spouse, child, child-in-law, other relative, friend, housekeeper-companion; n = 23) <u>judgement pairs</u> : 23	determine differences between patients' and primary caregivers' views of oncology hospice patients' quality of life	<u>measure</u> : VAS (QLI: Padilla, Present, Grant, Metter, Lipsett & Heide, 1983) <u>instruction</u> : pain during last week <u>time</u> : patient and judge were seen at the same time, filled in questionnaire in presence of investigator	<u>means</u> : 75.9 (p), 63.0 (j) <u>t-test</u> : patients reported sig. less pain than caregivers (p=0.03)	- agreement defined by results of t-test

Authors	Sample	Research question	Pain intensity measure ¹ and time of measure	Relevant results	Comments
(7) Dar, Beach, Barden & Cleeland, 1992	<p>patients: cancer patients (n = 40)</p> <p>judges: spouses (n = 40)</p> <p>judgement pairs: 40</p>	examine cancer pain in the context of patients and their spouses	<p>measure: NRS 0-10 (BPI: Cleeland & Ryan, 1994; Daut, Cleeland & Flannery, 1983)</p> <p>instruction: worst, least and average pain</p> <p>time: patient and spouse were interviewed separately by two investigators</p>	<p>means:</p> <p>worst: 5.45 (p), 5.50 (j)</p> <p>average: 3.6 (p), 4.3 (j)</p> <p>t-test: no sig. differences between groups</p>	<p>- agreement defined as result of t-test</p>
(8) van der Does, 1989	<p>patients: burn patients (n = 17)</p> <p>judges: nurses on ward (n = 30)</p> <p>judgement pairs: 126, in 19 cases 2 nurses, i.e. 145 judgements</p>	<p>provide more data on extent of pain suffered by burn patients during wound care procedures</p>	<p>measure: VAS</p> <p>instruction: worst, overall pain</p> <p>time: after dressing changes every morning; patients and nurses rated pain independently</p>	<p>means</p> <p>3.22 (p), 3.79 (n)</p> <p>agreement:</p> <p>worst: 31.5 %</p> <p>overall: 37.7 %</p> <p>mean abs. diff.:</p> <p>worst: 2.53</p> <p>overall: 2.07</p> <p>correlation:</p> <p>worst: 0.36 (p<0.001)</p> <p>overall: 0.38 (p<0.001)</p>	<p>- agreement defined as ± 1 cm</p> <p>- proportion of over-, underestimation not reported</p>
(9) Drayer, Henderson & Reidenberg, 1999	<p>patients: acutely hospitalised patients (cancer, sickle cell disease, miscellaneous lesions, no identifiable lesions) (n = 50)</p> <p>judges: nurses, house staff physicians caring for patients</p>	<p>clarify the barrier to improving pain control that health care professionals are often aware of patients' pain perceptions but think that patients overstate its intensity</p>	<p>measure: 0-10 (which scale?)</p> <p>instruction: not explicitly stated; current pain? time: judges within 24 hours after patient (for 48 of 50 patients)</p>	<p>mean difference:</p> <p>0.18 (p-s); 0.58 (p-j)</p> <p>(difference = patient - judge)</p>	<p>- no definition of agreement</p> <p>- no proportion of agreement, over-, underestimation provided</p>

Authors	Sample	Research question	Pain intensity measure and time of measure	Relevant results	Comments
(10) van Dulmen, Fennis, Mokkink, van der Velden & Bleijenberg, 1994	patients: functional abdominal complaints/irritable bowel syndrome (n = 120) judges: doctors in outpatient clinic for internal medicine (n = 13)	explore extent to which doctors are able to perceive patients' cognitions and emotions	measure: NRS 1-5 (dichotomised: more and less than moderate) instruction: ? time: <u>patient</u> : before first consultation of outpatient clinic <u>judge</u> : immediately after each consultation	agreement: 60 % underest.: 32 % overest.: 8 % sign test: doctors sig. underestimated pain severity (p=0.01) <u>kappa</u> : 0.15 (p>0.01)	- agreement defined as result of sign test
(11) Elliott, Murray, Braun & Johnson, 1996	<u>patients</u> : cancer patients (n = 122) <u>judges</u> : family members (n = 122)	investigate relation of specific knowledge and attitudes and patients' and family members' reports of pain due to cancer	<u>measure</u> : NRS 0-10 (BPI: Cleeland & Ryan, 1994; Daut et al., 1983) <u>instruction</u> : current pain, worst, least and average pain during last week; judges without current pain <u>time</u> : separate interviews, patient first	<u>mean</u> : (averaged over average, worst, least pain) 9.9 (j), 8.7 (p) <u>correlation</u> : 0.67 (p=0.0001) t-test: no sig. differences between patients' and family members' ratings (p>0.05)	- agreement defined as result of t-test
(12) Ferrell, Rhiner, Cohen & Grant, 1991	<u>patients</u> : cancer patients (n = 85) <u>judges</u> : family member (parent, spouse, child, other; n = 85)	enhance understanding of social concerns domain of quality of life model	<u>measure</u> : 0-100 (which scale) <u>instruction</u> : during past few days <u>time</u> : patients and judges interviewed separately	<u>means</u> : 70 (j), 45.5 (p)	- no definition of agreement
(13) Forrest, Hermann & Andersen, 1989	<u>patients</u> : acute abdominal pain patients (n = 50) <u>judges</u> : doctors who had taken case history and examined patient (n = 8)	compare pain estimates by doctors and patients who experience pain	<u>measure</u> : VAS <u>instruction</u> : current <u>time</u> : patient and judge independent of each other after examination	<u>medians</u> : 3.4 (j), 6.1 (p) <u>Wilcoxon</u> : doctors rated pain sig. lower than patients (p<0.05) <u>correlation</u> : 0.64 (p<0.05)	- agreement defined as result of Wilcoxon test

Authors	Sample	Research question	Pain intensity measure ¹ and time of measure	Relevant results	Comments
(14) Geisser, Bingham & Robinson, 1995	<p><u>patients</u>: burn patients (n = 11)</p> <p><u>judges</u>: nurses</p> <p><u>judgement pairs</u>: 107</p>	readdress relationship between patients' and nurses' pain ratings and anxiety before and during burn dressing changes.	<p><u>measure</u>: VAS</p> <p><u>instruction</u>: average, maximum pain level during dressing change</p> <p><u>time</u>: immediately after dressing change</p>	<p><u>means</u>:</p> <p><u>average</u>: 4.4 (p), 3.7 (j)</p> <p><u>worst</u>: 4.9 (p), 4.7 (j)</p> <p><u>agreement</u>:</p> <p><u>average</u>: 25.2 %</p> <p><u>worst</u>: 27.1 %</p> <p>mean abs. diff.:</p> <p><u>average</u>: 2.4</p> <p><u>worst</u>: 2.7</p>	<ul style="list-style-type: none"> - agreement defined as ± 1 cm - proportion of over-, underestimation not reported
(15) Gohar, Lunenfeld, Poshnik & Glezerman, 1993	<p><u>patients</u>: women undergoing oocyte pickup (n = 57)</p> <p><u>judge</u>: gynaecologist present at OPU but not participating in it</p>	evaluate patient's comfort while undergoing oocyte pickup as assessed by patient and observing physician	<p><u>measure</u>: VAS</p> <p><u>instruction</u>: pain during oocyte pickup</p> <p><u>time</u>:</p> <p><u>patients</u>: 8-10 hours after oocyte pickup</p> <p><u>judges</u>: ?</p>	<p><u>means</u>:</p> <p>4.7 (p), 3.0 (j)</p> <p>t-test: judge rated pain sig. lower than patients (p<0.05)</p>	<ul style="list-style-type: none"> - agreement defined by result of t-test
(16) Graffam, 1981	<p><u>patients</u>: acute pain (n = 75), chronic pain patients (n = 25)</p> <p><u>judges</u>: nurses (n = 51)</p> <p><u>judgement pairs</u>: 100</p>	examine whether expectations of nurses and patients differ	<p><u>measure</u>: ?</p> <p><u>instruction</u>: current pain</p> <p><u>time</u>: at time of pain report and 30-60 minutes following relief</p>	?	<ul style="list-style-type: none"> - no definition of agreement
(17) Grossman, Sheidler, Swedeon, Mucenski & Piantadosi, 1991	<p><u>patients</u>: cancer inpatients (n = 103)</p> <p><u>judges</u>: nurse, house officer, medical oncology fellow of patient</p> <p><u>judgement pairs</u>: 103</p>	determine if cancer patients and their doctors and nurses perceive patient's pain to be similar in intensity	<p><u>measure</u>: VAS (1-10)</p> <p><u>instruction</u>: average pain since admission</p> <p><u>time</u>:</p> <p><u>patients</u>: within 48 hours of admission to hospital</p> <p><u>judges</u>: on same day as patient</p>	<p><u>means</u>:</p> <p>3.4 (p), 2.4 (n), 3.4 (ho), 3.0 (of), 2.9 (average all j)</p> <p><u>correlations</u>:</p> <p><u>nurses</u>: 0.46</p> <p><u>house off</u>: 0.35</p> <p><u>onc fell</u>: 0.42</p> <p>(all p's=0.0001)</p>	<ul style="list-style-type: none"> - agreement defined by sig. correlation coefficients

Authors	Sample	Research question	Pain intensity measure ¹ and time of measure	Relevant results	Comments
(18) Hall-Lord, Larsson & Steen, 1998	patients: 65 years and older in intensive care unit (n = 51) judges: nurses (n = 44), assistant nurses (n = 37) responsible for patient on day of data collection	investigate intensive care unit patients' experiences of pain and distress and compare these experiences with the way nurses assess patients' responses related to these issues	measure: NRS (range?) instruction: on day of interview (d), during whole period of hospitalization (wp) time: on same day patients were interviewed and separately nurses filled out questionnaire	means patients: 2.35 (d), 3.13 (wp) nurses: 2.06 (d), 2.87 (wp) ass. Nurses: 1.91 (d), 2.48 (wp) Kruskal-Wallis: non sig. Differences	- agreement defined as result of Kruskal-Wallis
(19) Harrison, 1993	patients: medical (n = 85), surgical (n = 89) and paediatric patients (n = 25) judges: nurses on duty responsible for patient judgement pairs: 199, 197 for worst pain	investigate agreement shown by nurses and patients from general medical, surgical and paediatric wards of a district hospital in Kuwait	measure: NRS 0-10 instruction: current pain, worst and least pain during last 24 hours time: patients and judges were interviewed independently of each other	means: current: 3.13 (p), 1.93 (n) worst: 5.99 (p), 3.76 (n) least: 2.15 (p), 1.29 (n) Kruskal-Wallis: current: $\chi^2=34.49$ worst: $\chi^2=36.44$ least: $\chi^2=21.96$ (all p's<0.05)	- agreement defined by results of Kruskal-Wallis
(20) Hodgkins, Albert & Daltroy, 1985	patients: rheumatology patients having needle aspiration/injection (n = 21) judges: physicians doing this	compare pain assessments by patients and physicians to evaluate patient and physician rating differences; investigate how well physicians predict experienced pain by their patients	measure: VAS instruction: before: predict pain after: rate pain caused by procedure time: before and after needle aspiration/injection; independent ratings	means: before: 4.76 (p), 4.15 (j) after: 4.01 (p), 3.36 (j) correlation: before: 0.54 (p<0.05) after: 0.62 (p<0.01) ANOVA: non sig. Differences between groups and for time	- agreement defined by results of ANOVA

Authors	Sample	Research question	Pain intensity measure ¹ and time of measure	Relevant results	Comments
(21) Krivo & Reidenberg, 1996	<u>patients</u> : admitted to hospital in pain (due to cancer, AIDS, sickle cell, other reasons; n = 48) <u>judges</u> : physicians (n = 48) and nurses (n = 48) responsible for patient	test hypothesis that medical staff do not know pain intensity of patient	<u>measure</u> : 0-10 (which scale?) <u>instruction</u> : current pain severity <u>time</u> : day after admission	<u>means</u> : judges are on average 2.1 (doctors) / 1.8 (nurses) units lower than patients <u>χ^2-test</u> : sig. Less physicians rated pain within 1 unit of patients	- agreement defined as ± 1 unit on scale - proportion of over-, underestimation not reported
(22) Krokosky & Reardon, 1989	<u>patients</u> : patients in acute care hospital (n = 50) <u>judges</u> : nurses and doctors	establish the problem that there are deficiencies in pain relief	<u>measure</u> : VRS 1-5 (Melzack & Katz, 2001) <u>instruction</u> : ? <u>time</u> : ?	<u>means</u> : 3.12 (p), 2.64 (n), 2.48 (d) <u>correlation</u> : <u>pat-nrs</u> : -0.01 (ns) <u>pat-drs</u> : 0.30 (ns) <u>ANOVA</u> : sig. Effect for groups (p<0.01); post-hoc analysis: sig. Differences between pat. And each group of judges, but not between doctors and nurses	- agreement defined by results of ANOVA
(23) Marquié, Raufaste, Lauque, Mariné, Ecoiffier & Sorum, 2003	<u>patients</u> : presenting with pain in emergency department (n = 200) <u>judges</u> : physicians examining patients (n = 48)	confirm miscalibration between pain ratings of patients and physicians and study factors associated with it	<u>measure</u> : VAS 10 cm <u>instruction</u> : current pain <u>time</u> : <u>patient</u> : on arrival, before examination, when leaving hospital; <u>physicians</u> : after examination, when patient left hospital; patients' pain ratings on arrival were accessible to physicians	<u>mean</u> : <u>on arrival</u> : 5.02 (p), 3.70 (j) <u>at exit</u> : 3.94 (p), 2.56 (j) <u>mean difference</u> : - 1.42 (difference = physician - patient) <u>t-test</u> : t = -9.67, (p<0.001)	- no definition of agreement - no proportion of agreement, over-, underestimation reported

