Mind the gap : explanations for the differences in utilities between respondent groups

Peeters, Y.

Citation

Peeters, Y. (2011, May 11). *Mind the gap : explanations for the differences in utilities between respondent groups*. Retrieved from https://hdl.handle.net/1887/17625

Colofon

ISBN: 978-90-8570-745-5
Cover illustration : with help of Jan Edelaar
Lay-out: with help of Janneke van der Niet
Printed by: CPI wöhrmann print service

Mind the Gap:

Explanations for the differences in utilities between respondent groups.

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van de Rector Magnificus prof. mr. P. F. van der Heijden,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 11 mei 2011
klokke 15.00 uur

door

Yvette Peeters

geboren te Venlo in 1981

Promotiecommissie

Promotor:	Prof. dr. A. M. Stiggelbout
overige leden:	Prof. dr. A. V. Ranchor (Universitair Medisch Centrum Groningen) Prof. dr. M. A. G. Sprangers (Universiteit van Amsterdam) Prof. dr. P. A. Ubel (Duke University,USA) Dr. T. P. M. Vliet Vlieland

The printing of this thesis was financially supported by the Department of Medical Decision Making of the LUMC and the Dutch Arthritis Association.

The work presented in this thesis was financially supported by the Netherlands Organization for Scientific Research NWO Innovational Research Incentives (grant number 917.56.356).

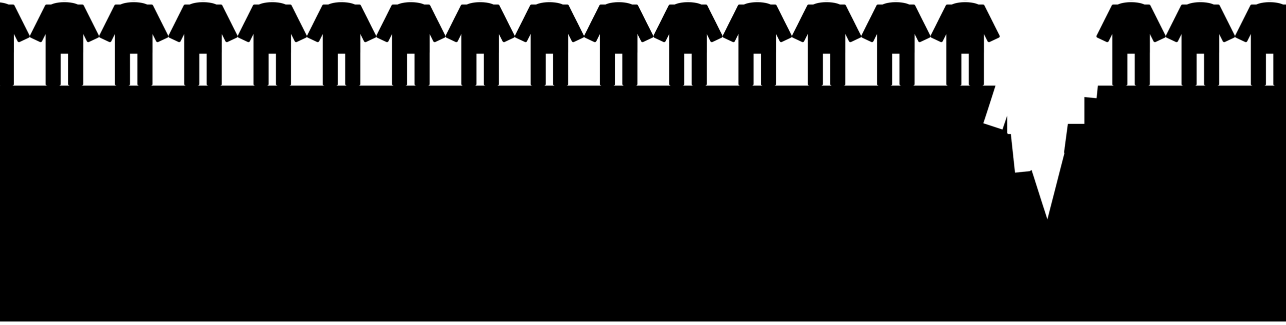
*W*atch with glittering eyes the whole world around you, because the greatest secrets are always hidden in the most unlikely places. Those who don't believe in the magic will never find it.

- Roald Dahl -

Contents

1

General Introduction



1.1 Cost-Utility analyses

In medical care resources are scarce and choices have to be made about how these resources are to be distributed. For example, should we vaccinate all Dutch students against mumps[1] or should we introduce cartilage transplant for patients with rheumatoid arthritis,[2] or maybe both? To judge the optimal allocation of medical resources, systematic economic evaluations are needed, comparing costs with the benefits of health services.[3] For these economic evaluations different techniques can be used, among which cost-utility analysis. Cost-utility analysis compares the costs of treatment to the outcomes obtained.

In cost-utility analysis preferences for a certain set of outcomes are measured using health state utilities. Health state utilities are strongly related to health related quality of life (HRQL) but they differ in that health state utilities measure both quality of life and the valuation of this quality of life compared to perfect health and death.[4] Health state utilities are values between 0 and 1 that represent individuals' preferences for health states. Preferences are elicited using different methods such as the Standard Gamble (SG), the Time-Trade-Off (TTO), and the Visual Analogue Scale (VAS).[5]

In the SG participants are asked to choose between a certain outcome, the health state to be valued, or a gamble with a probability (p) of receiving the best possible outcome, perfect health, and a probability $(1 - p)$ of receiving the worst possible outcome, usually death. By varying p, the indifference point is searched, the probability at which the participant is not able to choose between the gamble and the sure outcome. The value obtained is the utility for the health state under valuation $(\mu = ((p \cdot 1) + ((1 - p) \cdot 0) = p))$. In the TTO participants are asked to choose between a number of years living in the health state to be valued or living a shorter period of time in perfect health. The time in perfect health is varied to obtain an indifference point, the number of years in perfect health equal to a higher number of years in the health state to be valued. The health state utility is calculated by $\frac{\sharp\ years\ in\ perfect\ health}{\sharp\ years\ in\ health\ state\ to\ be\ valued}$[a]. In the VAS, participants are asked to give a valuation for the health state to be valued by placing a mark on a 100 mm. horizontal line ranging from perfect health to the worst possible outcome, usually death. The health state utility is the number of mm. from the death anchor divided by 100.

Which method should be used when eliciting health state utilities has been

---

[a] For states worse than death slightly different procedure is used. But in this thesis only the described procedure is used.

topic of debate.[4] Initially the SG had a reputation of being the gold standard since it meets the axioms, as described by Neumann and Morgenstern, of expected utility theory.[6] Nowadays the feasibility and validity of the SG is questioned. Participants experience the SG as a complex method,[7] and answers elicited by the SG are vulnerable to probability weighting.[8,9] The TTO on the other hand, is simpler to elicit, is not vulnerable to probability weighting, and appears to have good face validity.[4,7] Nevertheless the TTO is vulnerable to other biases, but these biases probably cancel one another out. The bias of the utility curvature which is downwards makes up for the upward bias cause by loss aversion and scale compatibility.[9] Furthermore, the time-line used in the TTO gives a good representation of decisions made in clinical settings.[10] The VAS is often used because of its feasibility, it is easy to elicit but its construct validity has been questioned.[7] Given the above reasons the TTO is now the most often used method to elicit health state utilities.

Health state utilities can be elicited directly, by asking patients or members of the public to give valuations to health states, or with indirect instruments. In studies using indirect utility instruments, health state utilities of members of the general public are based on patients' answers to a short descriptive questionnaire. These answers are fed into a model estimated from an earlier study,[11] which generates the utility values of the general public. The EQ-5D-tariff is such an indirect instrument that is widely used. The EQ–5D consists of five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension is described according to one of three levels of severity: no problems(1), some problems (2) and extreme problems (3). In total the EQ-5D can thus create 243(3^5) theoretically possible health state descriptions. A selection of these health state descriptions has been valued by a large sample of members of the public, and based on these valuations a model has been estimated from which utilities for each of the 243 descriptions can be generated.[11]

Quality-Adjusted Life-Years (QALY) can be computed based on these elicited health state utilities. QALYs define the overall utility for a certain time path or life expectancy. To explain the concept of QALY, I revert to the cartilage transplant as illustration. When patients with RA receive cartilage transplant their utility might deteriorate initially (due to surgery) but will reestablish over time (assuming that this treatment will give no long term side effects). For example, a patient with a life expectancy of 30 years might initially have a health state utility of 0.6. Due to surgery this utility will deteriorate to 0.5 for a year, however after this year it will increase to a health state utility of 0.8 which will remain for the next 29 years.

The QALY is then $(1 \cdot 0.5) + (29 \cdot 0.8) = 23.7$. The utility of patients who do not receive the transplant will not deteriorate initially (they do not have to recover from surgery), but over a longer period of time these patients will continue having pain complaints affecting their utility. For example a patient with a life expectancy of 30 years will continue to have a health state utility of 0.6 for 30 years long; leading to a QALY of $(30 \cdot 0.6) = 18$. Gain in QALYs from transplant is computed by comparing the QALYs of patients with transplant compared to the QALYs of patients without transplant. In cost-utility analyses this gain in QALYs will be compared with the costs that have to be made, resulting in cost per QALY gained.

1.2 Public or patients' preferences

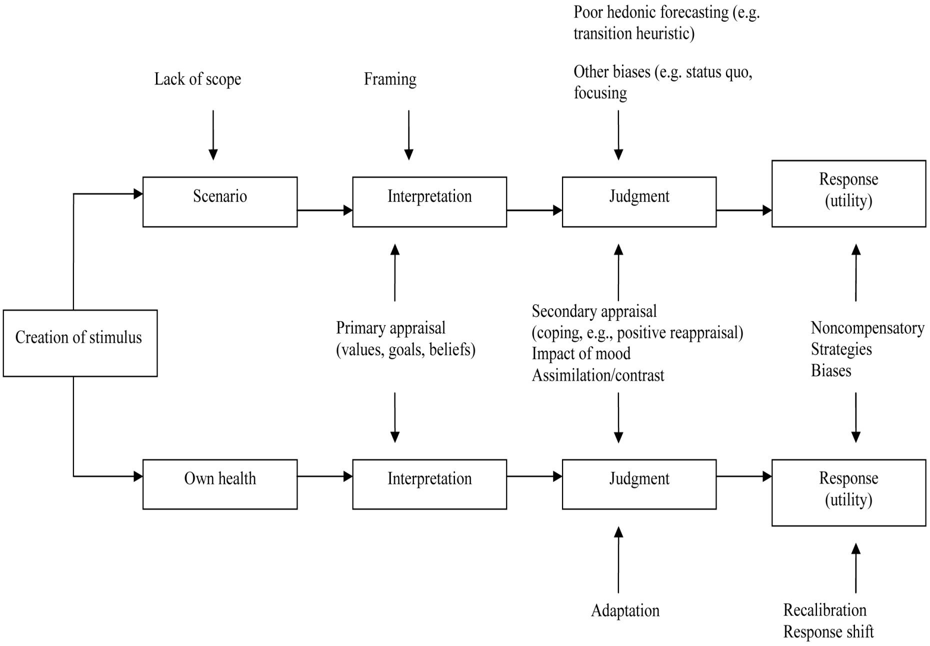
Cost-utility analyses for allocation decisions are recommended to be made from the societal perspective. This implies that health state utilities should be elicited from members of the public. Since cost-utility analyses should lead to a just allocation of resources these analyses should not only be based on the opinion of those who gain health but also on that of those who pay for it.[12] Organizations involved in developing guidelines on the use of new and existing treatments, such as the National Institute for Health and Clinical Excellence (NICE), the panel of the U.S. Public Health Service, and the Dutch Health Care Insurance Board (CvZ), advise the use of the societal perspective, in which health state utilities elicited from a fully informed representative sample of members of the public are preferred.[12–14] However it might be challenging to fully inform members of the public. Instead, health state utilities of patients might be informative given that certainly patients experiencing an illness are well-informed.[12, 15] The panel of the U.S. Public Health Service already suggested that in cost-utility analyses in which alternative interventions are compared patient preferences might be the better choice.[12]

Whose' preferences are used in cost-utility analyses does matter. Preferences of members of the public are often found to be lower than patients' preferences.[16] Several explanations for this gap in health state valuations between patients and public have been provided by research from different fields.[17–19] To make more evidence-justified suggestions about whose preferences to use, the mechanisms underlying this gap in health state utilities have to be understood.[15] Whose valuations are most valid depends on the explanations for this gap. Is this gap caused by errors in the method used, or rather due to cognitive mechanisms?[18]

1.3 Mechanisms underlying the gap between patients and public

Stiggelbout and de Vogel-Voogt[15] systematically described mechanisms that might cause the gap between utilities given by patients and members of the public, by using stimulus response models (Figure 1.1). This resulted in a framework in which the different valuation processes of patients and members of the public are presented simultaneously. A short description of the mechanisms provided in their framework, and of mechanisms mentioned by other researchers is provided below. The outline of the mechanisms described here is not intended to be conclusive. By combining information from different research fields one can always come up with additional mechanisms that are more or less related to the ones described below.