Authors	Sample	Research question	Pain intensity measure ¹ and time of measure	Relevant results	Comments
(24) McKinley & Botti, 1991	patients: randomly selected from medical, surgical, critical care, rehabilitation, obstetrics and other wards in one hospital (n = 115) judges: nurses responsible for patient that day (n = 115) judgement pairs: 115	compare prevalence of pain in patients and ratings of pain identified by nurses; determine relationship between patients' and nurses' ratings of pain intensity	measure: VAS instruction: current pain time: patient before nurse; on same day; data collected separately	prevalence: 63 % (p), 84 % (n) medians: 28 (p), 20.5 (n) correlation: 0.35 (p<0.001) χ^2 -test: nurses reported sig. more frequently that their patients were in pain than patients (p<0.001)	- agreement defined as result by χ^2 -test
(25) McMillan, 1996	patients: hospice cancer patients (n = 48) judges: primary caregivers (spouse, child, other relative, friend; n =48) judgement pairs: 48	compare hospice patients pain report on admission and after three weeks; investigate extent to which caregivers can report patients' pain	measure: NRS 1-10 (item added to HQLI: McMillan & Mahon, 1994) instruction: worst pain time: interviews within 48 hours after admission in separate rooms and three weeks later	means: on admission: 7.1 (p), 6.0 (j) three weeks later: 7.0 (p), 6.1 (j) correlation: on admission: 0.58 three weeks later: 0.31	- no definition of agreement
(26) McMillan & Mahon, 1994	patients: oncology hospice patients (n = 68) judges: caregivers of patients (spouse, child, parent, other relative, friend, nursing assistant; n = 68) judgement pairs: 44 (for pain ratings)	evaluate validity and reliability of the Hospice Quality of Life Index	measure: VAS (item added to HQLI: McMillan & Mahon, 1994) instruction: worst pain time: within 48 hours after admission to hospice	means: 59.7 (p), 68.4 (j) correlation: 0.42 (p<0.005) t-test: ns difference between patients and judges (p=0.2)	- agreement defined by results of t-test
(27) O'Brien & Francis, 1988	patients: cancer patients (n = 42) judges: next-of-kin (n = 42) judgement pairs: 42	investigate validity of proxy estimation of subjective experiences	measure: VRS (0-4), VRS (Meizack & Katz, 2001)	means (VRS 0-4): 0.90 (p), 0.98 (j)	- agreement defined as sig. Kappa coefficient

Authors	Sample	Research question	Pain intensity measure ¹ and time of measure	Relevant results	Comments
(27) continued			<p><u>instruction</u>: typical, worst and least pain in week before interview with patient</p> <p><u>time</u>: interview with patient first, interviews with next-of-kin 30 days (median) later</p>	<p><u>kappa</u>: <u>VRS 0-4</u>: 0.32-0.37 <u>VRS</u>: 0.24-0.22</p>	<p>– large time gap between patient and next-of-kin interview</p>
(28) Olden, Jordan, Sakima & Grass, 1995	<p><u>patients</u>: patients undergoing caesarean section (n = 20)</p> <p><u>judges</u>: nurses who were primary caregivers for patients</p>	<p>compare nurses' and patients' pain and sedation assessments in patients receiving epidural or intravenous PCA</p>	<p><u>measure</u>: VAS</p> <p><u>instruction</u>: current pain time: 2 and 6 hours after surgery, 9am and 5 pm on 1st and 2nd day after surgery</p>	<p><u>average diff.</u>: 20 ± 2 mm</p> <p><u>overest.</u>: 55 %</p> <p><u>underest.</u>: 43 %</p> <p><u>correlation</u>: 0.08 (ns)</p>	<p>– no definition of agreement</p>
(29) Paice, Mahon & Faut-Callahan, 1991	<p><u>patients</u>: postsurgical patients diagnosed with cancer (n = 34)</p> <p><u>judges</u>: nurses and physicians</p>	<p>examine pain experience of surgical oncology patients and explore the relationship between patient's and nurse's and physician's pain assessments</p>	<p><u>measure</u>: NRS 0-10</p> <p><u>instruction</u>: ?</p> <p><u>time</u>: patient interviewed 1-31 days after surgery, judges within one hour of patient's interview</p>	<p><u>mean</u>: 3.0 (p)</p> <p><u>correlation</u>: <u>pat-phys</u>: 0.12 <u>pat-nur</u>: 0.06 <u>phys-nur</u>: 0.37 (all ns)</p>	<p>– agreement defined by sig. correlation coefficient</p>
(30) Peteet, Tay, Cohen & MacIntyre, 1986	<p><u>patients</u>: patients with cancer related pain (n = 25)</p> <p><u>judges</u>: physicians of these patients (n = 25)</p>	<p>obtain more information about pain treatment and reasons for inadequate pain relief in individual cases</p>	<p><u>measure</u>: 0-5 (which scale?)</p> <p><u>instruction</u>: pain during last week; judges: same week as patient time; physician was interviewed 1-31 days (61.5 % within 7 days) after patient interview</p>	<p><u>agreement</u>: 59 %</p> <p><u>underest.</u>: 36 %</p> <p><u>overest.</u>: 5 %</p> <p><u>mean difference</u>: 0.73</p>	<p>– agreement defined as ± 1 unit on scale</p> <p>– big time gap between patient and physician interview</p>

Authors	Sample	Research question	Pain intensity measure ¹ and time of measure	Relevant results	Comments
(31) Rae, Gallagher, Watson & Kinsella, 2000	<u>patients</u> : burn patients (n=30) <u>judges</u> : plastic surgeons (n=21), nurses (n=32) working with burn patients	attempt to quantify severity of pain suffered during burn dressing changes in burn unit and assess patient satisfaction with analgesia provided	<u>measure</u> : VAS (p), VRS (0-3; p, j) <u>instruction</u> : worst pain during dressing change <u>time</u> : <u>patients</u> : during dressing change <u>judges</u> : interviewed separately; not asked about specific patients	<u>VRS</u> : <u>patients</u> : majority rated pain as 0 or 1; <u>judges</u> : majority rated pain as 2 or 3	- no definition of agreement - proportion of over-, underestimation not reported
(32) Rafferty, Smith-Coggins & Chen, 1995	<u>patients</u> : patients presenting in Emergency Department (ED) with headache, neck or back pain (n=190) <u>judges</u> : medical students, interns residents, nurse practitioners, attending physicians (n=84)	determine whether gender of patient or health care provider influences analgesic practice in ED	<u>measure</u> : VAS <u>instruction</u> : judges were asked about the intensity of pain reported by patient and about their perception of patients' pain intensity <u>time</u> : immediately after patient was discharged	<u>means</u> : <u>patients</u> : 68.5±20.5 mm <u>judges</u> : 55.5±20.5 mm	- no definition of agreement - proportion of over-, underestimation not reported
(33) Redinbaugh, Baum, Demoss, Fello & Arnold, 2002	<u>patients</u> : home hospice patients (n = 31) <u>judges</u> : family caregivers (71 % spouses) (n = 31)	investigate caregiver perception of patients' pain specifically examining relationships between perceptions of pain and quality of life	<u>measure</u> : 0-10 (which scale?) (p), FPQ: Ferrell, Rhiner & Rivera, 1993 (j) <u>instruction</u> : current pain <u>time</u> : data collected separately	<u>mean</u> : 3.4 (p), 6.1 © <u>t-test</u> : -4.9 (p<0.01)	- agreement defined as result of t-test - no proportion of agreement, over-, underestimation reported
(34) Salmon & Manyande, 1996	<u>patients</u> : undergoing minor abdominal/rectoanal surgery (n = 56) <u>judges</u> : nurses in charge conducting relevant drug rounds (n = 15)	Investigate whether analgesic administration frequency is related to patients' feelings of inability to cope with pain (or to nurses' estimates of their inability to cope) rather than to pain intensity	<u>measure</u> : VAS <u>instruction</u> : current pain <u>time</u> : <u>patient</u> : afternoon before surgery and on two postoperative days before drug round;	<u>correlations</u> : <u>day 1</u> : 0.58 (p<0.01) <u>day 2</u> : 0.66 (p<0.01) ANOVA: no group effect, i.e. ns difference between patients' and nurses' ratings (F<1.0)	- agreement defined by results of ANOVA

Authors	Sample	Research question	Pain intensity measure ¹ and time of measure	Relevant results	Comments
(34) continued		examine whether nurses' perception of patients' pain is inaccurate in that it underestimates patients' ability to cope with it	<u>judge</u> : when recruited and on two postoperative days during drug round; data collected separately		
(35) Schuler, Neuhauser, Hauer, Oster, Razus & Hacker, 2001	<u>patients</u> : geriatric patients (n=126) <u>judges</u> : doctors, nurses, physiotherapists, logopaedics, occupational therapists responsible for patient (n=59)	examine work-related variables which impact on agreement and certainty of judgement	<u>measure</u> : pain present or absent <u>instruction</u> : current pain, pain during previous week <u>time</u> : patients and judges were interviewed between 11 am and 1 pm on the same day	<u>agreement</u> : <u>current</u> : 21.9-97.7 % <u>previous week</u> : 37.1-73.3 %	- exact agreement in terms of dichotomous presence/absence of pain - proportion of over-, underestimation not reported
(36) Schulte, Marvin & Sandidge, 1987	<u>patients</u> : burn patients <u>judges</u> : burn nurses <u>judgement pairs</u> : 324	examine whether nurses can assess pain	<u>measure</u> : VAS <u>instruction</u> : overall pain; worst pain patients only <u>time</u> : after wound care; data collected separately	<u>agreement</u> : 53.4 % <u>t-test</u> : ns difference (p>0.28)	- agreement defined as within \pm 1 cm - proportion of over-, underestimation not reported
(37) Seers, 1987	<u>patients</u> : admitted for elective abdominal surgery (n = 80) <u>judges</u> : nurses	investigate whether pain does matter, whether it does influence recovery or anxiety and whether there are differences between nurses' and patients' ratings of pain	<u>measure</u> : VRS (10 cm long) <u>instruction</u> : ? <u>time</u> : interviews before surgery, twice a day after surgery for seven days	<u>agreement</u> : 23 % <u>underest.</u> : 54 % <u>overest.</u> : 13 %	- no definition of agreement - no instruction for scale reported

Authors	Sample	Research question	Pain intensity measure¹ and time of measure	Relevant results	Comments
(38) Sneeuw, Aaronson, Osoba, Muller, Hsu, Yung, Brada & Newlands, 1997	patients: brain cancer patients (n = 103) judges: proxies (spouses, parents, children, siblings, friends; n = 103) judgement pairs: 103	examine level of agreement between health-related quality of life ratings provided by patients and significant others	measure: VRS 1-4 (QLQ-C30: Aaronson et al., 1993) instruction: pain and interference of pain with daily life during past week time: initial clinic visit, one week later and clinic visit at least 4 weeks after initial visit	means: (transformed to 0-100 scale) 14.9 (p), 18.3 (j) exact agreement: 62.6 % approx. agreement: 89.3 %	- exact agreement: same response category chosen - approximate agreement: ± 1 category
(38) continued					- proportion of over-, underestimation not reported
(39) Stein & Miech, 1993	patients: hospice cancer patients (n=239) judges: admission and primary nurses	assess relationship between subjective pain assessment and other clinical variables	measure: NRS 0-5 (p); pain present/absent (j) instruction: current pain (?) time: admission, subsequent visits	agreement: 84 % for patients not reporting pain, 93 % for patients reporting pain false positive: 6 % false negative: 3 %	- agreement defined as matching categories (presence/absence of pain)
(40) Stephenson, 1994	patients: postsurgical inpatients (n = 25) judges: nurses (n = 11) judgement pairs: 23	comparison of nurses' and patients' perceptions of postsurgical pain and pain relief	measure: VAS, VRS 0-5 (Melzack & Katz, 2001) instruction: current pain time: when patient complained of pain during first 24 hours after surgery (t1) and 45 min after complaint (t2)	means: VAS: 5.67 (p), 4.63 (j) VRS: 2.87 (p), 2.43 (j) correlation: VAS : 0.04-0.39 (ns) VRS: 0.47-0.59 (sig.) t-test: nurses rate pain as lower on both scales at t2 (p = 0.025)	- agreement defined by results of t-tests
(41) Striebel, Hackenberger & Wessel, 1992	patients: postoperative patients (hysterectomy; n = 30) judges: nurse (n = 1) judgement pairs: 240	differences in pain evaluations between patient and observer	measure: VAS instruction: current pain time: 10, 20, 30, 40, 60, 100, 120, 240 minutes after awakening from anaesthesia	Mann-Whitney-U-Test: judges' rating sig. lower on each measurement time than patients' ratings (p<0.0001)	- agreement defined by results of Mann-Whitney-U-Test