Figure 1.1 Framework of mechanisms underlying the gap by Stiggelbout & Vogel-Voogt[15]



Lack of Scope When eliciting health state utilities patients are generally asked to value their own health of the previous week, whereas members of the public have to value a health state based on a description. Health state descriptions can be developed by researches based on experience of physicians or patients,[20] or they can be based on classification systems, such as the Health Utility Index (HUI)[21] or the EQ-5D.[22] A lack of correspondence between health state descriptions and the actual experience of a health state might cause the gap in health state valuations between patients and public.[23] Insinga and Fryback[23] found that participants gave different health state valuations for their own experienced health than for an EQ-5D health state description of their own health. Possibly the EQ-5D health state description is too sparse. Jansen et al.[20] found similar results in a sample of patients undergoing radiation therapy. The own experienced health, while being treated with radiation therapy, was valued higher than the health state description of this radiation therapy.

Framing Framing of the health state description influences how a health state is interpreted. Most health state descriptions tend to include only limitations and handicaps caused by the health state. Due to this negative framing, members of the public focus on the limitations of a health state whereas patients might also think about positive aspects in their lives.[24]

Focusing illusion The fact that members of the public focuses on limitations is probably not only caused by the negative framing of health state descriptions. People have a natural tendency to focus on the difference between their current situation and an imaginary situation; they will overestimate the differences and overlook the similarities.[25] This focusing illusion has e.g. been demonstrated among assistant professors who were asked to imagine that they would not achieve tenure[26] and among football fans.[27] However, among members of the public imagining a disability no evidence was found for this focusing illusion.[24]

Status Quo Bias Status quo bias indicates that people value goods more highly when they own them. Evidence for status quo bias has been shown previously. Participants who were randomly assigned to a car would not part with this car, even if they were given the opportunity to choose a different car without penalty.[28] In medical decision making evidence for this status quo bias has also been found. Salkeld et al.[29] studied preferences for a bowel cancer prevention test. On average patients were willing to pay more for a test that they had used before (status quo) instead of starting to use a new test. Both tests were equal on all attributes. Regarding the gap in health state valuations status quo bias might cause patients to give higher valuations. Patients valuing their own life are probably less willing to trade-off own life years than members of the public valuing a hypothetical health state.[15]

Loss Aversion Related to status quo bias is loss aversion. People evaluate outcomes as gains and losses and are more sensitive to losses than to gains. For instance people value the loss of €10,- worse than the gain of the same amount of money. The difference between losses and gains in health state valuations depends on the reference point.[30] In patients whose reference point is their own health, the loss of life years that a patient has to trade will get more weight than the health that is gained leading to an upward bias.[9, 15] Further, for patients trading of life years or increasing risk of dying, it might feel as tempting fate. Members of the public are probably less concerned about trading life years or increasing risk of dying since the situation remains hypothetical.

Adaptation Adaptation can be defined as a response that diminishes or remains constant over time despite and increase in the stimulus.[31] When confronted with adverse circumstances such as an illness people tend to adapt peculiarly well.[31] Therefore adaptation is often suggested to explain the patients' relative high reported quality of life.[32–36] Adaptation takes place on physical and psychological level.[31] Physically patients learn to handle handicaps and mentally they learn to deal with the illness. Different processes are suggested to initiate psychological adaptation, among which coping strategies and benefit finding. When providing health state valuations patients will include their ability to adapt whereas members of the public may fail to anticipate on this ability to adapt.[37] Members of the public have the tendency to overpredict the duration of emotional reactions to future events[27] and underestimate their cognitive mechanisms which alleviate this reaction.[26] Tentative support has been found for this failure to anticipate on adaptation. Members of the public who were made aware of their ability to adapt gave higher valuations on a person trade-off (PTO) and on a VAS measuring quality of life,[24, 38] but not on the TTO and SG.[39]

Valuation shift Dolan[40] suggested that experiencing poor health might result in higher valuations of other hypothetical health states, a process they called valuation shift. Dolan showed that participants in poor health assigned higher valuations to various EQ-5D scenarios than did participants in good health. Scale recalibration To measure health state utilities, subjective scales are used, which are susceptible to different interpretations between people, but more importantly within people.[37] When people experience illness they might change their internal standards, leading to a change in interpretation of these scales.[41] For instance a patient with RA who first valued her joint pains as 8 on a VAS scale ranging from 0 (no pain) to 10 (major pain), recalibrated her pain to a 5 after experiencing kidney stones. The pain she experienced due to kidney stones was significantly more intense than any pain she had ever experienced before, resulting in a recalibration of her valuation of major pain.

Implicit theories of stability and change Implicit theories of stability and change are heuristics that people use to recall emotions. To recall emotions, people first note their present status and then decide if their status has changed over time. This reconstruction of emotions is guided by theories that include specific beliefs regarding the inherent stability of an attribute.[42] For instance people have the implicit belief that they will become happier over time. When people are asked to give an estimation of their previous happiness, they assume that they had been less happy than they are now.[43] Depending on the method used to investigate health state utilities, implicit theories might cause bias. Such may be the case in the increasingly popular method of asking patients to recall how their health state has changed over time.[37]

1.4 Objective and outline of the thesis

Several mechanisms have been suggested to cause the gap between valuations given by patients and members of the public, of which a number have been examined empirically. However still no conclusive suggestions can be made, and more research is necessary to enhance our knowledge. Although adaptation is often mentioned it has never been tested empirically. Other mechanisms have only been studied among members of the public and not among patients, such as focusing illusion, or the reverse, lack of scope. The overall objective of this thesis is to further examine some of the mechanisms proposed to cause the gap between health state valuations, in order to gain insight in the relative validity of health state utilities of patients and of members of the public.

In Chapter 2 first a meta-analytical comparison of health state valuations of patients and members of the public is presented. Previously, studies described contrasting findings[16, 44] about the difference in health state valuations between patients and members of the public. The aim of our study was to investigate the influence of respondent group on health state valuations. Post hoc, other design-effects were tentatively studied using moderator analyses.

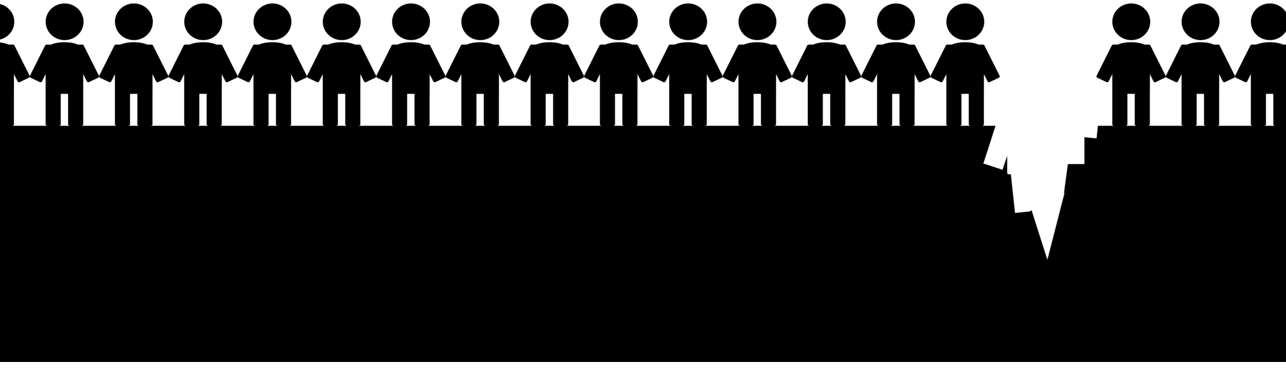
In Chapters 3 through 8 mechanisms potentially underlying the difference between patients and members of the public were studied. Chapter 3 starts with the influence of lack of scope and framing of a health state description. Patients with RA valued their own experienced health, an EQ-5D description of their own health, and an enriched EQ-5D description of their own health. These valuations were compared to investigate the influence of differences in health state descriptions. Next, in Chapter 4 the effects of focusing illusion and adaptation were examined, as well as the sparseness of the EQ-5D description (lack of scope). In this study open-ended questions were used to assess aspects important to patients with RA and to members of the public imagining having RA. In Chapter 5 the effect of lack of scope and framing was investigated further. Here the effect of a health state description was not only investigated among patients, but also among partners of patients and among members of the public. All participants valued their own imagined/experienced health state, a standard EQ-5D description of this health state, and an enriched EQ-5D description of this health state. By comparing the valuations given by partners of patients to the valuations of patients and of members of the public the effect of vicarious experience could also be examined. Chapter 6 describes a cross-sectional study among patients with RA in which the effect of adaptive abilities on health state valuations is examined. Adaptive abilities were based on Cognitive Adaptation Theory (CAT) as suggested by Taylor.[32,45] Chapter 7 further describes adaptation and valuation shift, investigated in a longitudinal study among patients with Spinal Cord Injury (SCI). Health state valuations of patients with recent onset acute SCI were assessed at three points in time. In Chapter 8 the effect of adaptation was also examined. Here the ability to anticipate on adaptation by patients experiencing new adversities as well as the effect of implicit theories of stability and change were studied.

While examining the mechanisms suggested to cause the difference in health state valuations between patients and members of the public we were challenged by often ambiguous descriptions of these mechanisms. Among others we felt that the language used by “response shift” gathers together different terms already existing in the scientific literature. In Chapter 9 the conceptual confusions related to the language of response shift is described.

2

Health State Valuations of Patients and the General Public Analytically Compared

Peeters, Y. & Stiggelbout, A.M. (2010). Health State Valuations of Patients and the General Public Analytically Compared: A Meta-Analytical Comparison of Patient and Population Health State Utilities. Value in Health, 13, 306-309.



Abstract ***Objectives:*** *To obtain quality-adjusted life-years, different respondent groups, such as patients or the general public, may be asked to value health states. Until now, it remains unclear if the respondent group has an influence on the values obtained. We assessed this issue through metaanalysis.* ***Methods:*** *A literature search was performed for studies reporting valuations given by patients and nonpatients. Studies using indirect utility instruments were excluded.* ***Results:*** *From 30 eligible studies, 40 estimators were retrieved revealing a difference between respondent group (Cohen's* $d = 0.20, p < 0.01$*). When elicitation methods were analyzed separately, patients gave higher valuations than nonpatients using the time trade-off (TTO)* ($N = 25$*, unstandardized* $d = 0.05, p < 0.05$) *and the visual analog scale (VAS)* ($N = 22$*, unstandardized* $d = 0.04, p < 0.05$)*. When the standard gamble was used, no difference was seen* ($N = 24$*, unstandardized* $d = 0.01, p = 0.70$)*.* ***Conclusion:*** *In contrast with Dolders et al., our results show that patients give higher valuations than members of the general public. For future cost-utility analyses, researchers should be aware of the differential effects of respondent group for the elicitation methods TTO and VAS.*

2.1 Introduction

Valuations used in decision analyses and cost-utility analyses can be given by different groups, such as patients or the general public. Three studies have investigated the effect of response group by summing results of empirical studies,[16],[4446]. Two of these studies, a review, and a meta-analysis of prostate cancer utilities, found higher valuations given by patients. The third, a meta-analysis on varying patient groups, did not find any difference. The latter two included indirect utility instruments like the European Quality of Life Five Dimensions EQ-5D-tariff[11] or Health Utilities Index Mark (HUI)[47] and included multiple health state valuations from the same study sample.

In studies using indirect utility instruments, only patients are approached to participate, members of the public are not included as a separate sample. Such studies calculate health state utilities of members of the general public from patients' answers to a short questionnaire. These answers are put in a model captured from an earlier study[11] which generates the utility values of the general public. Therefore, including more than one study using indirect utility instruments leads to multiple health state valuations from the same subject sample, which is a violation of the assumption of independent data points. This may have led to a distortion of the standard error, an inflated sample size, and an overrepresentation of certain studies.[48] The aim of our study was to investigate through meta-analysis the influence of the respondent group on valuations avoiding this bias.

2.2 Methods

2.2.1 Search and retrieval of studies

Studies reporting valuations given by patients and by members of the general public, professionals, or proxies (which we from now on refer to as "nonpatients") were retrieved through the computerized databases PsychInfo and PubMed. Studies published between 1970 and October 2008 were searched using preferences, utility, patient, public and, respectively, time trade-off (TTO), standard gamble (SG), or visual analog scale (VAS) as key words. With the so-called snowball method, the bibliographic information of De Wit et al.,[16] Dolders et al.,[44] Bremner et al.,[46] and other retrieved studies were searched for additional studies. With the database Web of Science, we retrieved studies for the citations of the already retrieved studies.

Abstracts were examined regarding the inclusion criteria. Studies were included if they reported valuations of both patients and nonpatients, used a standard utility method (TTO, SG, or VAS), included participants 18 years, and were written in English. Studies that used indirect instruments (classification systems), that investigated mental health states, or in which nonpatients answered what they thought the patient would have answered, were excluded.

2.2.2 Data extraction

A detailed coding system was used to extract data. From each study, the mean valuations and SDs for each evaluated health state were coded for every group. If these data were not reported, authors were contacted. We excluded studies when the authors did not respond after three attempts or could not reveal the mean valuations. If only the SDs were missing, we estimated these by the weighted sum of the SDs reported in the included studies. We further coded: elicitation method, nature of the nonpatient respondent group, and various types of information about the health state description used. With the elicitation method it was coded if the TTO, VAS, or SG was used. Non-patient respondent groups were coded as professionals/proxies or members of the general public. Information about the health state description included three aspects. First, it was recoded if the patients valued a description or if they valued their own experienced health state. Second, it was denoted what kind of health state description was used; a standard EQ-5D health state description, a standard HUI health state description, or a specifically developed health state description. Thirdly, it was coded if the health state description provided an illness label. Information of the retrieved studies was independently rated by two judges (A.M.S. and Y.P.) with satisfactory agreement for most variables (Cohen's κ between 1 and 0.77). Agreement on the variable "own health state or hypothetical health state" was low, (Cohen's $\kappa = 0.61$) in three of 30 ratings the judges disagreed. All dissimilar ratings were compared and discussed until agreement was found.

2.2.3 Statistical analyses

Before all meta-analyses, the standard mean differences and sample sizes were checked for outliers. One outlier for the sample size of nonpatients was found. Specifically, Smith et al.[49] included 567 nonpatients. Studies with larger sample size are given more weight as these are assumed to be more precise. In such weighted estimation, studies with extremely large sample size can define the entire meta-analysis if these are given according weights.[50] Therefore, we recoded this study sample into the highest nonextreme sample size of nonpatients ($N = 246$). Next, we compared the results obtained with the original sample size to those obtained with the recoded sample size. Because the results remained almost unchanged, we present the data including the original sample size.

One overall meta-analysis and three subanalyses by elicitation method were performed. Before any of the analyses, data within each of the retrieved studies were combined. If more than one health state was valued in one single study, a meta-analysis on the level of this primary study was performed. The differences between patients and nonpatients were estimated for each health state and were then combined into one estimator through metaanalysis. This estimated mean difference was then used as estimator for this study in the overall meta-analysis. In studies that included more than one respondent group in either the patient or the nonpatient group, estimations of both subsamples were included. The sample size of the other group was divided by two, and used twice to compare each of the subsamples. In studies using more than one elicitation method, a meta-analysis on the level of the primary study was performed.

Using the software Comprehensive Meta-Analysis (version 2.2.046),[51] the standard mean difference, Cohen's d, and 95% confidence interval were estimated.We used Cohen's d to control for the difference in the numerical scales of TTO, SG, and VAS. For each analysis by elicitation method, the unstandardized difference was estimated, instead of Cohen's d.

The homogeneity of the sample was checked with the Q-statistic.[52] If the sample of reports appeared to be heterogeneous, random effect models were used and moderator variables were analyzed to investigate if these could explain this heterogeneity. The significance of the six moderating effects was checked using the Q-statistic. A significant contrast means that the moderator variable explains some of the heterogeneity between the groups, but it does not necessarily imply that one of the subsamples is homogeneous. For each subsample, we again investigated the Q-statistic and Cohen's d. The Duval and Tweedie's trim and fill procedure[53] gave no indication for publication bias in the overall meta-analysis, nor in the subanalyses.

2.3 Results

The search yielded 36 studies of which 30 could be included in the analyses. Two studies were excluded due to differences in elicitation method used for patients and non-patients[54, 55] and two studies were excluded since the reported data was already included in another study.[56, 57] In another two studies the same group of non-patients was used.[58,59] We decided to divide the sample size of this group of non-patients by two and keep the estimations of both studies in the analyses.

Of the remaining 31 studies, five studies reported other data than mean valuations. The authors of these studies were contacted. From three of these studies the authors sent the mean valuations and standard deviations by mail.[60–63] Of one study additional not-reported data was sent.[63] No mean valuations and standard deviations could be retrieved from the other two studies.[64,65] In Appendix A data of the included studies is shown.[16,49,58–63,66–85] In 23 studies, participants rated more than one health state, and in 13 studies, more than one elicitation method was used. In these studies, meta-analyses on the level of the primary study were performed.

2.3.1 Overall meta-analysis

From the included 30 studies, 40 mean differences in health state valuations between patients and nonpatients, from now on referred to as "estimators," were extracted. The total set of estimators was heterogeneous [$Q(39) = 398.25, p < 0.01$]. Using the random effects model, the overall combined effect size for the total set was significant (Cohen's $d = 0.20, SD = 0.06, p < 0.01$). Patients gave higher valuations compared to nonpatients. Figure 2.1presents the standardized mean differences for each study. Two moderators showed a significant contrast (Table 2.1).

Patients' and nonpatients' valuations were more distinct when no label was provided than when it was. Furthermore, valuations were more similar between groups when they both valued a health state description than when patients valued their own health. In terms of heterogeneity, the Q-statistic reveals that all subsamples remain heterogeneous, except for the subsample of studies without illness label. We want to emphasize that this sample consisted of only three studies. Because this subsample was homogeneous, the fixed effect model was used to test the group difference. For each subsample, the group difference is reported as Cohen's d.