Authors	Sample	Research question	Pain intensity measure ¹ and time of measure	Relevant results	Comments
(42) Sutherland, Wesley, Cole, Nesvacil, Daly & Gepner, 1988	<p>patients: presenting in family practice clinic (n = 401)</p> <p>judges: physicians</p>	measuring similarities and differences between patient and physician ratings of patient pain	<p>measure: NRS 0-10</p> <p>instruction: current pain time:</p> <p>patients: before seeing physician</p> <p>judges: while seeing/ after having seen patient</p>	<p>means: 4.37 (p), 3.52 (j)</p> <p>correlation: 0.66 (p=0.001)</p> <p>t-test: judges' ratings sig. lower than patients' (p<0.001)</p>	<p>– agreement defined by results of correlation and t-tests</p> <p>– physicians were allowed to see rating of patient</p>
(43) Teske, Daut & Cleeland, 1983	<p>patients: surgical pain patients (n = 34), chronic pain patients (n = 37)</p> <p>judges: 2 nurses saw first 21 patients, then 1 nurse judgement pairs: 92</p>	investigation of relationship of patients' pain self-report to observer's inferences about pain based on observations of non-verbal behaviours	<p>measure: VAS (p, n), NRS 0-6 (n)</p> <p>instruction: VAS: current pain; NRS: global pain time: patients after nurse; data collected separately</p>	<p>correlations: VAS-VAS: 0.38 MRS-VAS: 0.41</p> <p>t-test: nurses rated pain as sig. lower than did chronic patients (p<0.001), but not than ratings by acute pain patients</p>	<p>– agreement defined as result of sig. correlation coefficient</p>
(44) Thomas, Borczuk, Shackelford, Ostrander, Silver, Evans & Stein, 1999	<p>patients: presenting in emergency department with severe (≥ 5 on VAS) undifferentiated abdominal pain (n = 30)</p> <p>judges: physicians working in emergency department (n = ?)</p>	determine level of agreement between patients and physicians as to severity of abdominal pain	<p>measure: VAS</p> <p>instruction:</p> <p>time: on arrival, 1 and 2 hours afterwards; data collected separately</p>	<p>means: arrival: 7.5 (p), 5.3 (j) 1 hour: 6.7 (p), 4.7 (j) 2 hours: 5.1 (p), 3.9 (j)</p> <p>t-tests: p < 0.0001, p = 0.002, p = 0.013</p>	<p>– agreement defined as ± 1.3 cm</p> <p>– no proportions of agreement, over-, underestimation reported</p>
(45) Thomas, Robinson, Champion, Mckell & Pell, 1998	<p>patients: undergoing elective surgery (n = 91)</p> <p>judges: nurses (n = 4)</p>	identify factors which might correlate with and potentially predict severe post-operative pain and dissatisfaction with analgesic management	<p>measure: VRS 0-5 (Melzack & Katz, 2001)</p> <p>instruction: current pain time: 3 occasions at random times on days 1-5</p>	<p>means: day 1: 2.3 (p), 0.5 (j) day 2: 1.8 (p), 0.6 (j) day 3: 1.6 (p), 0.4 (j) day 4: 1.6 (p), 0.4 (j) day 5: 1.4 (p), 0.3 (j)</p>	<p>– no definition of agreement</p> <p>– proportion of over-, underestimation not reported</p>

Authors	Sample	Research question	Pain intensity measure ¹ and time of measure	Relevant results	Comments
(46) Todd, Lee & Hoffman, 1994	patients: presenting with isolated extremity trauma in ER (n = 207) <u>judges</u> : physicians (n = 65) <u>judgement pairs</u> : 207	investigate the effect of ethnicity on physician estimates of pain severity	<u>measure</u> : VAS <u>instruction</u> : current pain <u>time</u> : <u>patient</u> : after consenting; <u>judge</u> : after first pain evaluation	<u>means</u> : 39.4 (p), 31.65 (j) <u>t-test</u> : patients' ratings were sig. higher than physicians' ratings (p's = 0.005)	- agreement defined as results of t-test
(47) Walkenstein, 1982	patients: burn patients (n = 15) <u>judges</u> : burn nurses (n = 8) <u>judgement pairs</u> : 45	determine how burned patients perceive their pain in comparison to how nursing staff members perceive the patients' pain	<u>measure</u> : Stewart Pain-Color Scale (Stewart, 1977) (range: 0-10) <u>instruction</u> : <u>time</u> : 3 randomly selected times; nurses were asked immediately after patients	<u>means</u> : 2.88 (p), 2.84 (j) <u>correlation</u> : 0.44 (p<0.0) <u>Wilcoxon-test</u> : ns differences	- agreement defined as result of Wilcoxon test
(48) Weiner, Ladd, Pieper & Keefe, 1995	<u>patients</u> : residents in nursing home (n = 93) <u>judges</u> : nurses in nursing home <u>judgement pairs</u> : 81	quantitatively assess staff awareness of resident pain complaints in a nursing home	<u>measure</u> : NGRS 0-10 <u>instruction</u> : average pain intensity during past month <u>time</u> : simultaneously, but without nurse asking resident	<u>median</u> 2.5 (p) <u>correlation (NGRS)</u> : 0.07 (ns)	- agreement defined by sig. correlation coefficient
(49) Weiner, Peterson & Keefe, 1999	patients: residents in nursing home with chronic pain (n = 42). <u>judges</u> : family members visiting at least twice a month (n = 42), primary nurse (n = 42)	confidence of nursing home and family caregivers to accurately assess pain and that residents display a sufficient amount of pain behaviours to allow for assessment	<u>measure</u> : NGRS 0-10, vertical verbal descriptor scale (thermometer) <u>instruction</u> : average pain over past two week	<u>medians</u> : 6 (p; over scales) <u>NGRS</u> : 4 (n), 4 (f) <u>therm</u> : 1.7 (n), 1.7 (f)	- agreement defined as sig. correlation coefficient

Authors	Sample	Research question	Pain intensity measure ¹ and time of measure	Relevant results	Comments
(49) continued	<p><u>judgement pairs</u>: 42 pat.-nurse, 42 pat.-fam. memb., 42 nurse-fam. memb.</p>	agreement between caregivers (nurse or family) and residents as well as both groups of caregivers regarding pain behaviour and pain intensity	<p><u>time</u>: <u>family members</u>: during telephone interview; <u>nurses</u>: during structured interview</p>	<p><u>correlation</u>: <u>res-fam</u>: - 0.19 (NGRS; ns), - 0.01 (therm; ns) <u>res-nurs</u>: 0.34 (NGRS; ns) 0.04 (therm; ns)</p>	
(50) Werner, Cohen- Mansfield, Watson & Pasis, 1998	<p><u>patients</u>: senior day care participants (n = 130) <u>judges</u>: family caregivers, day care staff members</p>	evaluating presence of pain in elderly day care participants and assessing the concordance among different raters	<p><u>measure</u>: VRS (1-6) <u>instruction</u>: <u>patients</u>: mean score created from current pain, worst and least pain during last two weeks <u>judges</u>: average pain during last two weeks <u>time</u>: approximately same time (up until one month)</p>	<p><u>means</u>: 1.5 (p), 1.9 (fc), 1.6 (sm) <u>agreement</u>: between 63.4 and 69.4 % <u>kappa</u>: 0.32-0.40</p>	<p>- exact agreement in terms of dichotomous presence/absence of pain</p>
(51) Yeager, Miaskowski, Dibble & Wallhagen, 1995	<p><u>patients</u>: cancer outpatients (n = 86) <u>judges</u>: caregivers (spouse, partner, other; n = 86) <u>judgement pairs</u>: 86</p>	determine if knowledge about pain and perception of pain experience differ when comparing outpatients with their family caregivers	<p><u>measure</u>: VAS <u>instruction</u>: current pain <u>time</u>: patients and judges completed questionnaire at the same time</p>	<p><u>means</u>: 41.7 (p), 51.2 (f) <u>t-test</u>: caregivers reported sig. more pain than patients (p=0.0002)</p>	<p>- agreement defined as result of t-tests</p>

¹ NRS: numerical rating scale; VAS: visual analogue scale; VRS: verbal rating scale; NGRS: numerical graphic rating scale

I.2 Scales measuring pain intensity

There are three commonly used approaches to quantitatively estimate the severity of pain: Verbal Rating Scales (VRS), Visual Analogue Scale [VAS; including Graphic Rating Scales (GRS)] and Numerical Rating Scales (NRS) (Jensen & Karoly, 2001). Some of them are included in broader pain questionnaires as for instance in the Brief Pain Inventory (BPI: Cleeland & Ryan, 1994; Daut et al., 1983) or the McGill Pain Questionnaire (MPQ: Melzack & Katz, 2001), others are included in questionnaires measuring quality of life. In the following I will not describe any of the questionnaires (their references are given in the prior appendix), but only the pain intensity measure used (VRS, VAS, NRS).

1. Verbal Rating Scales (VRS)

VRSs consist of lists of words or phrases describing various levels of pain intensity. Participants are asked to read over the list and select the word or phrase that best describes their level of pain. VRSs are commonly scored by ordering the words according to the pain intensity they express and assigning a number to each word, with higher numbers expressing higher pain intensities. Choiniere and colleagues (Choiniere, Melzack, Girard, Rondeau & Paquin, 1990), for example, in their study use a VRS consisting of five adjectives/phrases: no pain, mild, moderate, severe and unbearable pain. The words are scored from 0 (no pain) to 4 (unbearable pain). Also the Present Pain Intensity (PPI) scale in the McGill Pain Questionnaire (Melzack & Katz, 2001) is a VRS with the following descriptors ranked from 0 to 5: no pain, mild, discomforting, distressing, horrible and excruciating pain. For validity and sensitivity of VRSs compare the chapter by Jensen and Karoly (2001).

Advantages of VRSs are that they are easy to administer and score for investigators as well as easy to comprehend and use for participants. Disadvantages are that participants may not find a word or phrase that best describes their pain and that the method of scoring the pain intensity assumes equal intervals between the single words.

2. Visual Analogue Scales (VAS)

A VAS consists of a line, usually 10 cm long, whose ends are labelled as the extremes of pain, for instance the left end as 'no pain' and the right end as 'pain as bad as it could be'. A Graphic Rating Scale (GRS) is a VAS with specific points along the line which can either be adjectives or numbers. VASs with adjectives along the line are commonly called Verbal Graphic Rating Scale (VGRS), those with numbers Numerical Graphic Rating Scale (NGRS).

No matter which of these specific scales are used, participants are instructed to indicate which point on the line best represents their pain intensity. Pain intensity is then quantified by measuring the distance (in cm or mm) from the left end of the scale to the mark made by participants. For validity and sensitivity of VRSs again compare the chapter by Jensen & Karoly (2001).

Advantages of the VASs, are that they appear to have ratio data quality and that they have a high number of response categories which makes them potentially more sensitive to changes in pain intensity. A disadvantage is that VAS have been shown to be more difficult to understand than other measures of pain intensity for elder people (e.g. Jensen, Karoly & Braver, 1986).

3. Numerical Rating Scales (NRS)

NRSs require participants to rate their pain on scales from 0 to one of various end-points, for example, from 0-10 (11-point scale), 0-20 (21-point scale) or 0-100 (101-point scale) assuming that 0 represents the end of the dimension with no pain, while 10, 20 or 100 represents the end of maximum pain intensity.

NRSs can be visually depicted in various ways. In the Brief Pain Inventory (BPI: Cleeland & Ryan, 1994; Daut et al., 1983), for example, numbers from 0 to 10 are enlisted on a horizontal line with 0 being defined as 'no pain' and 10 as 'pain as bad as it can be'. Participants are asked to rate their pain by circling the one number that best describes their pain intensity. The score is always the number indicated by the participant.

Advantages of NRSs are that they are easy to administer and score as well as simply to comprehend and use. A disadvantage is that they may not have ratio scale qualities. For information on sensitivity and validity see Jensen & Karoly (2001).

I.3 Bibliography for I.1 and I.2

- Aaronson, N. K., Ahmedzai, S., Bergman, B., Bullinger, M., Cull, A., Duez, N. J., Filiberti, A., Flechtner, H., Fleishman, S. B., de Haes, J. C. J. M., Kaasa, S., Klee, M., Osoba, D., Razavi, D., Rofe, P. B., Schraub, S., Sneeuw, K., Sullivan, M. & Takeda, F. (1993). The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *Journal of the National Cancer Institute*, *85*, 365-376.
- Au, E., Loprinzi, C. L., Dhodapkar, M., Nelson, T., Novotny, P., Hammack, J. & O'Fallon, J. (1994). Regular use of a verbal pain scale improves the understanding of oncology inpatient pain intensity. *Journal of Clinical Oncology*, *12*, 2751-2755.
- Bergh, I. & Sjöström, B. (1999). A comparative study of nurses' and elderly patients' ratings of pain and pain tolerance. *Journal of Gerontological Nursing*, *25*, 30-36.
- Bondestam, E., Hovgren, K., Johansson, F. G., Jern, S., Herlitz, J. & Holmberg, S. (1987). Pain assessment by patients and nurses in the early phase of acute myocardial infarction. *Journal of Advanced Nursing*, *12*, 677-682.
- Choiniere, M., Melzack, R., Girard, N., Rondeau, J. & Paquin, M.-J. (1990). Comparisons between patients' and nurses' assessment of pain and medication efficacy in severe burn injuries. *Pain*, *40*, 143-152.
- Cleeland, C. S. & Ryan, K. M. (1994). Pain assessment: global use of the Brief Pain Inventory. *Annals Academy of Medicine Singapore*, *23*, 129-138.
- Clipp, E. C. & George, L. K. (1992). Patients with cancer and their spouse caregivers. Perceptions of illness experience. *Cancer*, *69*, 1074-1079.
- Cohen-Mansfield, J. (2002). Relatives' assessment of pain in cognitively impaired nursing home residents. *Journal of Pain and Symptom Management*, *24*, 562-571.
- Curtis, A. E. & Fernsler, J. I. (1989). Quality of life of oncology hospice patients: a comparison of patient and primary caregiver reports. *Oncology Nursing Forum*, *16*, 49-53.