2.3.2 Meta-analysis of studies by estimation method

The set of 25 TTO estimators was heterogeneous [$Q(24) = 263.85, p < 0.01$]. The overall combined effect size revealed a difference between the response groups unstandardized $d = 0.05, SD = 0.02, p < 0.05$). Moderator analyses showed a significant contrast between studies with own health and studies with a health state description [$Q(1) = 5.93, p < 0.01$]. When patients valued their own health ($N = 3$), their valuations were different from those of nonpatients (unstandardized $d = 0.24, p < 0.01$). When both groups valued a health state description ($N = 22$), the

Figure 2.1 The 40 mean differences from the 30 included studies

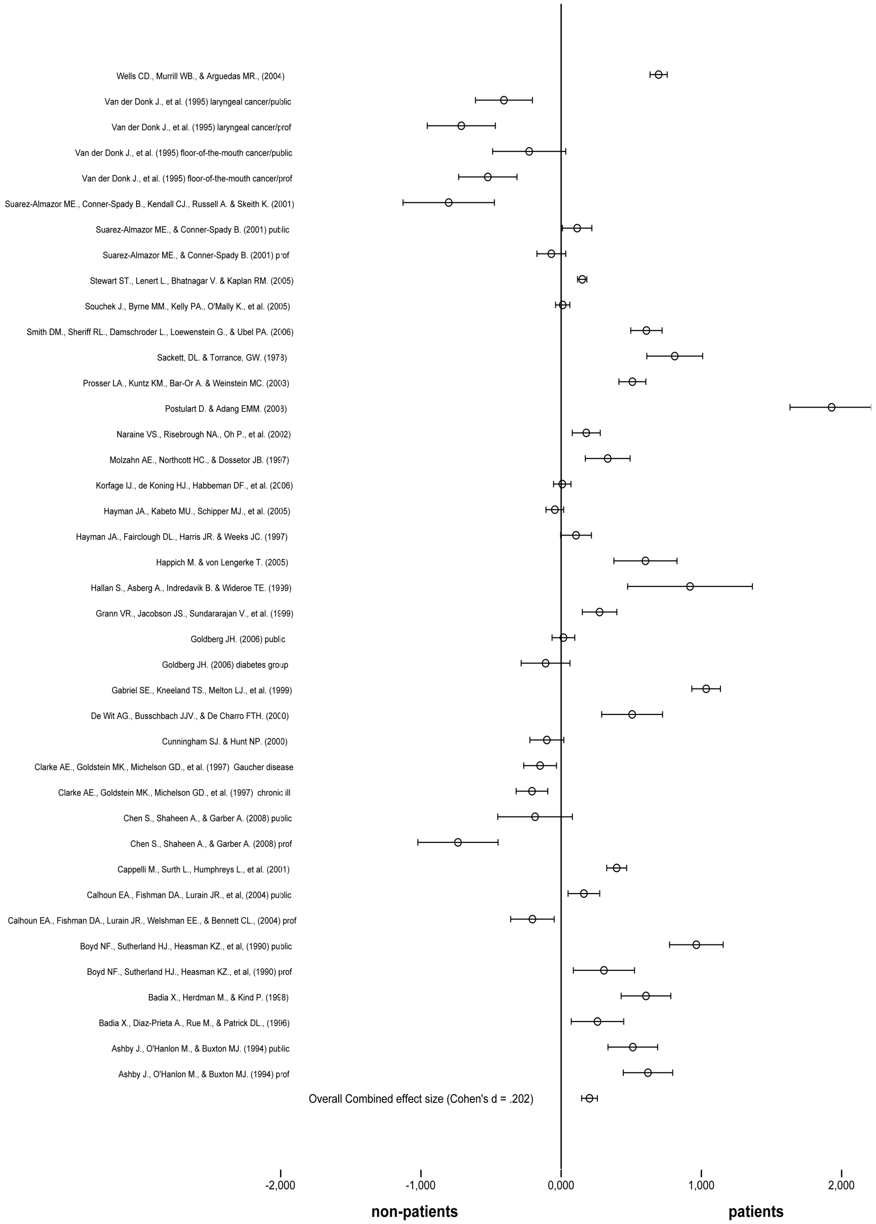


Table 2.1 Moderator variables; contrast for each of the moderators.

	N^d	Cohen's d	95% CI^a	Q^b	$Contrast^c$
Type of control group					2.54
Members of the general public	28	$0.27^{\ddagger}$	0.15 – 0.38	259.88**	
Professionals/proxies	12	0.00	−0.03 – 0.24	131.21**	
Own/hypothetical					4.63*
Scenario	34	$0.14^{\dagger}$	0.03 – 0.24	263.55**	
Own (patients valuing their own health)	6	$0.67^{\ddagger}$	0.20 – 1.15	54.71**	
Type of scenario					0.28
Description	28	$0.12^{\dagger}$	0.00 – 0.24	240.70**	
EQ-5D	6	0.18	0.00 – 0.36	17.56**	
Illness Label					4.81*
With Label	37	$0.18^{\ddagger}$	0.07 – 0.30	390.29**	
Without Label	3	$0.46^{\ddagger}$	0.24 – 0.68	1.87	
Non-patients actual/hypothetical					0.53
Actual (imagining health patient)	2	-0.20	−1.31 – 0.91	9.75**	
Hypothetical	38	$0.22^{\ddagger}$	0.10 – 0.33	388.22**	
Administration method					0.12
Computerized interview	7	0.27	−0.01 – 0.54	112.78**	
Interview without computer	30	$0.21^{\ddagger}$	0.07 – 0.35	271.62**	

^{a}CI = Confidence Interval, ^{b}Q = heterogeneity statistic, ^{c}Contrast between sets of studies, in Q, ^{d} the total N does sometimes not add up to 41 due to missing data,*Contrast for the moderator variable is significant $p < 0.05$, † Effect size of the subsample is significant $p < 0.05$, ‡ Effect size of the subsample is significant $p < 0.01$, **Subsample is heterogeneous $p < 0.01$

valuations of the two groups were similar (unstandardized $d = 0.02, p < 0.05$). The set of 24 SG estimators was heterogeneous [$Q(23) = 116.36, p < 0.01$]. There was no significant difference between response groups (unstandardized $d = 0.01, SD = 0.01, p < 0.05$); therefore, search for moderator factors was not performed. The set of 22 VAS estimators was heterogeneous [$Q(21) = 189.47, p < 0.01$]. A difference was seen between respondent groups (unstandardized $d = 0.04, SD = 0.02, p < 0.01$). Patients valued health states higher compared with nonpatients. A significant contrast was found between professionals/proxies and members of the general public [$Q(1) = 9.53, p < 0.01$]. Professionals/proxies ($N = 6$) did not value health states different from patients (unstandardized $d = -0.04, p < 0.05$), whereas members of the general public ($N = 16$) gave lower valuations compared with patients (unstandardized $d = 0.07, p < 0.01$).

2.4 Discussion

In this meta-analysis using 40 estimators from 30 studies, we found a small to moderate difference in valuations between patients and nonpatients. This finding contrasts with the findings of Dolders et al.[44] The exclusion of studies that used indirect instruments is unlikely to have caused this, as Dolders et al. did find a difference in valuations between respondent groups in studies using indirect instruments. A smaller number of included studies is not an explanation either, because we included 29 studies compared with only 11 by Dolders et al. From these 11 studies, seven studies were selected for the current meta-analyses; of the remaining four studies included in Dolders et al., three were based on indirect health state valuations, (the EQ-5D) and one study valued health states worse than death and reported that the majority of patients were unable to complete or understand the measurement tasks. Newly published studies ($N = 10$) included in our study may partly explain the difference. Finally, the difference might be explained by the inclusion of multiple effect sizes by Dolders et al.[44] which might have led to errors.

The results of the current study showed that states providing an illness label were rated more similar by patients and nonpatients than states not providing an illness label. Possibly, healthy subjects, like patients, will not use the whole utility continuum for labelled health states.[86] Another contrast was shown between studies in which patients valued their own health and studies in which patients valued a health state description. Valuations were more similar between groups when they both valued a description. This might be explained by a so-called loss aversion, patients giving higher valuations when they "own" a health state.[15] Initially, in three studies, the judges disagreed on this moderator variable, but after reading through the studies again, agreement was easily found. The disagreement was in two studies due to poor reporting and in one study due to a poor definition.

Only in the meta-analysis including studies for the VAS was an effect for the type of nonpatient group found. Valuations of professionals/proxies were more similar to those of patients than valuations of the general public, probably because of their experience with patients. In future meta-analyses, it may be worthwhile to start off by stratifying by both disease label and type of health state valued by patients (own health vs. scenarios), as these had moderating effects.

Despite the use of several moderator factors, all samples remained heterogeneous, except for three studies without illness label. Different explanations may be given for this heterogeneity. First, a great diversity was seen between the type and severity of the health states. As shown by Insinga and Fryback,[23] the difference between valuations given by different respondent groups may depend on the severity of the health state. Second, patients as well as members of the general public differ in the extent of their experience with different health states, which creates heterogeneous groups.[17] Unfortunately, we were not able to control for the differences in experience and the choice of the particular health states.

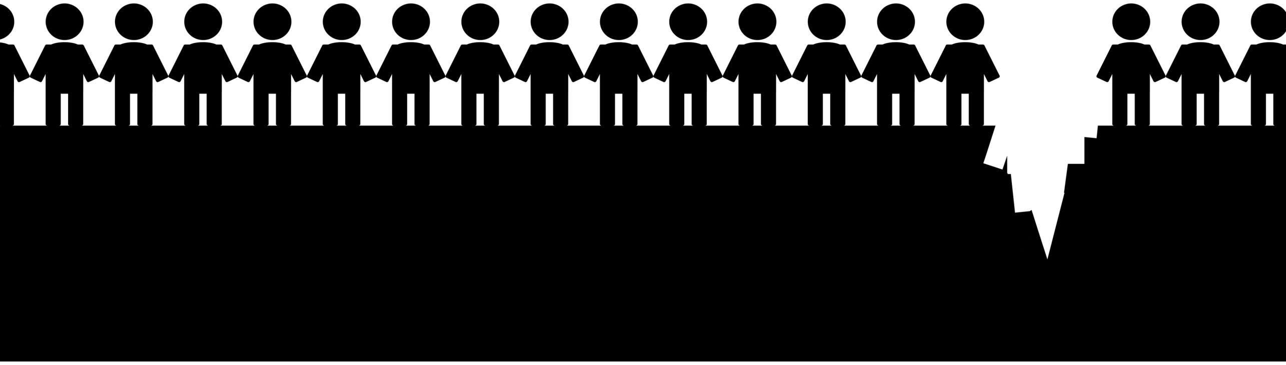
In this study, multiple significance tests were carried out, which might have led to multiplicity. Using Bonferroni correction, the main results of the elicitation subsamples remained the same. Correcting the moderator variables in the overall metaanalysis and in the meta-analysis of studies by elicitation method, nonsignificant contrasts for all samples were found. However, it has been argued that tests performed to investigate heterogeneity should not be adjusted for multiple testing.[87] Given our results, future studies should take the impact of respondent group into account. Which respondent group should assign health state valuations depends on the research question of the study. For cost-utility analysis, the implications of our findings can be best illustrated using the unstandardized differences. Mean unstandardized difference in studies using the TTO or the VAS was 0.05 and 0.04 with a 95% confidence interval of 0.01-0.08 for the TTO and 0.01-0.07 for the VAS. The influence of such a difference on a cost-utility ratio depends on other characteristics included in the analysis, for example the period for which the effect of treatment lasts. In studies using the SG, no effect of respondent group was seen, probably due to ceiling effects caused by risk aversion.[15] Given the small sample sizes and different findings between the meta-analyses, we feel that we cannot claim implications for the findings of the moderator analyses. These results should be corroborated in future research.

We would like to thank those authors who provided additional information for their cooperation.

3

Valuing Health: Does Enriching a Scenario Lead to Higher Utilities?

Peeters, Y. & Stiggelbout, A.M. (2009). Valuing health: does enriching a scenario lead to higher utilities? Medical Decision Making, 29, 334-342.



Abstract ***Objectives:*** *Patients have been found to value their own experienced health state higher than an investigator constructed scenario of that health state. The aim of this study was to investigate if patients value their own experienced health state higher than a standard EQ-5D scenario of their health state and if "enriching" this scenario by adding individualized attributes reduces the differences between experienced health and the scenario.* ***Methods:*** *Face-to-face interviews were held with* 129 *patients with rheumatoid arthritis. Patients were asked to value in a time tradeoff their own experienced health; 6 standard EQ-5D scenarios, of which the 5th (untold to them) represented their own health state; and a standard EQ-5D scenario of their health state (identified as such) enriched with individual attributes.* ***Results:*** *The own experienced health state was not valued differently from the own standard EQ-5D state and was lower compared to the own enriched EQ-5D state of that same health state. An interaction effect was found for health status. Patients with better health did not report different values for their own experienced health compared with their own standard EQ-5D description; their own experienced state was rated lower than their own enriched EQ-5D description. Patients with poor health valued all 3 health states similarly. Surprisingly, utilities for scenarios enriched with exclusively negative individual attributes were not lower than those for the own standard EQ-5D description.* ***Conclusion:*** *The hypothesis that disparities in valuation can be attributed to EQ-5D description being too sparse was not confirmed.*

3.1 Introduction

Utilities of health states are important in health decisions. Health state utilities are used to compare investments in cost of a therapy with the benefits in health. Utilities can be elicited in members of the general public but also in patients. Which group should be used is still a matter of discussion.[16,44] Many studies[16,88] but not all[81] have found valuations of patients to be different from valuations of members of the general public.

Patients are often asked to value their own experienced health state, whereas members of the general public are asked to value descriptions of these health states. Jansen and others[20] found that patients' ratings of their own experienced health state were higher than their valuation of a description of that same health state. The authors explained this difference in rating by hypothesizing that the description of the health state may not have matched the own experienced health state despite an evidence-based development process. Similar results were revealed in a meta-analysis of utilities assigned to prostate cancer.[46] Patients with prostate cancer rated a description of their health lower than their experienced health state.

Valuing a description of a health state instead of valuing an experienced health state might cause differences in the interpretation and integration of the information. These differences in interpretation and integration could result in different utilities.[15] In particular, patients interpret information in light of their experience, whereas healthy participants are limited to the information that is provided in the health state description.

Moreover, descriptions of health states are developed in several ways. Jansen and others[20] developed health state descriptions based on the literature and experiences of physicians and patients. However, others have developed descriptions on the basis of health state classification systems, such as the Health Utilities Index (HUI)[21] and the EuroQol EQ-5D.[89] Dissimilarities in the construction of health state descriptions might lead to different interpretations and valuations as well. In addition, health state descriptions are often framed in negative terms. This leads to a focus on the negative impact of the health state, which might cause healthy participants to overestimate the negative impact of a disease.

Insinga and Fryback[23] asked members of the general public to value a selection of all possible EQ-5D health state descriptions as well as their own experienced health. By chance, several participants' experienced health matched one of the EQ-5D descriptions they had valued. It turned out that ratings of the own experienced health differed from the ratings of the matching EQ-5D description. Specifically, participants with mild health problems valued their own experienced health lower than the corresponding EQ-5D health state description, whereas patients with moderate health problems estimated their own experienced health higher than the corresponding EQ-5D health state description. The authors concluded that an EQ-5D profile of a health state does not resemble the own experienced health state because it is too sparse and lacks positive aspects.[23] Possibly, EQ-5D descriptions should be enriched to create more resemblance between self-ratings and self-identified EQ-5D ratings.

In cost-utility research, enriched EQ-5D descriptions have already been used to explore preferences for different medication types. Medication-related attributes added to the EQ-5D description induced differences in preferences between treatments.[90] Smith and others[49] suggested that formerly treated patients should rate their past health state more similar to patients than to members of the general public, assuming that differing valuations result from descriptions being sparse and lacking scope. In contrast to their expectations, ratings of formerly treated patients were more similar to the ratings of members of the general public than to the ratings of patients currently undergoing treatment.[49] This finding indicates that providing more detailed information about a health state might still not eliminate patient-public differences.[49] Nevertheless, information that makes the health state description more personal might improve health state descriptions. For instance, Llewellyn-Thomas and others[91] found that with objective health outcomes, individual health state descriptions were better explained than standardized health state descriptions.

The aim of this study was to investigate if patients value their own experienced health state higher than their own standard EQ-5D scenario and if "enriching" this scenario by adding individualized attributes leads to smaller differences between the valuations of the own experienced health and the scenario. To this purpose, patients had to value their own health state in 3 different ways. They valued their own experienced health, a standard EQ-5D description of this health state, and an enriched EQ-5D description of this health state. Based on the findings of Smith and others[49] and Llewelyn-Thomas and others[91] we chose to enrich the own standard EQ-5D description with individual patient attributes instead of giving more detailed but standard information. Considering the results of Insinga and Fryback[23] we expected the valuation of the own standard EQ-5D description in relation to the other valuations to depend on the current health of the patient.

3.2 Methods

3.2.1 Participants and procedures

The sample consisted of patients with rheumatoid arthritis (RA) aged 18 to 76 years old who had visited their treating rheumatologist in the past 6 months. From the database of the Leiden University Medical Center, 300 patients who visited their rheumatologist in the last year were randomly selected. In the selection method, we oversampled men to get an equal male/female distribution because RA is more prevalent in women.

Medical records of the selected patients were assessed for comorbid conditions and true diagnosis of RA. From the 300 selected patients, 50 patients had not been diagnosed with RA, and 7 had comorbid conditions. The remaining 243 eligible patients received information about the survey by mail, including an informed consent form. If patients did not return the informed consent form within 3 weeks, they were called as a reminder. Data were collected using self-report questionnaires and a semistructured interview. The medical ethics committee of the Leiden University Medical Center approved the study protocol.

3.2.2 The interview

Face-to-face interviews were performed by 3 trained interviewers following a strict interview protocol. The interviews took place at the patients' preferred location: at home, in the hospital, or at work. Patients who were interviewed in the hospital came to the hospital; they were not hospitalized at the time of the interview. The interview started with the valuation of each participant's own experienced health of the previous week. This was followed by the EQ-5D questionnaire, a 5-item health-related quality-of-life questionnaire with the dimensions mobility, selfcare, usual activities, pain/discomfort, and anxiety/depression.[89] Patients answered this questionnaire on a 3-point scale: no problems, some problems, and no function at all or, in the case of pain, extreme pain. After this EQ-5D questionnaire, 2 filler questionnaires followed -that is, the Patient Satisfaction Questionnaire[92] and the Rosenberg Self-Esteem Scale[93]- to distract patients' attention from the answers they gave on the EQ-5D questionnaire. In the next part, participants were asked to value[46] standard EQ-5D states. Five of these EQ-5D states were retrieved from previous research with patients with RA, covering the full utility range from 0 to 1 according to the UK tariff.[11] A description of these health states can be found in appendix B. Unknown to the patients, the sixth health state was their own standard EQ-5D state of the previous week, as indicated in the EQ-5D questionnaire. The computer retrieved the answers of the patient earlier in the interview and created the own standard EQ-5D state for this patient. All standard EQ-5D states were randomly presented except for the patient's own standard EQ-5D state, which was always presented as the 5th state. The description of the patient's own standard EQ-5D state was similar to that of the other standard EQ-5D health states. Patients were not informed that it was their own standard EQ-5D health state. If 1 of the 5 preselected EQ-5D states happened to be the same as the own standard EQ-5D state, this state was replaced automatically with the EQ-5D state that should have been presented in the 6th place.

After valuing their own experienced health and the 6 EQ-5D descriptions, patients answered an open-ended question asking them to indicate attributes important to the own experienced health state. The interviewer entered these attributes in the computer. It was impossible to add a full description of each attribute; consequently, a key word was used. The interviewer and the patient together created suitable key words for each of the individual attributes. Only key words on which the patient agreed were used. These individual attributes were then combined with the patient's own standard EQ-5D state of the previous week to create an own enriched EQ-5D state. On the computer screen, the description of the own enriched EQ-5D state was shown with the individual attributes represented beneath the 5 standard attributes. It was made clear to the patients that the order in which the attributes were presented was arbitrary and that it was up to the patients how important the attributes were to them. Furthermore, patients were told that the description as stated on the computer fit their own health state.

If this were not clear, the interviewer explained how this description was created and made sure that the patient understood that it was his or her own health state. After the valuation of this own enriched EQ-5D state, patients indicated their level of functioning on the individual attributes that they had named before as important to their quality of life of the previous week. To rate this functioning, we used the same scale as was used in the EQ-5D questionnaire. Patients stated if they had no problems, some problems, or were not able to perform an individual attribute. At the end of the interview, all patients were asked whether they had recognized among the 6 EQ-5D states their own standard EQ-5D state that described their own health state. A general overview of the different elements of the interview is shown in (Figure 3.1).