- Dar, R., Beach, C. M., Barden, P. L. & Cleeland, C. S. (1992). Cancer pain in the marital system: a study of patients and their spouses. *Journal of Pain and Symptom Management*, 7, 87-93.
- Daut, R. L., Cleeland, C. S. & Flannery, R. L. (1983). Development of the Wisconsin Brief Pain Questionnaire to assess pain in cancer and other diseases. *Pain*, 17, 197-210.
- Drayer, R. A., Henderson, J. & Reidenberg, M. (1999). Barriers to better pain control in hospitalized patients. *Journal of Pain and Symptom Management*, 17, 434-440.
- Elliott, B. A., Elliott, T. E., Murray, D. M., Braun, B. L. & Johnson, K. M. (1996). Patients and family members: the role of knowledge and attitudes in cancer pain. *Journal of Pain and Symptom Management*, 12, 209-220.
- Ferrell, B., Rhiner, M. & Rivera, L. M. (1993). Development and evaluation of the Family Pain Questionnaire. *Journal of Psychosocial Oncology*, 10, 21-35.
- Ferrell, B. R., Rhiner, M., Cohen, M. Z. & Grant, M. (1991). Pain as a metaphor for illness. Part I: impact of cancer pain on family caregivers. *Oncology Nursing Forum*, 18, 1303-1308.
- Forrest, M., Hermann, G. & Andersen, B. (1989). Assessment of pain: a comparison between patients and doctors. *Acta Anaesthesiologica Scandinavica*, 33, 255-256.
- Geisser, M. E., Bingham, H. G. & Robinson, M. E. (1995). Pain and anxiety during burn dressing changes: concordance between patients and nurses ratings and relation to medication administration and patient variables. *Journal of Burn Care & Rehabilitation*, 16, 165-171.
- Gohar, J., Lunenfeld, E., Potashnik, G. & Glezerman, M. (1993). The use of sedation only during oocyte retrieval for in vitro fertilization: patients' pain self-assessments versus doctors' evaluations. *Journal of Assisted Reproduction and Genetics*, 10, 476-478.
- Graffam, S. (1981). Congruence of nurse-patient expectations regarding nursing interventions in pain. *Nursing Leadership*, 4, 12-15.
- Grossman, S. A., Sheidler, V. R., Swedeen, K., Mucenski, J. & Piantadosi, S. (1991). Correlation of patient and caregiver ratings of cancer pain. *Journal of Pain and Symptom Management*, 6, 53-57.

- Hall-Lord, M. L., Larsson, G. & Steen, B. (1998). Pain and distress among elderly intensive care unit patients: comparison of patients' experiences and nurses' assessments. *Heart & Lung, 27*, 123-132.
- Harrison, A. (1993). Comparing nurses' and patients' pain evaluations: a study of hospitalised patients in Kuwait. *Social Science and Medicine, 36*, 683-692.
- Hodgkins, M., Albert, D. & Daltroy, L. (1985). Comparing patients' and physicians' assessments of pain. *Pain, 23*, 273-277.
- Jensen, M. P. & Karoly, P. (2001). Self-report scales and procedures for assessing pain in adults. In D. C. Turk & R. Melzack (Eds.), *Handbook of pain assessment* (pp. 15-34). New York: Guilford Press.
- Jensen, M. P., Karoly, P. & Braver, S. (1986). The measurement of clinical pain intensity: a comparison of six methods. *Pain, 27*, 117-126.
- Krivo, S. & Reidenberg, M. M. (1996). Assessment of patients' pain. *New England Journal of Medicine, 334*, 59.
- Krokosky, N. J. & Reardon, R. C. (1989). The accuracy of nurses' and doctors' perceptions of patient pain. In S. G. Funk, E. M. Tornquist, M. T. Champagne, L. A. Cropp & R. A. Wiese (Eds.), *Key aspects of comfort: management of pain, fatigue and nausea* (pp. 127-134). New York: Springer.
- Marquié, L., Raufaste, E., Lauque, D., Mariné, C., Ecoiffier, M. & Sorum, P. (2003). Pain rating by patients and physicians: evidence of systematic pain miscalibration. *Pain, 102*, 289-296.
- McKinley, S. & Botti, M. (1991). Nurses' assessment of pain in hospitalised patients. *The Australian Journal of Advanced Nursing, 9*, 8-14.
- McMillan, S. C. (1996). The quality of life of patients with cancer receiving hospice care. *Oncology Nursing Forum, 23*, 1221-1228.
- McMillan, S. C. & Mahon, M. (1994). Measuring quality of life in hospice patients using a newly developed Hospice Quality of Life Index. *Quality of Life Research, 3*, 437-447.
- Melzack, R. & Katz, J. (2001). The McGill Pain Questionnaire: appraisal and current status. In D. C. Turk & R. Melzack (Eds.), *Handbook of pain assessment* (pp. 35-52). New York: Guilford Press.
- O'Brien, J. & Francis, A. (1988). The use of next-of-kin to estimate pain in cancer patients. *Pain, 35*, 171-178.

- Olden, A. J., Jordan, E. T., Sakima, N. T. & Grass, J. A. (1995). Patients' versus nurses' assessments of pain and sedation after cesarean section. *Journal of Obstetric, Gynecologic, and Neonatal Nursing*, 24, 137-141.
- Padilla, G. V., Presant, C., Grant, M. M., Metter, G., Lipsett, J. & Heide, F. (1983). Quality of life index for patients with cancer. *Research in Nursing & Health*, 6, 117-126.
- Paice, J. A., Mahon, S. M. & Faut-Callahan, M. (1991). Factors associated with adequate pain control in hospitalized postsurgical patients diagnosed with cancer. *Cancer Nursing*, 14, 298-305.
- Rae, C. P., Gallagher, G., Watson, S. & Kinsella, J. (2000). An audit of patient perception compared with medical and nursing staff estimation of pain during burn dressing changes. *European Journal of Anaesthesiology*, 17, 43-45.
- Rafferty, K. A., Smith-Coggins, R. & Chen, A. H. M. (1995). Gender-associated differences in emergency department pain management. *Annals of Emergency Medicine*, 26, 414-421.
- Redinbaugh, E. M., Baum, A., DeMoss, C., Fello, M. & Arnold, R. (2002). Factors associated with the accuracy of family caregiver estimates of patient pain. *Journal of Pain and Symptom Management*, 23, 31-38.
- Salmon, P. & Manyande, A. (1996). Good patients cope with their pain: postoperative analgesia and nurses' perceptions of their patients' pain. *Pain*, 68, 63-68.
- Schuler, M., Neuhauser, T., Hauer, K., Oster, P., Razus, D. & Hacker, M. (2001). Schmerzerkennung bei geriatrischen Patienten durch ein interdisziplinäres Team: Urteilssicherheit und Einflussfaktoren. *Zeitschrift für Gerontologie und Geriatrie*, 34, 376-386.
- Schulte, J., Marvin, J. A. & Sandidge, C. H. (1987). Can nurses assess pain? Fact or fallacy. In American Burn Association (Ed.), *Proceedings of the American Burn Association* (pp. 168). Washington DC: American Burn Association.
- Seers, K. (1987). Perceptions of pain. *Nursing Times*, 83, 37-39.
- Sneeuw, K. C. A., Aaronson, N. K., Osoba, D., Muller, M. J., Hsu, M.-A., Yung, W. K. A., Brada, M. & Newlands, E. S. (1997). The use of significant others as proxy raters of the quality of life of patients with brain cancer. *Medical Care*, 35, 490-506.
- Stein, W. M. & Miech, R. P. (1993). Cancer pain in the elderly hospice patient. *Journal of Pain and Symptom Management*, 8, 474-482.

- Stephenson, N. L. N. (1994). A comparison of nurse and patient. Perceptions of postsurgical pain. *Journal of Intravenous Nursing*, 17, 235-239.
- Stewart, M. L. (1977). Measurement of clinical pain. In A. K. Jacox (Ed.), *Pain: a source book for nurses and other health professionals* (pp. 107-137). Boston: Little Brown & Co.
- Striebel, H. W., Hackenberger, J. & Wessel, A. (1992). Beurteilung der postoperativen Schmerzintensität. Selbst- vs. Fremdbeurteilung. *Der Schmerz*, 6, 199-203.
- Sutherland, J. E., Wesley, R. M., Cole, P. M., Nesvacil, L. J., Daly, M. L. & Gepner, G. J. (1988). Differences and similarities between patient and physician perceptions of patient pain. *Family Medicine*, 20, 343-346.
- Teske, K., Daut, R. L. & Cleeland, C. S. (1983). Relationships between nurses' observations and patients' self-reports of pain. *Pain*, 16, 289-296.
- Thomas, S. H., Borczuk, P., Shackelford, J., Ostrander, J., Silver, D., Evans, M. & Stein, J. (1999). Patient and physician agreement on abdominal pain severity and need for opioid analgesia. *American Journal of Emergency Medicine*, 17, 586-590.
- Thomas, T., Robinson, C., Champion, D., McKell, M. & Pell, M. (1998). Prediction and assessment of the severity of post-operative pain and of satisfaction with management. *Pain*, 75, 177-185.
- Todd, K. H., Lee, T. & Hoffman, J. R. (1994). The effect of ethnicity on physician estimates of pain severity in patients with isolated extremity trauma. *Journal of the American Medical Association*, 271, 925-928.
- van der Does, A. J. W. (1989). Patients' and nurses' ratings of pain and anxiety during burn wound care. *Pain*, 39, 95-101.
- van Dulmen, A. M., Fennis, J. F. M., Mokkink, H. G. A., van der Velden, H. G. M. & Bleijenberg, G. (1994). Doctors' perception of patients' cognitions and complaints in irritable bowel syndrome at an out-patient clinic. *Journal of Psychosomatic Research*, 38, 581-590.
- Walkenstein, M. D. (1982). Comparison of burned patients' perception of pain with nurses' perception of patients' pain. *Journal of Burn Care & Rehabilitation*, 3, 233-236.
- Weiner, D., Peterson, B. & Keefe, F. (1999). Chronic pain-associated behaviors in the nursing home: resident versus caregiver perceptions. *Pain*, 80, 577-588.

- Weiner, D. K., Ladd, K. E., Pieper, C. F. & Keefe, F. J. (1995). Pain in the nursing home: resident versus staff perceptions. *Journal of the American Geriatrics Society*, 43, SA2.
- Werner, P., Cohen-Mansfield, J., Watson, V. & Pasis, S. (1998). Pain in participants of adult day care centers: assessment by different raters. *Journal of Pain and Symptom Management*, 15, 8-17.
- Yeager, K. A., Miaskowski, C., Dibble, S. L. & Wallhagen, M. (1995). Differences in pain knowledge and perception of the pain experience between outpatients with cancer and their family caregivers. *Oncology Nursing Forum*, 22, 1235-1241.

APPENDIX II: Study 1

II.1 Correspondence with Ethics Committee: initial letter and reply

Guy's and St Thomas' Hospital 

NHS Trust

INPUT Pain Management Unit
St Thomas' Hospital, London, SE1 7EH
Tel 0207 922 8107, Fax 0207 922 8229

St Thomas' Hospital
Lambeth Palace Road
London SE1 7EH

Tel: 020 7928 9292

Dr G du Mont
Chair, Research Ethics Committee
St Thomas' Hospital

6th October 2001

Dear Dr du Mont

When I contacted you before (5th May 1999) concerning an MSc project which sampled health professionals but not patients, you replied (7th May 1999) that there were no ethical issues and thus no need for Ethics Committee approval. I now propose to undertake three projects for my PhD, for which my supervisor at GKT & St Thomas Hospital is Dr Amanda Williams, all of which sample the influences on judgements about pain by health care professionals and by the spouses and family members of pain patients. None of the studies involves any patient contact, since the judgements will be made using vignettes and videotapes of people in pain. Both groups of participants will be asked to volunteer for the studies.

I would be most grateful to know if the studies we are planning require me to complete an Ethics Form, so that I could submit it as soon as possible.

Yours sincerely



Judith Kappesser

PhD Student

MSc Health Psychology



Amanda Williams

Sen. Lecturer in Clinical Health Psychology

& Consultant Clinical Psychologist,

INPUT Pain Management Unit

ST THOMAS' HOSPITAL RESEARCH ETHICS COMMITTEE

*Ethics Committee Office
Block 5, South Wing*

Chairman - Dr G du Mont
Administrator - Ms S Hirsch

St Thomas' Hospital
Lambeth Palace Road
London SE1 7EH

Tel: 020 7928 9292

24 October 2001

Ext 2097 fax 020 7922 8163
stella.hirsch@gstt.sthames.nhs.uk

Judith Kappesser
PhD Student
MSc Health Psychology
INPUT Pain Management Unit
St Thomas' Hospital

Dear Ms Kappesser

Re: Influences on judgements about pain by health care professionals and by the spouses and family members of pain patients

Dr du Mont has seen your letter dated 6 October 2001 in which you ask if it is necessary to complete an ethics application form for the above proposed project.

If you plan to involve family members, you will need to make a formal application to the Research Ethics Committee. Please let me know if you would like the form/guidelines electronically, in which case I need your email address, otherwise I can post a copy.

The next closing date for applications is 5th November.

I look forward to hearing from you.