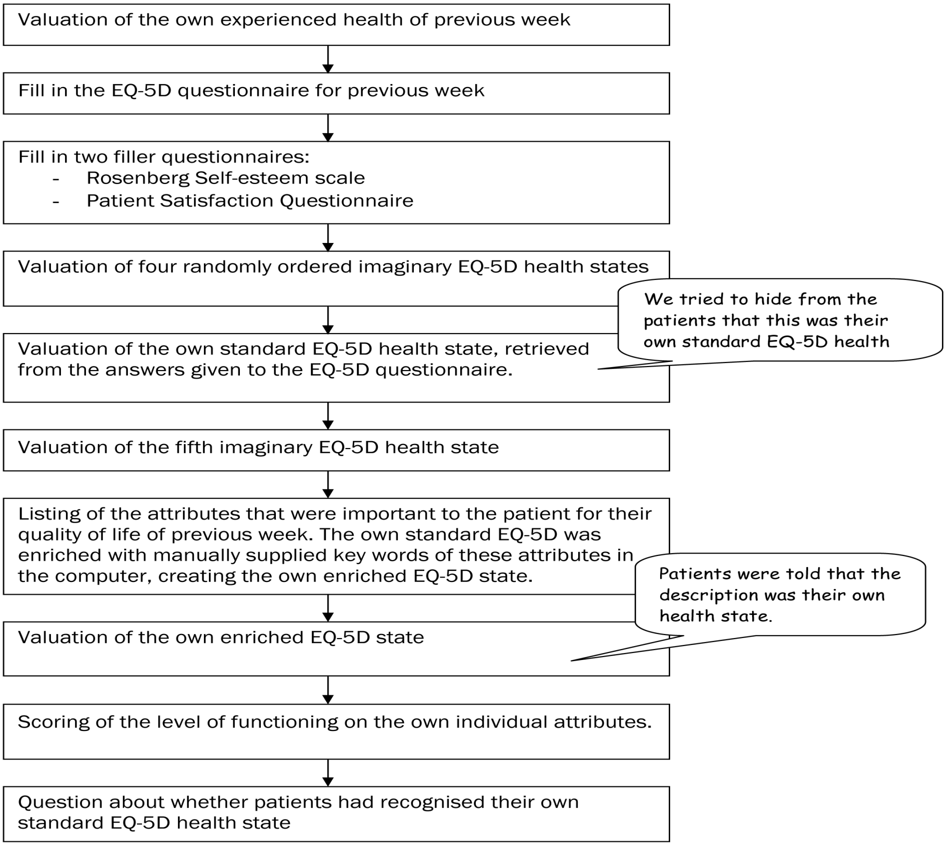


Figure 3.1 The interview process.

All health states were valued using a time tradeoff (TTO). Patients rated how many years (x) of their remaining life expectancy (y), derived from Dutch life expectancy tables [17], they were willing to trade to obtain perfect health. Utility was calculated as $\frac{y-x}{y}$. The computer program Ci3^{94} was used to elicit the utilities based on a pingpong search procedure. On the computer screen, a short description of perfect health and the health state to be valued were presented. Perfect health was described as full well-being, physically, psychologically, and regarding social activities. While completing the TTOs, patients were asked to think aloud.

After the interview, patients were asked to complete the Health Assessment Questionnaire (HAQ)^{95} at home and to return it by mail. The HAQ is a 24-item disease-specific health questionnaire. Patients reported the number of problems they perceived in performing several daily activities and whether they had to use devices for these activities. The total HAQ score was used in this study as an indicator of the patients' health status, with higher scores indicating worse functioning.

3.2.3 Data analysis

Prior to the main analyses, all variables were examined for uni- and multivariate outliers, linearity, and normality. Missing data were excluded listwise. Differences between valuations were analyzed using within-subjects analysis of variance (ANOVA). Using the Bonferroni pairwise comparison, post hoc contrasts were performed to investigate the valuations of the own experienced health, own standard EQ-5D, and own enriched EQ-5D pairwise. On the basis of statements made during the think-aloud procedure and the open-ended question, patients were divided into 2 groups depending on whether they had recognized their own standard EQ-5D state. To investigate if recognizing the own standard EQ-5D influenced the valuation of this health state, we performed a $t - test$.

Two interviewers judged independently whether the individual attributes named in the open question used to enrich the own standard EQ-5D states were positive, negative, or neutral. The agreement between the ratings of the interviewers was good (Cohen's $\kappa = 0.90$). Divergent evaluations were compared, and agreement was found through listening to the taped interview and by discussion. We expected the valuation of patients' own enriched EQ-5D to be higher when this description was made more positive compared to their own standard EQ-5D.

Inversely, we expected the valuation of the own enriched EQ-5D to be lower compared to the own standard EQ-5D when adding the individual attributes made this description more negative. Examples of negative attributes were pain, fatigue, and mobility; examples of positive attributes were grandchildren, good emotional functioning, and leisure activities. Naturally, the positive effect of the positive attributes would only hold if patients stated to have no problems on this attribute. Similarly, the negative effect would only hold if patients stated to have some problems or were not able to perform the attribute. To determine this valence of the attributes, we analyzed each attribute for the number of problems that patients stated to have with that particular attribute: no problems, some problems, or unable to perform. Only positively evaluated attributes with no problems were judged to add positive information, and negatively evaluated attributes with some problems or unable to perform were judged as negative added information. For example, when patients named their partner as an additional attribute, this was expected to increase the valuation of the enriched EQ-5D health state only if the patient stated that he or she had no problems with his or her partner. If the patient reported having some or severe problems with his or her partner, we could not be sure if the enriched EQ-5D would become more positive by adding the partner as an additional attribute. The effect of added attributes on the valuation of the health state was assessed with descriptive statistics and a paired sample $t-$test. Finally, ANOVA was used to assess if current health influenced the relative valuations of the 3 health states, with current health based on the dichotomized total HAQ score.

3.3 Results

A total of 132 patients of 243 patients approved the interview, a response rate of 54%. Of these responders, 1 patient with emotional problems and 2 patients who were not able to speak and understand Dutch were excluded. No differences in age and time since diagnosis between responders and nonresponders were found. Data of 2 participants created multivariate outliers and were excluded from further analyses; Mahalanobis distance, $F(3) = 31.07$ and $F(3) = 18.05$. All variables met the assumptions for linearity and normality, except for the variables “own experienced health,” “own standard EQ-5D,” and “own enriched EQ-5D.” Because we found similar results with nonparametric tests as with parametric tests, we decided to present the results of the parametric tests. These tests give more information and made it possible to test an interaction effect.

The interviews took place at the patients’ preferred location: at the hospital ($N = 82$), at the respondent’s home ($N = 44$), or at work ($N = 1$). Patients were not hospitalized at the time of the interview. The interview took 1.5 to 2 hours. Patients interviewed at home had on average more health problems based on the HAQ total score than patients interviewed in the hospital. Table 3.1 presents the demographic information of the 127 respondents who were included.

3.3.1 Valuations of own experienced health state

Table 3.2 shows the means and standard deviations of the 3 health state valuations. We found small differences among the ratings of the 3 health states: own experienced health state, the own standard EQ-5D, and the own enriched EQ-5D, $F(2, 242) = 3.83, p = 0.03$. Post hoc analyses showed that this effect resulted principally from the patient’s own experienced health state scoring somewhat lower

Table 3.1 Patient Characteristics ($N = 127$)

	Mean	SD	N	(%)
Age	58	11		
Gender				
Female			61	(48%)
Education[a]				
Nine years or less			38	(30%)
Between 10 and 12 years			62	(49%)
13 years or more			24	(19%)
Children				
Yes			61	(48%)
Marital status				
Married			38	(30%)
Divorced/Widow			62	(49%)
Single			24	(19%)

[a] Numbers do not add up to 127 due to missing data.

than the patient's own enriched EQ-5D state ($p = 0.03$). No significant differences were found between the ratings of the patient's own experienced health state and the patient's own standard EQ-5D state description or between the standard and the own enriched EQ-5D state descriptions.

Table 3.2 Means and SD of the valuations the different health states

($N = 122$)	Mean	SD
Own experienced health state	0.79	0.23
Own standard EQ-5D state[a]	0.81	0.25
Own enriched EQ-5D state	0.83	0.22

[a] No differences in the valuations of the own standard EQ-5D state were found between patients who had versus who had not recognized their own standard EQ-5D state, $t(123) = 0.651, p = 0.51$.

Table 3.3 Means and SD of the valuation of the different health state descriptions according to the severity of the patients' current health.

	Patients $HAQ \geq 0.94^b$ ($N = 59$)		Patients $HAQ < 0.94^b$ ($N = 62$)	
	Mean	SD	Mean	SD
Own experienced health state	0.75	0.26	0.83	0.20
Own standard EQ-5D state	0.74	0.29	0.88	0.18
Own enriched EQ-5D state	0.76	0.27	0.90	0.14

[b] Higher HAQ scores indicate worse functioning.

3.3.2 Differences in ratings between patients based on the severity of their current health state

To investigate the effect of the patients' current health, we performed a median split based on the HAQ total score ($0.00 - 0.94$ vs. $0.95 - 3.00$). There were no differences in gender or age between the 2 groups. When the dichotomous HAQ score was added to the ANOVA for the different valuations, a trend was found for an interaction, $F(2, 238) = 2.5, p = 0.09$. Table 3.3 shows the means and standard deviations of the 3 health state valuations for the 2 groups.

The effect seen in the total group turned out to occur only in patients in better health. There was a small difference between the ratings of the 3 health states, $F(3, 183) = 7.94, p < 0.01$. Bonferroni pairwise comparisons showed that this difference resulted principally from the lower valuation of the own experienced health state than the valuation of the own enriched EQ-5D state description ($p = 0.01$). Patients in poorer health rated the 3 health state descriptions as similar, $F(2, 116) = 0.55, p = 0.55$.

3.3.3 Own enriched EQ-5D state description

To the open-ended question, most patients named both positive as well as negative attributes ($N = 96; 76\%$). Fourteen (11%) patients named exclusively positive attributes, and 16(13%) named exclusively negative attributes. Patients who named exclusively positive attributes indeed gave slightly higher valuations to their own enriched EQ-5D state ($mean = 0.92, SD = 0.13$) compared to their own standard EQ-5D state ($mean = 0.90, SD = 0.15$); however, this difference was not statistically significant, $t(13) = 0.03, p = 0.11$. Contrary to our expectations, patients who named exclusively negative attributes similarly did not rate their own enriched EQ-5D state lower than the standard EQ-5D state description ($mean = 0.79, SD = 0.23$ vs. $mean = 0.76, SD = 0.26$).

3.4 Discussion

In this study, we investigated if patients valued their own experienced health state higher than an EQ-5D scenario describing their health state and if enriching this scenario by adding individualized attributes led to smaller differences between the valuations of the own experienced health and the scenario. Contrary to our hypotheses, the own experienced health state was not valued differently from the own standard EQ-5D state and was found to be lower when compared to the own enriched EQ-5D state of that same health state. We found an indication of an interaction between one's current health and valuations. Patients in relatively good health rated the own standard EQ-5D state description somewhat higher than their own experienced health but not statistically significantly, whereas we did not see differences for patients with poorer health. These findings are in line with the results of Insinga and Fryback.[23] These authors suggested that when individuals rate their own experienced health, they might consider minor decrements in 1 or more of the 5 EQ-5D dimensions that fall between "no problems" and "some problems" or that patients consider health decrements within attributes not specified by the EQ-5D dimensions. Likewise, in our study, patients with better health might have thought about minor problems when they valued their own experienced health.