Yours sincerely



Stella Hirsch
Administrator
Research Ethics Committee

cc: Dr Amanda Williams

II.2 Poster asking for volunteers

VOLUNTEERS WANTED FOR RESEARCH

Who?	Nurses and doctors working for A & E
What is it?	Research project on pain judgements
When and where?	Before or after your shift, I will be in your staff room from the beginning of October onwards
How long?	10 - 15 minutes
Contact	Judith Kappesser Research Psychologist INPUT Pain Management Unit, STH ext 1426 or judith.kappesser@t-online.de

II.3 Questionnaire

No _____

Please state your

1. Age in years _____

2. Sex male female

3. Are you a therapist doctor nurse physio- other _____

(a) If you are a doctor, are you a house officer senior house officer specialist registrar consultant

(b) If you are a nurse, which grade are you? health care assistant D E F G H I

(c) If you are a physiotherapist, which grade are you? Junior Senior II Senior I Super-intendent III Super-intendent II Super-intendent I

4. How long have you been working in your profession? _____ years

5. Do you work in A&E Oncology Orthopedics Palliative Care
 Physiotherapy Rheumatology Other _____

6. How long have you been working in your specialty? _____ years

7. Please order the following factors according to the **importance they have for you when judging pain**. Assign 1 to the most important, 2 to the second most important and so on.

- results of medical investigations _____
- protective movements and postures _____
- patient mentions disability claim or sickness certification _____
- patient has history of psychiatric illness _____
- physiological indicators (e.g. being pale, sweaty) _____
- patient gets much attention and care from his/her family _____
- facial expression of pain _____
- patient has history of drug abuse _____
- patient reports pain _____
- patient moans, groans or sighs _____

Is there anything missing which is important for you when judging pain?

8. Please order the following factors according to **how easy they are for the patient to control or manipulate**. Assign 1 to the one easiest to control, 2 to the second easiest and so on.

- protective movements and postures _____
- report of pain _____
- facial expression _____
- sounds (e.g. moaning, sighing) _____
- results of medical investigations _____
- physiological indicators (e.g. being pale, sweaty) _____

9. Please order the following factors according to **how cautious they make you when judging pain**. Assign 1 to the one making you most cautious, 2 to the one making you second most cautious and so on.

- no sounds (e.g. moaning, sighing) _____
- no protective movements and postures _____
- no facial expression of pain _____
- no report of pain _____
- no physiological indicators (e.g. being pale, sweaty) _____
- attention and care from his/her family _____
- all relevant medical investigations are normal _____
- disability claim or sickness certification _____
- history of drug abuse _____
- psychiatric illness _____

Is there anything missing which may make you cautious when judging pain?

10. Please complete each of the sentences.

(a) If a patient has genuine pain, it is my job to

(b) If I find out that a patient is faking pain, it is my job to

11. At best guess, what proportion of patients do you think

- fake pain? _____ %
- exaggerate pain? _____ %
- minimise pain? _____ %
- hide pain? _____ %

Thank you for participating!

II.4 Feedback sheet

Which cues are used to judge pain? Perceptions of health care professionals.

Judith Kappesser & Amanda C de C Williams
INPUT Pain Management Unit, St Thomas' Hospital, London SE1 7EH

The aim of this study was to investigate which cues health care professionals who often judge pain (1) regard as important when judging pain, (2) consider to be easily manipulated by patients and (3) regard as making them cautious when judging pain.

Method

Nurses and doctors working in A & E, oncology / palliative care, orthopaedics and physiotherapists (n = 65) volunteered to participate. They were given lists with ten possible cues in random order. Cues were five pain behaviours (verbal report, para-linguistic vocalisations, facial expression, movements, physiologic reactions) and five context cues (results of medical investigations, psychiatric history, attention from family, disability claim, drug abuse history). The task was to (1) rank order the ten cues according to the importance they have when judging pain, (2) rank order the five pain behaviours plus medical investigation results according to how easy they are to manipulate and (3) rank order all ten cues according to how cautious they make observers when judging pain, the pain behaviours and medical investigation results by their absence, the other context cues by their presence.

Results

1. Importance

- Rank ordering (most important cue first): verbal report, physiological indicators, facial expression, posture/movement, sounds, results of medical investigations, attention from family, drug abuse history, psychiatric history and disability claim.
- High level of concordance, i.e. all volunteers are applying essentially the same standard in ranking the ten cues

2. Manipulation

- Rank ordering (cue most easy to manipulate first): verbal report, sounds, facial expression, posture/movement, physiological indicators and results of medical investigations
- High level of concordance

3. Cautiousness

- Rank ordering (most cautious making cue first): no verbal report, no physiological indicators, no posture/movement, no facial expression, drug abuse history, no sounds, no results of medical investigations, disability claim, psychiatric history and attention from family
- Low level of concordance (but still sign. different from chance); interestingly concordance is higher within each specialty than concordance for all specialties together

Thank you for your help and support with this research!

If you wish to get more information, please email to judith.kappesser@t-online.de.

II.5 Post-hoc analysis of differences between cues regarding their importance: results of Wilcoxon signed rank test

pair of cues being compared	z-value ¹	p-value ²
verbal report - facial expression	- 2.40 ³	0.016
verbal report - sounds	- 4.64 ³	< 0.001
verbal report - postures	- 2.81 ³	0.005
verbal report - phys. indicators	- 1.68 ³	0.093
verbal report - results	- 6.37 ³	< 0.001
verbal report - disability claim	- 6.58 ³	< 0.001
verbal report - drug abuse	- 6.69 ³	< 0.001
verbal report - attention	- 6.56 ³	< 0.001
verbal report - psychiatric history	- 6.62 ³	< 0.001
facial expression - sounds	- 3.37 ³	0.001
facial expression - postures	- 0.53 ³	0.599
facial expression - phys. indicators	- 0.56 ⁴	0.574
facial expression - results	- 5.38 ³	< 0.001
facial expression - disability claim	- 6.58 ³	< 0.001
facial expression - drug abuse	- 6.79 ³	< 0.001
facial expression - attention	- 6.83 ³	< 0.001
facial expression - psychiatric history	- 6.71 ³	< 0.001
sounds - postures	- 3.00 ⁴	0.003
sounds - phys. indicators	- 2.99 ⁴	0.003
sounds - results	- 3.73 ³	< 0.001
sounds - disability claim	- 6.25 ³	< 0.001
sounds - drug abuse	- 6.70 ³	< 0.001
sounds - attention	- 6.43 ³	< 0.001
sounds - psychiatric history	- 6.34 ³	< 0.001
postures - phys. indicators	- 1.32 ⁴	0.186
postures - results	- 5.63 ³	< 0.001
postures - disability claim	- 6.86 ³	< 0.001
postures - drug abuse	- 6.87 ³	< 0.001
postures - attention	- 6.92 ³	< 0.001
postures - psychiatric history	- 6.89 ³	< 0.001
phys. indicators - results	- 6.11 ³	< 0.001
phys. indicators - disability claim	- 6.80 ³	< 0.001
phys. indicators - drug abuse	- 6.86 ³	< 0.001
phys. indicators - attention	- 6.83 ³	< 0.001
phys. indicators - psychiatric history	- 6.92 ³	< 0.001
results - disability claim	- 5.32 ³	< 0.001
results - drug abuse	- 4.75 ³	< 0.001
results - attention	- 3.98 ³	< 0.001
results - psychiatric history	- 5.53 ³	< 0.001
disability claim - drug abuse	- 1.14 ⁴	0.254
disability claim - attention	- 1.74 ⁴	0.082
disability claim - psychiatric history	- 1.23 ⁴	0.220
drug abuse - attention	- 1.12 ⁴	0.264
drug abuse - psychiatric history	- 0.01 ⁴	0.994
attention - psychiatric history	- 1.11 ³	0.268

¹ based on sum of positive and negative ranks

² p-value significant if $p < 0.05/45 = 0.001$

³ based on negative ranks

⁴ based on positive ranks

II.6: Summary of missing importance cues as provided by participants

- Is pain chronic? Chronic pain patient may not lock in pain, but has learnt to live with it
- compensation claim, length of time experiencing pain
- compensation, employment
- culture
- current medication
- dependence on analgesics, impact on daily life
- effect of pain on daily activities
- effect of pain on lifestyle
- effect of pain on quality of life
- emotional expression of pain
- ethnic background
- functional impairment caused by pain, objective signs of muscular atrophy
- I feel different factors will differ for each person due to history and personality
- length of time with pain
- length of time with pain, mental state, e.g. had bad news
- medical condition e.g. rheumatoid arthritis, abscess
- obvious bruising/swelling
- other objective signs (loss of movement range, history of pain, effect of pain on function)
- pain is more than physical, makes judging 1-10 difficult
- patient is very quiet/does not complain
- physiological finding on examination
- previous exposure to analgesia, e.g. what medication they were admitted on
- tension, flinching, guarding on physical exam

II.7 Post-hoc analysis of differences between cues regarding their ease of manipulation: results of Wilcoxon signed rank test

pair of cues being compared	z-value ¹	p-value ²
verbal report - facial expression	- 5.15 ³	< <u>0.001</u>
verbal report - sounds	- 3.55 ³	< <u>0.001</u>
verbal report - postures	- 5.76 ³	< <u>0.001</u>
verbal report - phys. indicators	- 7.03 ³	< <u>0.001</u>
facial expression - sounds	- 3.48 ³	< <u>0.001</u>
facial expression - postures	- 3.25 ³	<u>0.001</u>
facial expression - phys. indicators	- 6.83 ³	< <u>0.001</u>
sounds - postures	- 5.29 ³	< <u>0.001</u>
sounds – phys. indicators	- 6.89 ³	< <u>0.001</u>
postures - phys. indicators	- 6.12 ³	< <u>0.001</u>

¹ based on sum of positive and negative ranks

² p-value significant if $p < 0.05/45 = 0.001$

³ based on negative ranks

⁴ based on positive ranks

II.8 Post-hoc analysis of differences between cues regarding their cautiousness: results of Wilcoxon signed rank test

pair of cues being compared	z-value ¹	p-value ²
verbal report - facial expression	- 2.43 ³	0.015
verbal report - sounds	- 3.97 ³	< 0.001
verbal report - postures	- 1.92 ³	0.054
verbal report - phys. indicators	- 1.83 ³	0.068
verbal report - results	- 3.59 ³	< 0.001
verbal report - disability claim	- 2.86 ³	0.004
verbal report - drug abuse	- 3.03 ³	0.002
verbal report - attention	- 3.52 ³	< 0.001
verbal report - psychiatric history	- 3.57 ³	< 0.001
facial expression - sounds	- 2.84 ³	0.004
facial expression - postures	- 0.54 ⁴	0.592
facial expression - phys. indicators	- 0.69 ⁴	0.493
facial expression - results	- 1.65 ³	0.099
facial expression - disability claim	- 1.66 ³	0.097
facial expression - drug abuse	- 1.42 ³	0.155
facial expression - attention	- 2.68 ³	0.007
facial expression - psychiatric history	- 2.36 ³	0.018
sounds - postures	- 2.76 ⁴	0.006
sounds - phys. indicators	- 2.71 ⁴	0.007
sounds - results	- 0.02 ³	0.984
sounds - disability claim	- 0.13 ³	0.893
sounds - drug abuse	- 0.49 ⁴	0.623
sounds - attention	- 1.16 ³	0.248
sounds - psychiatric history	- 0.60 ³	0.552
postures - phys. indicators	- 0.29 ⁴	0.769
postures - results	- 2.74 ³	0.006
postures - disability claim	- 2.06 ³	0.040
postures - drug abuse	- 1.71 ³	0.087
postures - attention	- 2.88 ³	0.004
postures - psychiatric history	- 2.54 ³	0.011
phys. indicators - results	- 2.83 ³	0.005
phys. indicators - disability claim	- 2.19 ³	0.028
phys. indicators - drug abuse	- 1.77 ³	0.077
phys. indicators - attention	- 3.08 ³	0.002
phys. indicators - psychiatric history	- 2.62 ³	0.009
results - disability claim	- 0.40 ³	0.691
results - drug abuse	- 0.49 ⁴	0.626
results - attention	- 0.71 ³	0.479
results - psychiatric history	- 0.39 ³	0.700
disability claim - drug abuse	- 0.68 ⁴	0.496
disability claim - attention	- 0.63 ³	0.527
disability claim - psychiatric history	- 0.07 ³	0.946
drug abuse - attention	- 1.36 ³	0.174
drug abuse - psychiatric history	- 1.09 ³	0.276
attention - psychiatric history	- 0.68 ⁴	0.499

¹ based on sum of positive and negative ranks

² p-value significant if $p < 0.05/45 = 0.001$

³ based on negative ranks

⁴ based on positive ranks

II.9: Summary of missing cautiousness-inducing cues as provided by participants

- age, previous experience
- attention given to other patients
- compensation claim
- compensation, behaviour, fear avoidance
- culture
- culture, religious, social and emotional factors
- dependence on analgesics, impact on daily life
- depends very much on individual
- discrepancy between physical signs and no report
- failure to attempt self-medication before coming to A&E
- if family insists pt has pain but pt denies
- inability to communicate due to disability
- insurance claims
- known to attend hospitals frequently
- legal claims
- litigation issues, patient's attitude (very focussed on pain)
- need to respect the patients' wishes
- no functional disability, ethnicity
- no response at any stage to any analgesic measures
- patients' attitudes and beliefs re: their diagnosis
- some patients like having certain pain degree, pain as indicator of disease progression
- taken own analgesics, made effort themselves
- total pain, e.g. emotional disturbance; current analgesic regime; pain history

APPENDIX III: Study 2

III.1: Letter and poster by which volunteers were contacted

1.1 Example of letters to contact oncology doctors

27th of July 2003

Dear

I am a PhD student based at INPUT Pain Management Unit at St Thomas'; my PhD is concerned with pain judgements of the sort that clinicians have to make every day. I was given your name by consultant Mr Adrian Timothy who also gave his permission to contact you. I would be grateful if you could consider taking part in one of the studies for my PhD.

Taking part would mean 10 minutes of your time. Your task would be to watch eight videotaped faces of patients and to rate your impressions regarding their pain (there are no 'right' or 'wrong' answers). The data will be collected anonymously.

I will be in the chemotherapy quiet room in St Thomas' the next two Thursdays (31st of July and 7th of August) from 3 pm onwards and grateful if you could make some time then.

If this time does not suit you or if you have any questions, you can most easily contact me by email (judith.kappesser@t-online.de), or at INPUT Pain Management Unit on St Thomas' ext 1426 (my supervisor's, Dr Amanda Williams, extension).

Thank you for considering this request!

Yours sincerely

Judith Kappesser
INPUT Pain Management Unit
St Thomas' Hospital

Supervisor: Dr Amanda C de C Williams,
Senior Lecturer in Clinical Health Psychology, GKT, and
Consultant Clinical Psychologist, INPUT Pain Management Unit

1.2 Example of poster to contact A&E doctors and nurses

VOLUNTEERS WANTED FOR RESEARCH

Who?	Nurses and doctors working for A & E
What is it?	Research project on pain judgements Watch brief video sequences and judge pain
When and where?	Before or after your shift, I will be in your staff room from the 22 nd July onwards
How long?	10 - 15 minutes
Contact	Judith Kappesser Research Psychologist INPUT Pain Management Unit, STH ext 1426 or judith.kappesser@t-online.de

III.2 Questionnaires

2.1 Questionnaire for condition 1: facial expression only

No _____

Please state your

1. Age in years _____
 2. Sex male female
 3. Are you a doctor nurse other _____
 - (a) If you are a doctor, are you a house officer senior house officer specialist registrar consultant
 - (b) If you are a nurse, which grade are you?
 health care assistant D E F G H I
 4. How long have you been working in your profession? _____ years
 5. Do you work in A&E Oncology Palliative Care
 6. How long have you been working in your specialty? _____ years
-

Introduction

In the following you will see eight videotaped faces. I would like you to imagine that all of them are your patients. From this information please could you rate whether they are in pain and, if so, how intense the pain is.

To rate the intensity of pain you are asked to choose the one word which best describes the maximum intensity shown in the video.

Also you will be asked which word **you think each patient used** to describe his/her state.

And lastly you will be asked how confident you felt with each of these judgements you made.

Do you have any questions before we start?

1. When watching the faces, did you have the impression one or several of the patients was/were

faking pain? yes no

exaggerating pain yes no

minimising pain? yes no

hiding pain? yes no

2. At best guess, what proportion of patients in general (not just those in the videos) do you think (percentages do not have to add up to 100%)

· fake pain? _____ %

· exaggerate pain? _____ %

· minimise pain? _____ %

· hide pain? _____ %

Thank you for participating!

2.2 Questionnaire for condition 2: facial expression and verbal report

No _____

Please state your

1. Age in years _____
 2. Sex male female
 3. Are you a doctor nurse other _____
 - (a) If you are a doctor, are you a house officer senior house officer specialist registrar consultant
 - (b) If you are a nurse, which grade are you? health care assistant D E F G H I
 4. How long have you been working in your profession? _____ years
 5. Do you work in A&E Oncology Palliative Care
 6. How long have you been working in your specialty? _____ years
-

Introduction

In the following you will see eight videotaped faces. I would like you to imagine that all of them are your patients. You will also be told what each of them said about how they felt. From the information please could you rate whether they are in pain and, if so, how intense the pain is.

To rate the intensity of pain you are asked to choose the one word which best describes the maximum intensity shown in the video. As this is **your** opinion based on all available information, your rating may or may not agree with their report.

And lastly you will be asked how confident you were of your judgements.

Do you have any questions before we start?

Face _____

1. Please choose from the following words the one that best describes the maximum pain experienced by the patient

- | | |
|--|--|
| <input type="checkbox"/> No pain | <input type="checkbox"/> Moderate |
| <input type="checkbox"/> Extremely weak | <input type="checkbox"/> Barely strong |
| <input type="checkbox"/> Faint | <input type="checkbox"/> Clear-cut |
| <input type="checkbox"/> Very weak | <input type="checkbox"/> Slightly intense |
| <input type="checkbox"/> Weak | <input type="checkbox"/> Strong |
| <input type="checkbox"/> Very mild | <input type="checkbox"/> Intense |
| <input type="checkbox"/> Mild | <input type="checkbox"/> Very intense |
| <input type="checkbox"/> Slightly moderate | <input type="checkbox"/> Extremely intense |

2. How confident are you with your judgement?

- not at all extremely

Face _____

1. Please choose from the following words the one that best describes the maximum pain experienced by the patient

- | | |
|--|--|
| <input type="checkbox"/> No pain | <input type="checkbox"/> Moderate |
| <input type="checkbox"/> Extremely weak | <input type="checkbox"/> Barely strong |
| <input type="checkbox"/> Faint | <input type="checkbox"/> Clear-cut |
| <input type="checkbox"/> Very weak | <input type="checkbox"/> Slightly intense |
| <input type="checkbox"/> Weak | <input type="checkbox"/> Strong |
| <input type="checkbox"/> Very mild | <input type="checkbox"/> Intense |
| <input type="checkbox"/> Mild | <input type="checkbox"/> Very intense |
| <input type="checkbox"/> Slightly moderate | <input type="checkbox"/> Extremely intense |

2. How confident are you with your judgement?

- not at all extremely

1. When watching the faces, did you have the impression one or several of the patients was/were

faking pain? yes no

exaggerating pain yes no

minimising pain? yes no

hiding pain? yes no

2. At best guess, what proportion of patients in general (not just those in the videos) do you think
(percentages do not have to add up to 100%)

· fake pain? _____ %

· exaggerate pain? _____ %

· minimise pain? _____ %

· hide pain? _____ %

Thank you for participating!

2.3 Questionnaire for condition 3: facial expression, verbal report and opioids

No _____

Please state your

1. Age in years _____
 2. Sex male female
 3. Are you a doctor nurse other _____
 - (a) If you are a doctor, are you a house officer senior house officer specialist registrar consultant
 - (b) If you are a nurse, which grade are you? health care assistant D E F G H I
 4. How long have you been working in your profession? _____ years
 5. Do you work in A&E Oncology Palliative Care
 6. How long have you been working in your specialty? _____ years
-

Introduction

In the following you will see eight videotaped faces. I would like you to imagine that all of them are your patients. You will also be told what each of them said about how they felt. From the information please could you rate whether they are in pain and, if so, how intense the pain is.

To rate the intensity of pain you are asked to choose the one word which best describes the maximum intensity shown in the video. As this is **your** opinion based on all available information, your rating may or may not agree with their report.

Be aware that when videotaped some people were faking pain to obtain opioid drugs.

And lastly you will be asked how confident you were of your judgements.

Do you have any questions before we start?

Face _____

1. Please choose from the following words the one that best describes the maximum pain experienced by the patient

- | | |
|--|--|
| <input type="checkbox"/> No pain | <input type="checkbox"/> Moderate |
| <input type="checkbox"/> Extremely weak | <input type="checkbox"/> Barely strong |
| <input type="checkbox"/> Faint | <input type="checkbox"/> Clear-cut |
| <input type="checkbox"/> Very weak | <input type="checkbox"/> Slightly intense |
| <input type="checkbox"/> Weak | <input type="checkbox"/> Strong |
| <input type="checkbox"/> Very mild | <input type="checkbox"/> Intense |
| <input type="checkbox"/> Mild | <input type="checkbox"/> Very intense |
| <input type="checkbox"/> Slightly moderate | <input type="checkbox"/> Extremely intense |

2. How confident are you with your judgement?

-
- not at all extremely

Face _____

1. Please choose from the following words the one that best describes the maximum pain experienced by the patient

- | | |
|--|--|
| <input type="checkbox"/> No pain | <input type="checkbox"/> Moderate |
| <input type="checkbox"/> Extremely weak | <input type="checkbox"/> Barely strong |
| <input type="checkbox"/> Faint | <input type="checkbox"/> Clear-cut |
| <input type="checkbox"/> Very weak | <input type="checkbox"/> Slightly intense |
| <input type="checkbox"/> Weak | <input type="checkbox"/> Strong |
| <input type="checkbox"/> Very mild | <input type="checkbox"/> Intense |
| <input type="checkbox"/> Mild | <input type="checkbox"/> Very intense |
| <input type="checkbox"/> Slightly moderate | <input type="checkbox"/> Extremely intense |

2. How confident are you with your judgement?

-
- not at all extremely

1. When watching the faces, did you have the impression one or several of the patients was/were

faking pain? yes no

exaggerating pain yes no

minimising pain? yes no

hiding pain? yes no

2. At best guess, what proportion of patients in general (not just those in the videos) do you think
(percentages do not have to add up to 100%)

· fake pain? _____ %

· exaggerate pain? _____ %

· minimise pain? _____ %

· hide pain? _____ %

Thank you for participating!

III.3 Feedback sheet

What causes underestimation of pain? Feedback on the results of the study you participated in this summer

Judith Kappesser & Amanda C de C Williams
INPUT Pain Management Unit, St Thomas' Hospital, London SE1 7EH

Introduction

Several studies have found that an individual's pain intensity can be underestimated by observers such as health care staff. Although cues used by observers to judge pain have been investigated, no clear explanation has emerged. This study investigated how pain judgements vary with cues using two hypotheses:

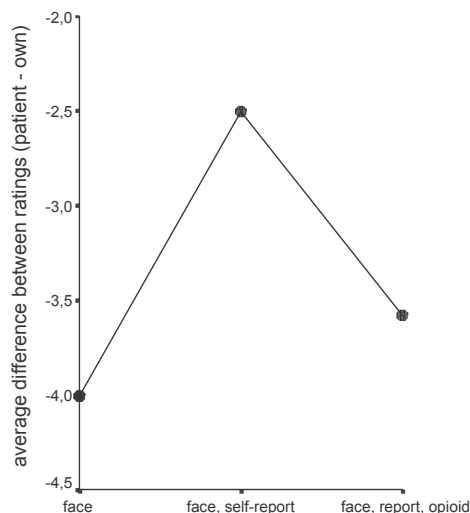
1. underestimation occurs when verbal report is absent;
2. underestimation occurs when a context cue suggests that the patient has an agenda other than pain relief.

Method

Doctors and nurses working for A & E (n = 60) or Oncology, Haematology and Palliative Care (n= 60) watched silent video sequences of patients' faces in pain and estimated the pain intensity on the same scale as the patients had done when video-taped. Group 1 (n = 40) had to base their judgements only on the facial expressions in the video. Group 2 (n = 40) was given two cues: patients' own pain ratings and facial expression. Group 3 (n = 40) judged on face and self-report and was informed that some patients were faking pain to obtain opioids.

Results

Only Group 2 (face and self-report) rated pain close to patients' own ratings. Group 1 (face only) and Group 3 (face, self-report, opioids) significantly underestimated pain.



Discussion

Results of the study show that underestimation can be exacerbated by absence of an important cue (self-report) as well as by a context factor (covert agenda).

**Thank you very much for your great help and support
with this study!**

And if you wish to get further information,
please do not hesitate to contact Judith by email: judith.kappesser@t-online.de

III.4 Correlations between all variables enclosed in the chain graph

variables ¹		difference	faking	exagg.	minim.	hiding	age	sex	profession	specialty	years in prof.
faking	<i>r</i>	-0,06									
	<i>p</i>	0,51									
exaggerating	<i>r</i>	-0,27**	0,49**								
	<i>p</i>	0,00	0,00								
minimising	<i>r</i>	0,15	-0,02	-0,13							
	<i>p</i>	0,10	0,85	0,16							
hiding	<i>r</i>	0,10	0,17	-0,02	0,76**						
	<i>p</i>	0,28	0,06	0,80	0,00						
age	<i>r</i>	0,00	-0,21*	-0,26**	0,01	-0,10					
	<i>p</i>	0,97	0,02	0,00	0,88	0,28					
sex	<i>r</i>	0,13	-0,04	-0,07	0,17	0,13	0,06				
	<i>p</i>	0,16	0,66	0,43	0,07	0,15	0,48				
profession	<i>r</i>	0,14	0,14	0,04	0,28**	0,20*	-0,01	0,45**			
	<i>p</i>	0,14	0,12	0,67	0,00	0,03	0,95	0,00			
specialty	<i>r</i>	0,16	-0,18*	-0,36**	0,31**	0,26**	0,20*	0,20*	0,00		
	<i>p</i>	0,08	0,04	0,00	0,00	0,00	0,03	0,03	1,00		
years in profession	<i>r</i>	0,06	-0,24**	-0,29**	0,01	-0,14	0,68**	0,09	0,03	0,17	
	<i>p</i>	0,51	0,01	0,00	0,93	0,14	0,00	0,31	0,73	0,06	
years in specialty	<i>r</i>	0,07	-0,23*	-0,24**	0,10	-0,12	0,55**	0,19*	0,05	0,21*	0,78**
	<i>p</i>	0,45	0,01	0,01	0,27	0,19	0,00	0,04	0,60	0,02	0,00

¹all $n = 120$; * $p \leq 0,05$; ** $p \leq 0,01$

III.5 Regression equations for variables excluded from the chain graph

Response variable: difference in pain intensity ratings

explanatory variables	<i>t</i>	<i>p</i>
sex	0.847	0.399
hiding	1.636	0.104
years in specialty	- 0.010	0.992
specialty	0.589	0.557
years in profession	- 0.162	0.871
faking	0.885	0.378
age	- 0.736	0.464
profession	1.833	0.069
minimising	2.037	0.044