By enriching the own standard EQ-5D state, we expected to make the description and, as a result, also the valuation more similar to the own experienced health state. However, the own standard EQ-5D states were also valued higher than the own experienced health by patients with better health. We can only speculate about this unexpected finding. Perhaps framing of the question about the own experienced health caused the difference between the valuations. Specifically, in the introduction of the own experienced health state at the beginning of the interview, patients were asked to think about their previous week, particularly about their physical, social, and emotional health in that week. In this introduction, the last week was emphasized, and as a result, patients with better health might have been provoked to think about minor problems of the previous week. In the open-ended question, following the 6 standard EQ-5D valuations, patients were asked to name the most important attributes relating to their own health state of the previous week. First, thinking about the ”most important aspects” might have overshadowed the minor problems patients thought about earlier. Second, patients might have become more aware of the positive aspects of their lives. The latter question was asked after patients had valued 5 health states that were for patients in relatively good health who were most often worse than their own health. Lacey and others[96] showed that health states are valued differently when context information is added. Their participants rated a severe health state more severe in a context of other less severe health states and rated a mild health state more mildly in a context of other more severe health states.

In our study, patients valued their own standard EQ-5D state and the own enriched EQ-5D after they had valued 4 other health states. We did so to create a situation where patients would not easily recognize their own standard EQ-5D state. To reduce a context effect of the other health states on the valuation of the EQ-5D states, however, we varied the sequence in which the health states were presented randomly, whereas in the study of Lacey and others,[43] patients viewed the health states in order from most severe to least severe, and patients were asked to view all health states first before they rated them. We believe that a context effect, if any were present in our study, would not have strongly influenced patients’ valuations because preselected EQ-5D states showed no order effect. That is, a health state presented first was not valued differently from a health state presented last [a] , even though the valuation of the enriched EQ-5D health state might have been influenced by the earlier 5 states for patients with better health. They might have recognized that their health state was not as bad as the other 5 health states, which could have led to a higher valuation of their own enriched EQ-5D health state. The finding that this effect was only seen in patients in better health may point toward such a contrast effect. For this group of patients, the states were more likely to be worse than the own health state. The fact that their own experienced health was always presented at the beginning of the interview might have had a negative effect on the valuation of this health state. We felt this to be inevitable because we wished to avoid having EQ-5D information to influence patients’ valuations of their own experienced health state.

Another finding in this study was that even when exclusively negative attributes were added, the own enriched EQ-5D state was not valued lower than the own

---

[a] The mean (0.69[0.28]) of the first presented health state did not differ from the mean (0.69[0.28]) presented last. Because of the randomization, all health states appeared an approximately equal number of times in each place. This made it possible to calculate the mean valuation of all health states presented in the first place and to compare this with the mean valuation of a health state presented in the last place.

standard EQ-5D. Because this result was based on valuations of only 16 patients, this finding has to be viewed with caution. However, if there is some ground for this finding, it might have been caused by a so-called status quo bias; people tend to value health states higher when they ”own” that health state.[15] In economic decision making, studies have revealed that people value goods more highly when they own these goods.[28, 29] People thus seem to prefer what they know, which may result in a preference for own health above an unknown health state. In our study, when patients were valuing the own enriched EQ-5D health state, they were told that it was their own health, in contrast to when they were valuing the own standard EQ-5D description of their own health. Although we found significant differences, they are small compared to the minimal importance difference (MID) for the EQ-5D. Walters and Brazier[97] revealed a mean MID of $0.074(-0.011$ to $0.140)$ for the EQ-5D with secondary analyses on 11 studies. However, the aim of this study was to understand why health states are valued differently. In future studies, it would be interesting to investigate the consequences of such differences for cost utility analyses.

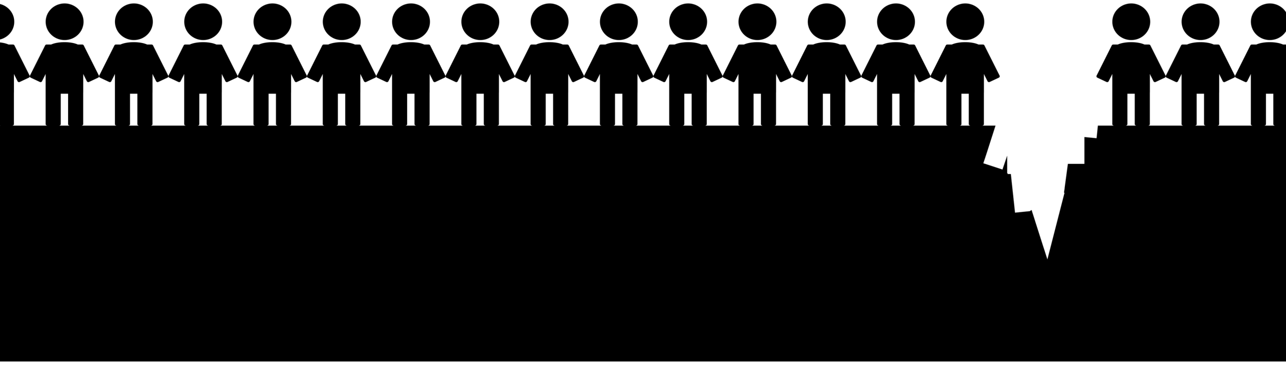
3.5 Conclusion

Only limited support was found for the contention that the EQ-5D state description might be too sparse. It remains uncertain if including personal information with a health state description will make hypothetical health states valued more similarly to experienced health state ratings.

4

Focusing illusion, adaptation and EQ-5D Health State Descriptions: The difference between patients and the general public.

Peeters Y., Vliet Vlieland T.P.M., & Stiggelbout A.M. Focusing illusion, adaptation and EQ-5D Health State descriptions: The difference between patients and the general public. Health Expectations 2011 Mar 3, Epub ahead of print.



Abstract ***Objectives:****Patients tend to assign higher utilities to health states compared to the general public. Several explanations have been given for this difference including focusing illusion -, caused in part by the We investigated whether patients and the public differ in which dimensions they find important. Furthermore, we compared whether the dimensions named by patients and the public obtained higher rankings of importance compared to the predefined EQ-5D dimensions. Within each nominated dimension we investigated whether the public used a more negative frame compared to patients. In addition, adaptation was investigated by comparing patients with high levels of adaptation and patients with low levels of adaptation.* ***Method:*** *Data were collected using semistructured interviews among 124 patients with rheumatoid arthritis and 64 members of the public. Participants indicated which aspects are important to them when they think about their life having rheumatoid arthritis and rated the importance of these aspects and of the EQ-5D dimensions.* ***Results:*** *In contrast to patients, the public named more often aspects related to sports & mobility, leisure activities, and work, and framed these aspects negatively. Compared to self-rated dimensions, the public ranked the EQ-5D dimensions as more important whereas patients found both groups of aspects equally important. Patients who showed higher levels of adaptation did not differ significantly from patients with lower levels.* ***Conclusion:*** *The public is focussed on life domains that are negatively influenced by the described health state whereas patients are focussed on both the positive and negative aspects of their lives.*

4.1 Introduction

In cost-utility analysis, utilities are used to estimate how much better the quality of life is in one health situation or 'state' compared with another. Utilities can be elicited from the public or from patients. For analysis from a societal perspective, it is recommended to use utilities assessed from a representative sample of the general public.[98] In contrast, utilities of patients who have experience with a health state might be more appropriate in clinical decision making and in certain policy decisions.[99] Whose utilities should be used is only relevant if patients and public differ. Literature dealing with this difference is somewhat controversial3 and generally supports the supposition that patients assign higher utilities compared to members of the public.[16, 46, 100]

Several explanations have been given for the difference between patients and the public. Patients typically assign utilities in light of their experiences, whereas members of the public are limited to a description of the health state. Even among patients, it has been shown that utilities for patients' own health are higher than utilities they assign to a description of their own health.[20, 46] When patients give utilities for their own health, a broad range of information can be used, whereas utilities based on health state descriptions are supposedly limited to the information provided.

The EQ-5D is widely used to provide health state descriptions. The EQ-5D consists of five dimensions:, mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension is described according to one of three levels of severity: no problems (1), some problems (2) and extreme problems (3). In total the EQ-5D can thus create 243(3^5)theoretically possible health state descriptions.[89] EQ-5D health state descriptions thus consist of five sentences stating the level of problems on the dimensions. Several publications have pointed out that the sparseness of these EQ-5D health state descriptions[23, 90, 101–104] limits their ability to comprehensively describe health states. The sparseness of EQ-5D descriptions is a result of the original purpose of the EQ-5D. The EQ-5D was developed as a non-comprehensive measurement alongside other more detailed measures of Health Related Quality of Life (HRQoL).[105] The five EQ-5D dimensions were selected from a review of existing instruments and tested against the results of a survey on the concepts of health of lay persons.[22] Given the current use of the EQ-5D as a full health state description it has been suggested that the comprehensiveness of the EQ-5D health states should be improved by adding new health state dimensions.[23] Studies investigating the effects of extra dimensions have based the selection of extra dimensions on expert groups and the general public's concepts of health.[102,104] However, by selecting EQ-5D dimensions from lay persons' health concepts, the sparseness of the EQ-5D health state descriptions might inadvertently prompt a focusing illusion in this group.

Focusing illusion is suggested as an important explanation for the difference between health state utilities assigned by patients and the public. When members of the general public are asked to imagine life in a certain health state they will concentrate on the differences between their current health state and the health state to be valued.[26] Life domains influenced by the health state receive disproportional attention, whereas domains that will not be affected are ignored. The public concentrates on the negative aspects of an illness while patients give utilities in light of their experiences in general.

Another explanation suggested for the differences in ratings between patients and the public is the adaptation of patients. When patients experience an illness they will adapt to it by changing their interests and goals and by lowering their expectations for specific dimensions in the future.[18] In contrast, it seems almost impossible for the public to anticipate such adaptation.[38,39] This study investigated focusing illusion and adaptation as explanations for the differences in utilities between patients and the public. We investigated whether patients with rheumatoid arthritis (RA) and members of the general public who were asked to imagine having RA differed in which dimensions they rated as important in their lives. We hypothesize that compared to patients, members of the public are more focussed on dimensions influenced by the illness. Furthermore, we asked participants to name important aspects in their lives and examined whether the dimensions named by patients and the public were given higher rankings of importance compared to the predefined EQ-5D dimensions. Given that the EQ-5D dimensions were based on concepts of lay persons we hypothesize that the EQ-5D will accurately reflect the opinion of the public but not that of patients. Within each named dimension we investigated whether the public used a more negative frame compared to patients. The public may focus on those dimensions for which they would experience limitations after developing an illness, thus leading to negative framing.