Response variable: faking

explanatory variables	<i>t</i>	<i>p</i>
group (dummy 1)	0.586	0.559
group (dummy 2)	1.121	0.265
age	- 0.664	0.508
years in specialty	- 0.675	0.501
sex	- 0.193	0.847
specialty	- 1.633	0.105
profession	1.696	0.093

Response variable: exaggerating

explanatory variables	<i>t</i>	<i>p</i>
years in specialty	0.117	0.907
sex	0.168	0.867
group (dummy 1)	0.583	0.561
group (dummy 2)	1.606	0.111
profession	0.552	0.582
age	- 0.674	0.502

Response variable: minimising

explanatory variables	<i>t</i>	<i>p</i>
sex	- 0.272	0.786
age	- 0.539	0.591
years in specialty	0.305	0.761
years in profession	- 0.642	0.522
group (dummy 1)	1.138	0.257
group (dummy 2)	- 0.600	0.549

Response variable: hiding

explanatory variables	<i>t</i>	<i>p</i>
age	- 1.726	0.087
sex	0.924	0.358
years in specialty	- 2.026	0.045
group (dummy 1)	2.044	0.043
group (dummy 2)	0.410	0.682
years in profession	- 2.088	0.039
profession	2.319	0.022

III.6 Correlations between confidence and all explanatory variables

variables ¹	confidence	
difference	<i>r</i>	-0,03
	<i>p</i>	0,76
faking	<i>r</i>	0,01
	<i>p</i>	0,89
exaggerating	<i>r</i>	0,12
	<i>p</i>	0,19
minimising	<i>r</i>	-0,22 *
	<i>p</i>	0,02
hiding	<i>r</i>	-0,14
	<i>p</i>	0,12
age	<i>r</i>	0,00
	<i>p</i>	0,97
sex	<i>r</i>	-0,04
	<i>p</i>	0,65
profession	<i>r</i>	0,04
	<i>p</i>	0,69
specialty	<i>r</i>	-0,26 **
	<i>p</i>	0,00
years in profession	<i>r</i>	0,05
	<i>p</i>	0,55
years in specialty	<i>r</i>	0,00
	<i>p</i>	0,96

¹ all $n = 120$; * $p \leq 0.05$; ** $p \leq 0.01$

III.7 Regression equations for variables found not to be associated with confidence

explanatory variables	<i>t</i>	<i>p</i>
exaggerating	- 0.201	0.841
sex	0.197	0.844
years in specialty	1.213	0.228
age	0.783	0.435
faking	- 0.411	0.682
hiding	- 1.157	0.250
profession	0.277	0.783
group (dummy 1)	1.144	0.255
group (dummy 2)	1.596	0.113
years in profession	1.690	0.094
minimising	- 1.728	0.087

APPENDIX IV: Study 3

IV.1 Ethics Committee application form including consent form

2000 version 4
<p><u>ST THOMAS' HOSPITAL LOCAL RESEARCH ETHICS COM- MITTEE</u></p> <p><u>APPLICATION FORM</u></p>

1)	<u>TITLE OF PROJECT</u>																		
	Patients' family caregivers' judgements of pain.																		
2a)	<u>PROPOSED STARTING DATE</u> May 2002																		
2b)	<u>DURATION OF STUDY</u> Approximately five months																		
3)	<u>INVESTIGATORS</u>																		
	<table style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 10%;"></th> <th style="width: 30%;"><u>Name</u></th> <th style="width: 15%;"><u>Status</u></th> <th style="width: 20%;"><u>Dept and address</u></th> <th style="width: 15%;"><u>Contact telephone no or bleep</u></th> <th style="width: 10%;"><u>Membership of MDU/MPS or other body</u></th> </tr> </thead> <tbody> <tr> <td style="padding-left: 20px;">a) Principal:</td> <td>Judith Kappesser, MSc</td> <td>Health Psychology Research Assistant</td> <td></td> <td style="text-align: center;">x 1426</td> <td></td> </tr> <tr> <td style="padding-left: 20px;">b) Others:</td> <td>Amanda Williams, PhD</td> <td>Senior Lecturer in Clinical Health Psychology & Consultant Clinical Psychologist INPUT Pain Management Unit St Thomas' Hospital</td> <td></td> <td style="text-align: center;">x 1426</td> <td style="text-align: center;">BPS</td> </tr> </tbody> </table>		<u>Name</u>	<u>Status</u>	<u>Dept and address</u>	<u>Contact telephone no or bleep</u>	<u>Membership of MDU/MPS or other body</u>	a) Principal:	Judith Kappesser, MSc	Health Psychology Research Assistant		x 1426		b) Others:	Amanda Williams, PhD	Senior Lecturer in Clinical Health Psychology & Consultant Clinical Psychologist INPUT Pain Management Unit St Thomas' Hospital		x 1426	BPS
	<u>Name</u>	<u>Status</u>	<u>Dept and address</u>	<u>Contact telephone no or bleep</u>	<u>Membership of MDU/MPS or other body</u>														
a) Principal:	Judith Kappesser, MSc	Health Psychology Research Assistant		x 1426															
b) Others:	Amanda Williams, PhD	Senior Lecturer in Clinical Health Psychology & Consultant Clinical Psychologist INPUT Pain Management Unit St Thomas' Hospital		x 1426	BPS														
4)	<u>WILL ANY OTHER STAFF BE INVOLVED IN THE RESEARCH PROJECT</u>																		
	<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 60%;">Nursing Staff (give particulars of involvement)</td> <td style="width: 40%; text-align: center;">No</td> </tr> <tr> <td>Any other staff</td> <td style="text-align: center;">No</td> </tr> <tr> <td>Social Workers</td> <td style="text-align: center;">No</td> </tr> <tr> <td>Laboratory staff</td> <td style="text-align: center;">No</td> </tr> <tr> <td>Any other staff</td> <td style="text-align: center;">No</td> </tr> </table>	Nursing Staff (give particulars of involvement)	No	Any other staff	No	Social Workers	No	Laboratory staff	No	Any other staff	No								
Nursing Staff (give particulars of involvement)	No																		
Any other staff	No																		
Social Workers	No																		
Laboratory staff	No																		
Any other staff	No																		

5) PLACES WHERE THE RESEARCH WILL BE DONE

Are the patients to be admitted?	N/A
If so is their admission part of a routine clinical admission?	N/A
If so will this research involve an extended stay in hospital?	N/A
If so, how long will this extension of their stay be?	N/A

6) BACKGROUND OF THE STUDIES

Pain is defined as a private, highly personal and subjective experience for the sufferers (IASP, 1979). Chronic pain patients consistently report not feeling understood, and in some settings, not even believed when they describe their pain and its effects on them. Pain sufferers communicate their suffering to others verbally and by their behaviour (Craig, Prkachin & Grunau, 1992). Most studies of how others judge pain involve medical and nursing staff, and close relatives or family members of patients. The latter show very variable results, which we believe may be explained by applying principles derived from evolutionary psychology (Williams, 2002).

Craig, KD, Prkachin, KM & Grunau, RVE (1992). The facial expression of pain. In DC Turk & R Melzack (Eds.), Handbook of pain assessment (pp. 257-276). New York: Guildford Press.

International Association for the Study of Pain (IASP; 1979). Pain terms: a list with definitions and notes on usage. Pain, 6, 249- 252.

Williams, ACdeC (2002). Facial expression of pain: an evolutionary account. Behavioural & Brain Sciences, in press.

7) AIMS OF THE STUDIES

(Please include anticipated clinical use of outcomes, the potential benefit to the patient and the potential benefit to medical science).

Information will be presented in vignettes which may determine variation in judgements of pain made by caregivers of pain patients.

Identification of judgement biases arising from prior information or in apparent discrepancies between patients' report of pain and their behaviour enables them to be countered by appropriate information and an emphasis on attending to neglected sources of information. It will also help patients to be more aware of the effects of such information and of discrepancies in pain expression on those around them.

8) DESIGN OF THE STUDIES

Caregivers of pain patients are given written vignettes with information about a fictitious pain patient's reported pain intensity (high or low), tendency to avoid/not avoid unpleasant tasks, and tendency to do/not do pleasant activities despite pain. Participants will then be asked to judge the amount of pain they think the fictitious patient is suffering. This is a replication of a study by Chibnall & Tait (1995) with their express permission, with substitution of family member caregivers for students posing as health professionals, and two behavioural variables for their three contextual variables.

Chibnall, JT & Tait RC (1995). Observer perceptions of low back pain: effects of pain report and other contextual factors. *Journal of Applied Social Psychology*, 25, 418-439.

9) **SIZE OF THE STUDY/STATISTICAL ANALYSIS**

Has the methodology and size of the study been discussed with a statistician? If yes indicate with whom – include contact details	No
(a) Is your application for a pilot study?	No
(b) If no, how was the size of the study determined?	It is a replication using the same N as in the original study but fewer variables so is more stringent.
(c) What is the primary endpoint?	Mean difference in each condition between pain level specified in vignette and pain rating by subjects
(d) What is the statistical power of the study?	Minimum alpha = 0.05, power = 0.8
(e) If subjects to be randomised, state method to be used	Random number table

10) **SUBJECTS**

a) Number of patients to be studied	None
b) The number of healthy volunteers caregivers	80 family member
c) Age range	20 years old minimum
d) Method of recruitment	Caregivers of chronic pain patients treated at INPUT Pain Management Unit will be contacted by letter (with consent form) which will be sent to them with the standard invitation to INPUT.
e) Exclusions	None
f) Details of any payments or other inducements to be made to the subjects	
i) Expenses	N/A
ii) Rewards	N/A
g) Are medical students to be involved?	No

11) DETAILS OF PROCEDURES

A: Does the study involve use of a new medicine product or the use of an existing product outside the licence? No

B: (a) Does the study involve a new medical device? No

If yes give details

C: Will any ionising or radioactive substances or X rays be administered? No

D: Are questionnaires to be used? Yes

Judgements of pain are made using a standard Numerical Rating Scale.

Will questionnaires be filled in by subject or administered by someone else?

The questionnaires will be filled in by participants themselves.

If someone else, by whom and where

What published evidence is there of validation of questionnaire design?

Numerical rating scales are widely used for the measurement of pain whether by patients or proxies (Jensen & Karoly 1992).

E: Will the study include the use of audio/video recording? No

F: Are other additional investigations, substances or agents required for the research? No

If yes give details

Jensen, MP & Karoly, P (1992). Self-report scales and procedures for assessing pain in adults. In: DC Turk, R Melzack (Eds.), Handbook of pain assessment (pp. 135-151), New York: Guilford Press.

12) **PHARMACEUTICAL COMPANY INVOLVEMENT**

Does the project involve participation or sponsorship by a pharmaceutical company? No

13) WHAT ASPECTS OF THE PROCEDURES DESCRIBED ARE NOT PART OF ROUTINE CLINICAL CARE?

The entire research is outside routine clinical care and involves no patients, only patients' caregivers. Those who volunteer will be asked to give 20 minutes of their time.

14) THE HEALTH AND COMFORT OF THE SUBJECTS

Will there be any risk of damage to the health of the subjects, or any pain, discomfort, distress, or inconvenience? No

If so please give an assessment of the seriousness of any possible damage to health, and of any pain, discomfort, etc, and the degree of risk N/A

15) CONSENT

- (a) Who will explain the study to the subject?
(If no one please justify.)
Judith Kappesser, Research Assistant
- (b) Will the subject be given a written information sheet or letter?
Relatives will be approached by letter.
- (c) Who will be the signatory if the initial approach is by letter:
Both investigators, i. e. Judith Kappesser and Amanda Williams.
- (d) Consent Form (Available from Medical Committee Office, Block 5, South Wing, St Thomas' Hospital):
Is the standard research consent form to be used? Yes
If not, please justify this departure and submit 16 copies of the substitute form which is to be used.

Caregivers of patients of INPUT Pain Management Unit will be approached by a letter accompanying the letter of invitation to INPUT which all of them receive as part of admission of their family member. Included in this approaching letter will be the consent form. Both, the letter as well as the consent form, are appended to this application form.

- (e) Who will seek the subject's consent?
Judith Kappesser, Research Assistant
- (f) How long will the subject have to decide whether to take part in the study? If less than 24 hours please justify.
As pain patients' relatives are contacted by letter they will have more than 24 hours.

16) INFORMATION TO THE G.P.

Will the General Practitioner be informed? No
If so, how?
If not, please justify No patients are involved.

17) COSTS

Have any arrangements been made to defray costs of the research to the District?
N/A

18) WHAT ARE THE ETHICAL PROBLEMS WHICH APPEAR TO THE APPLICANTS FROM THIS APPLICATION?

Please set them out and add any comments considered likely to assist the Committee

There appear to be no potential hazards or any discomfort or distress to pain patient relatives. Neither related studies already published nor personal communication from other researchers has suggested any discomfort or distress in participants. Rather than being an intrusion, most spouses and family members of chronic pain patients who attend the INPUT Pain Management Unit express particular interest in how they might better understand pain communication by the pain patient who is their spouse, partner or family member.

SIGNATURE OF INVESTIGATOR(S) :

DATE:

SIGNATURE OF HEAD OF DEPARTMENT:

DATE:

CONSENT FORM FOR PARTICIPATION IN RESEARCH PROJECTS & CLINICAL TRIALS

Title of Project: Judgements of pain from vignettes

Principal Investigator: Judith Kappesser Other Investigator/s: Amanda Williams Date: 20 May 2002	Ethics Committee St Thomas Hospital Code No: EC02/074 Version No:
---	--

Dear partner, family member or friend

I am inviting you to take part in a research study which aims to improve our understanding of how those who are close to a chronic pain sufferer judge how bad the pain is from what the patient says about it and how she or he behaves. Knowing another person's pain is difficult, but since you live with someone suffering from chronic pain, you certainly have experience of trying. I am a research worker at INPUT Pain Management Unit and this study is part of my PhD (a higher degree).

I would be grateful if you would consider taking part in this study. It will take about 15 minutes of your time and involves reading some short pieces of information about fictional people suffering from pain and then answering a few questions on your impressions. Apart from this consent form the data is collected anonymously (without your name and address) and all information will be kept strictly confidential. You cannot be recognised from your questionnaire by INPUT staff and I am not involved in the clinical work at INPUT. In addition, you cannot be recognised from any information from this project which is used outside the hospital.

This study has been approved by St Thomas' Hospital Research Ethics Committee. If you wish to get further information, you can contact me at INPUT Pain Management Unit. The number is (020) 7928 9292 extension 1426. Please feel free to ask about anything that is not clear and take time to decide whether or not you wish to take part.

If you wish to take part in the study, please sign the two consent forms (one is for you to keep), fill in the questionnaire added to the consent forms, put both (one consent form and the questionnaire) in the provided envelope and post it.

Thank you for considering this request.

Yours sincerely,

Judith Kappesser MSc
 Research Psychologist

 (Name)

 (Address)

I hereby consent to take part in the above investigation, the nature and purpose of which have been explained to me. Any questions I wished to ask have been answered to my satisfaction. I understand that I may withdraw from the investigation at any stage without necessarily giving a reason for doing so.

 (Date and signature)

3 copies required:- one for researcher, one for patient/volunteer, one for patient's notes

IV.2 Succeeding correspondence

Guy's and St Thomas' Hospital 

NHS Trust

ST THOMAS' HOSPITAL RESEARCH ETHICS COMMITTEE

*Medical Committee Office
Block 5, South Wing*

Co-Chairmen - Dr A Hopper/Dr A Williams
Administrator - Ms S Hirsch

2 May 2002

Ms J Kappesser
Research Assistant
INPUT
St Thomas' Hospital

St Thomas' Hospital
Lambeth Palace Road
London SE1 7EH

Tel: 020 7928 9292

Ext 2097 Fax 020 7922 8163
Stella.hirsch@gstt.sthames.nhs.uk

Dear Ms Kappesser

EC02/074 Patients' family caregivers' judgements of pain
Consent form V 1

The above project was considered at the Research Ethics Committee meeting on 30 April 2002. The Committee required the following amendments/clarification:

Provide a copy of the letter to be sent to patients with the consent form, as stated in 10) of application.

Once I have received written confirmation that the above have been addressed (including a single copy of any amended documents), a final letter of approval will be issued.

Please note that this study carries a reference number, noted above, which must be quoted in any future correspondence.

The St Thomas' Hospital LREC is compliant with the ICH GCP requirements.

Yours sincerely



Dr A Williams
Co-Chairman
Research Ethics Committee

Dr A. Williams
Co-Chairman
Research Ethics Committee
St Thomas' Hospital

8 May 2002

EC02/074 Patients' family caregivers' judgements of pain

Dear Dr Williams

Enclosed is a copy of the letter to be sent to patients with the consent form for their caregivers. I know that from an ethical point of view it can be an issue to ask for reasons for not taking part in a study, but as you know it is very important for conclusions concerning generalisation of the study. I have, therefore, worded the reply carefully and made it anonymous. I hope you are happy with this.

Yours sincerely

Judith Kappesser
Research Assistant

IV.3 Letter to chronic pain patients

Dear

I am writing to you to ask you for your help concerning a research study I am doing as part of my PhD (a higher degree); I am a research worker at INPUT Pain Management Unit.

The study aims to improve our understanding of how those who are close to someone with chronic pain judge how bad the pain is. Of course, knowing someone's pain is extremely difficult, but people who are close to pain sufferers certainly have experience of trying.

If you can help, please could you give the enclosed envelope to the person who is closest to you and knows most about your pain. This may be your partner, husband or wife, a relative (child, parent, brother, sister or any other relative), or a friend. What is important is that you feel she or he is close to you and knows about your pain.

If you cannot help, I would be grateful if you would briefly tell me why below and send this section back in the envelope provided. This will be anonymous (without your name and address).

This study has been approved by St Thomas' Hospital Research Ethics Committee. If you wish to get further information, you can contact me at INPUT Pain Management Unit. The number is (020) 7928 9292 extension 1426. Please feel free to ask about anything that is not clear.

Thank you for considering this request.

Yours sincerely

Judith Kappesser MSc
Research Psychologist

Version _____

Reasons why I cannot help. Please tick the box for the answer that seems most suitable.

- Nobody knows me/my pain well enough.
 - The person I chose cannot take part.
 - Other reason (please give details if you wish).
-

ST THOMAS' HOSPITAL RESEARCH ETHICS COMMITTEE

Medical Committee Office
Block 5, South Wing

Co-Chairmen - Dr A Hopper/Dr A Williams
Administrator - Ms S Hirsch

8 May 2002

Ms J Kappesser
Research Assistant
INPUT
St Thomas' Hospital

St Thomas' Hospital
Lambeth Palace Road
London SE1 7EH

Tel: 020 7928 9292

Ext 2097 Fax 0171 922 8163
Stella.hirsch@gstt.sthames.nhs.uk

Dear Ms Kappesser

EC02/074 Patients' family caregivers' judgements of pain Ms J Kappesser, Dr A Williams
Consent form V 1, letter to patients

Thank you for addressing the queries raised by the Research Ethics Committee at its meeting on 30 April 2002. This is satisfactory and I am happy for the study to commence. Approval extends to the Guy's site.

Please note the following conditions to the approval:

- The project number and the principal investigator must be clearly stated on the consent form (if applicable). If approval is given to named investigators only, these names must also be stated on the form.
- In the case of research on patients, a copy of the consent form (if applicable) must be placed in the patient's medical records, together with a note of the date of commencement of his/her participation in the research. A label must appear on the outside cover of the records when the patient is participating in the research.
- Any amendments to the protocol must be notified to the committee for approval.
- Approval is for the length of time specified in your application. If you require an extension, a letter from the principal investigator to the Chairman, is required to extend the research.
- The committee should be notified of any serious adverse events (please apply for standard SAE report form), or if the study is terminated prematurely
- The investigators must adhere to the published Guidelines of the Committee and provide the Chairman with annual progress reports and an end of study report. The research should start within 12 months of the date of approval.

This project carries a reference number, noted above, which must be quoted in any future correspondence.

The St Thomas' Hospital LREC is compliant with the ICH GCP requirements.

Yours sincerely



Dr A Hopper/Dr A Williams
Co-Chairmen
Research Ethics Committee

Encl.

IV.4: Questionnaire

Questionnaire

The following questionnaire consists of three parts:

1. On the first page you will be asked to answer some questions about yourself.
2. In the second part (page 2-5) you will be asked to read four stories about a fictional person in pain and to answer some questions about your impression of this person.
3. On page 6 you will be asked some questions about your experience of being close to someone with chronic pain.

On the very last page of this questionnaire you will find space to make any comments on the questionnaire or for anything else that you would like to let me know.

After finishing with the questionnaire please remember to put one of the signed consent forms with the questionnaire in the envelope provided. The other consent form is for you to keep.

Thank you for participating!

III. Lastly, some questions which are concerned with your experience of being close to somebody who has chronic pain. Please tick the box for the answer that seems most suitable.

1. Because of the pain do you think that your relationship with the person in pain

has worsened a lot or has worsened or has not changed or has improved or has improved a lot

2. The person in pain does things he/she likes despite the pain or has stopped doing things he/she likes because of the pain.

3. The person in pain does things he/she dislikes despite the pain or has stopped doing things he/she dislikes because of the pain.

4. (a) If the person in pain has stopped doing some things he/she dislikes because of the pain, do you now do more of those things? yes or no

(b) If you answered yes, how fair do you think it is that you are now doing the things the person in pain is not doing because of the pain?

0 1 2 3 4 5 6
 very unfair very fair

5. Do you feel that the person in pain tries to make up for the things you do for him/her because of the pain?

0 1 2 3 4 5 6
 not at all very much

6. Do you think that if you had chronic pain, the person in pain would do what you do for him/her?

0 1 2 3 4 5 6
 very unlikely very likely

7. How satisfied are you with your relationship to the person in pain?

0 1 2 3 4 5 6
 not at all satisfied highly satisfied

8. People with chronic pain express their pain to different extents.

(a) Do you think the person in pain you know tends to

0 1 2 3 4 5 6
 hide his/her pain make a big deal/ complain strongly about his/her pain

(b) Do you think pain sufferers in general tend to

0 1 2 3 4 5 6
 hide their pain make a big deal/ complain strongly about their pain

You are now finished with the study. On the back of this page is space for any comments you may have. Thank you for participating!

Now please put one signed consent form with the added questionnaire in the provided envelope and post it. Remember, the other consent form is for you to keep. THANK YOU!

If you have any **comments** on the questionnaire or anything else that you would like to let me know, please use this space.

IV.5 Feedback sheet

Pain intensity as judged by relatives of chronic pain patients. Which factors do matter?

Introduction

Factors impacting on pain judgements of health care professionals have often been investigated. However, not very much is known about factors that might have an effect on pain judgements of lay people. Therefore, aim of the present study was to investigate which of a limited set of factors have an impact on pain judgements made by lay people close to chronic pain patients.

Previous studies found that health care professionals judge pain to be more intense if medical findings supporting the pain were present and if the intensity of pain as experienced by the sufferer was high. Accordingly, both were included in the present study. Further, it was investigated whether (dis)continuation of (un)pleasant tasks had an effect on pain judgements of lay people close to chronic pain patients. These variables were included due to the extensive literature about suspicion of others' pain: Possibly behaviours that seem unfair in the sense of taking advantage of one's pain by avoiding unpleasant but continuing pleasant tasks could lead to lower pain judgements than other behaviours (i.e. continuing pleasant and unpleasant tasks, avoiding pleasant and unpleasant tasks, continuing unpleasant and avoiding pleasant tasks).

Method

127 participants who were close to a chronic pain patients read four vignettes in which hypothetical pain patients were described with regard to their experienced pain, the presence or absence of medical findings and the continuation or discontinuation of pleasant or unpleasant tasks. After reading each vignette participants were asked to give their opinion on the pain patients' pain intensity.

Results

Results revealed that pain intensity as reported by the patient as well as (dis)continuation of pleasant tasks in association with (dis)continuation with unpleasant tasks had an effect on pain judgements. Medical findings, on the contrary, did not matter.

Relatives rated patients' pain intensities higher when the reported pain was high. They, further, rated pain intensity to be higher for patients who stopped rather than continued doing pleasant tasks. Pain of patients who stopped pleasant and unpleasant tasks was rated as most intense, followed by pain of patients who stopped pleasant but continued unpleasant tasks. Pain of patients who continued pleasant and unpleasant tasks was rated as higher and pain of patients was rated lowest if they continued pleasant and stopped unpleasant tasks.

Results further revealed that pain intensities as rated by participants differed significantly from pain intensities reported in the vignettes: When the reported pain intensity was low, participants rated the patient's pain significantly higher. When the reported pain intensity was high, participants rated the patient's pain significantly lower.

Discussion

This study found similarities and differences between pain judgements of lay people and health care professionals. Similar effects were obtained with regard to pain intensity: the higher the reported pain intensity, the higher they judged the pain to be. However, when rated and reported pain intensity were compared directly, both groups rated low pain intensities as higher and high pain intensities as lower than the patients reported their pain to be. With regard to medical findings, however, there were differences between both groups: judgements of lay people were not affected by medical findings whereas judgements of health care professionals have repeatedly found to be affected.

Investigating the effect of behaviours regarding pleasant and unpleasant tasks obtained promising results: as assumed, continuing pleasant and avoiding unpleasant tasks led to lowest pain judgements in relatives

Lebenslauf

Name	Judith Kappesser
geboren	6. Februar 1973 in Mainz
Schulbildung	1979-1983 Martinusgrundschule, Mainz 1983-1989 Rabanus-Maurus Gymnasium, Mainz 1989-1990 Deutsche Schule Rom 1990-1992 Rabanus-Maurus Gymnasium, Mainz 1992 Abitur
Studium	1992-1998 Psychologiestudium an der Johannes Gutenberg-Universität, Mainz 1998 Diplomprüfungen 1998-1999 Studium der Gesundheitspsychologie an der Universität Bath gefördert durch ein Postgraduiertenstipendium des Deutschen Akademischen Austausch Dienstes 1999 MSc in Health Psychology
Promotion	2001-2004 Dissertation an der Fakultät für Informations- und Kognitionswissenschaften, Eberhard Karls Universität Tübingen gefördert durch ein Promotionsstipendium der Studienstiftung des deutschen Volkes 2002-2003 empirische Studien für die Dissertation in INPUT Pain Management Unit, St Thomas' Hospital London 2004 Promotion

**klinisch-
psychologische
Ausbildung /
Tätigkeit**

- seit 2000 Ausbildung zur psychologischen Psychotherapeutin (Verhaltenstherapie; bis 2002 IFKV in Bad Dürkheim, seit 2003 Johannes Gutenberg-Universität Mainz)
- 2000-2002 Psychologin im Praktikum in der Psychosomatischen Fachklinik Bad Dürkheim
- 2003 Zwischenprüfung
- seit 2003 psychotherapeutische Tätigkeit an der Poliklinischen Institutsambulanz für Psychotherapie, Johannes Gutenberg-Universität Mainz