Similarly, adaptation was investigated by comparing patients with high levels of adaptation and patients with low levels of adaptation. We investigated if patients with high levels of adaptation named different dimensions compared to patients with low levels of adaptation, if within each subgroup the dimensions named obtained higher importance rankings compared to the predefined EQ-5D dimensions and within each of the dimensions we investigated whether these two groups differed in their framing.

4.2 Methods

4.2.1 Patient subject recruitment

From the patient database of the Leiden University Medical Center, 300 patients aged 18 to 76 years old who had visited their rheumatologist in the past 6 months were identified. Men were oversampled to get an equal distribution of males and females. Medical records of the selected patients were reviewed to confirm the diagnosis of rheumatoid arthritis (RA) and to identify comorbid conditions. Of the 300 identified patients, 50 had not been diagnosed with RA, and 7 had severe comorbid conditions. The remaining 243 eligible potential patient subjects were mailed information about the interview and an informed consent form and asked to return the signed consent form if they were willing to participate. Those who did not return the form within 3 weeks were telephoned once and asked if they were interested, and whether they needed a copy of the survey and consent form mailed to them.

4.2.2 Recruitment of members of the public

Members of the public were recruited through advertisements in local newspapers distributed in Leiden and the surrounding mid-west region of the Netherlands. To meet inclusion criteria, participants had to be between 35 and 76 years old. Those who had RA or whose partner had RA were excluded. Of the 69 people who responded, two were excluded; one of the excluded participants had RA and the other one had a partner who had RA. The 67 individuals who were invited to participate were mailed information about the survey including an informed consent form.

4.2.3 Data collection

Data were collected using semistructured interviews and self administered questionnaires. Face-to-face interviews were performed by three trained interviewers. The interview took place at the patient subjects' preferred location: at home, at our department (an office area of the LUMC), or at work. Members of the general public were requested to visit our department. Patient participants who were interviewed at our hospital office were not hospitalized at the time of the interview. In this paper only the part of the interview used to gather the information for this study is described. A full description of the interview can be found elsewhere.[100]

The semistructured interview was slightly different for patients and members of the public. Both groups received the same questions but answered them from a different point of view. Patients were asked to answer questions about their own health during the last week. Members of the public read a health state description of RA (see Appendix C) and were asked to imagine that they had the RA state described, and to thus imagine their previous week with this health state. At any time point during the interview the RA health state description could be reread by the participant.

During the interview, patients and members of the public answered the EQ-5D questionnaire by filling out the level of problems on each of the EQ-5D dimensions for their own health or the health state description, respectively. Next, after some filler questions, patients were asked to name aspects important to their health during the last week, and members of the public were asked the same questions imagining that they had the RA state as described. An open-ended question was asked to elicit up to a maximum of five dimensions. Each of the dimensions named by the participant was summarized to one keyword after consultation with the participant. These keywords were then written down on separate sheets of paper. Subsequently, the five EQ-5D dimensions were also written down on separate sheets. All dimensions were ordered by importance to the participant's life with rheumatoid arthritis (actual or imagined). If aspects were ranked equally important they were put next to each other. The most important dimension was rated as 1, the next one 2, and so on.. If aspects were equally important they received the same importance rating.

After the interview, participants received a questionnaire which they completed at home and returned by mail. For patients this questionnaire included the Illness Cognition Questionnaire (ICQ) to assess adaptation. This questionnaire consists of three scales: hopelessness, benefit finding, and acceptance. Patients rated how much they agreed with 18 statements on a four point Likert scale ranging from 'not at all' to 'completely'. The medical ethics committee of the Leiden University Medical Center approved the study protocol.

4.2.4 Coding

The aspects named by the patients were initially consolidated into 10 dimensions based on often recurring themes in the interviews. Each interviewer independently coded the aspects based on these dimensions. Based on a comparison of the coding, a final coding system was developed. All aspects were recoded by each of the two interviewers, differences were compared, and discrepancies were resolved through discussion and reviewing the audio tapes of the interviews. A full description of the coding system can be found in Appendix D.

The dimensions named by members of the public were coded using the same coding system. Both interviewers coded all aspects. Agreement between the interviewers was substantial, with a Cohen's $\kappa = 0.79$.[106] Divergent evaluations were compared, and discrepancies were again resolved through discussion and reviewing the audio tapes of the taped interviews. All aspects named fitted in one of the dimensions of the coding system. The interviewers also judged independently whether the aspects were framed as positive, negative or neutral, based on the descriptions in the interview schemes. The agreement between these ratings was almost perfect for the dimensions named by patients (Cohen's $\kappa = 0.90$) and moderate for the public (Cohen's $\kappa = 0.51$).[106] Divergent evaluations were compared. By listening to the taped interviews agreement was found.

4.2.5 Analysis of data

1. Do patients and the public differ in which dimensions they find important, in their ranking, or in the framing of the aspects named?

To answer this research question, the number of patients was compared to the number of members of the public that named at least one aspect within a dimension. For each dimension the frequencies were compared by using a χ^2-test. Furthermore the overall importance of all EQ-5D dimensions was compared to the overall importance of all self-named aspects using the Wilcoxon signed rank test. A lower number indicated a higher importance. For this analysis we calculated for each participant the mean rank-order of all EQ-5D dimensions and that of the self-named dimensions. Next, for each of the EQ-5D and self-named dimensions the mean importance was calculated. The mean ranks of the coded dimensions based on the own named aspects were based only on the rank-order of participants who actually named an aspect in this dimension.

To investigate framing, we first compared the overall framing between patients and the public. For each participant we summed the positive, negative and neutral codings and divided these by the number of aspects. That is, a participant who named 3 aspects of which two positive and one neutral obtained the value $0.67(\frac{(1+1+0)}{(3)})$. The overall framing value between patients and the public was compared using the median test. Next we calculated for each dimension the number of participants who framed the aspect as positive, negative, or neutral. If participants named more than one aspect with unequal codings in the same dimension this participant was judged as “mixed”. For example, if a participant named “I often swim” and “I am not able to walk long distances” both aspects were coded into the dimension sports & mobility with one as positive and one negative. For each dimension the positive and negative frequencies of patients were compared to the frequencies of the public using a χ^2-test. Neutral and mixed were excluded.

2. Do patients with low acceptance of their illness differ from patients with high acceptance in which dimensions they find important, their ranking, or in the framing of the aspects named?

All analyses described above to compare patients and the public were also performed to compare patients with high acceptance of their illness and patients with low acceptance of their illness. Low and high acceptance of the illness was based on a median split of the acceptance scale of the ICQ. Bonferroni correction was used to correct for multiple comparisons where necessary.

4.3 Results

4.3.1 Participants

Of the 243 patients identified, 132 patients agreed to participate in the study (54%). No differences in age or time since diagnosis between responders and non-responders were found. Of the responders, one patient with emotional problems and two patients who were not able to speak and understand Dutch were excluded. Five patients who could not finish the interview due to time or cognitive and concentration problems were excluded from further analyses. In total 124 patients were included in the analyses. The interviews took place at the hospital ($N = 82$), at the patients’ home ($N = 41$), or at work ($N = 1$). In total 67 members of the public originally agreed to be interviewed. Two participants canceled the interviews, both due to tight time schedules. All interviews took place at the hospital ($N = 65$).

Table 4.1 and Table 4.2 present the demographic information of the participants. As can be seen from Table 4.2, the health state of patients was less severe than the health state description imagined by members of the public both based on the UK-tariff estimated from the ratings participants gave on the EQ-5D questionnaire. No differences were found on the aspects named and rankings given between patients with different educational levels. Past and/or present significant health

Table 4.1 Participant Characteristics

	Patients ($N = 124$)		Public ($N = 65$)	
	N	(%)	N	(%)
Gender				
Female	59	(48%)	r50	(77%)
Education[a]				
Nine years or less	33	(28%)	5	(8%)
Between 10 and 12 years	65	(52%)	29	(45%)
13 years or more	24	(19%)	31	(48%)
Children				
Yes	103	(83%)	49	(75%)
Marital status				
Married	109	(88%)	37	(57%)
Divorced/Widow	9	(7%)	13	(20%)
Single	6	(5%)	8	(12%)
Work Status[a]				
Fulltime	20	(16%)	7	(11%)
Part-time	22	(18%)	23	(35%)
Household/ volunteer/ unemployed	17	(14%)	14	(22%)
Unable to work	30	(24%)	6	(9%)
Early retirement	22	(26%)	15	(23%)
Retirement	2	(2%)	0	(0%)
Illness or disability				
Yes	No information available		40	(62%)

[a] Numbers do not add up to total sample size due to missing data.

Table 4.2 Participant characteristics; continuous

	Patients ($N = 124$)		Public ($N = 65$)	
	Mean (Range)	SD	Mean (Range)	SD
Age	58(29-75)	10.88	56(38-75)	10.86
Time since diagnose RA (years)	13(2-47)	9.26		
EQ-5D UK-tariff^a	0.63 ($-0.43 - 1.00$)	0.27	0.53($-0.20 - 0.69$)	0.27

^{a}EQ-5D UK-tariff of the health state to be valued

Table 4.3 Members of the public who currently have or previously had significant health problems:

Type of health problem:	N	(%)
Stroke	2	(3%)
Cardiac disease	4	(6%)
Cancer	7	(11%)
Migraine / severe headache	8	(12%)
High blood pressure	11	(17%)
Asthma/ bronchitis etc.	1	(2%)
Severe back problems/ herniated disc	3	(5%)
Parkinson	1	(2%)
Severe problems in joints	4	(6%)
Other	19	(29%)

problems experienced by participants from the general public are depicted in Table 4.3.

4.3.2 Patients vs. Public

1. Do patients and the public differ in which dimensions they find important, their ranking, or in the framing of the aspects named?

Table 4.4 shows the frequencies and percentages of participants who named at least one aspect in a dimension. Patients named significantly more often aspects in the dimension illnesses not RA-related ($\chi^2(1) = 15.25; p < 0.005$), and less often in sports and mobility ($\chi^2(1) = 9.67; p < 0.005$), and work ($\chi^2(1) = 8.01; p < 0.005$) compared to the public. The significance level was corrected for the number of tests,

Table 4.4 Number of aspects named in each dimension by patients and public

	Frequency and percentage dimension named:			
	Patient %	(N)	General public %	(N)
Physical inconvenience RA-related	77	(96)	66	(43)
Illness/inconvenience not RA-related	23	(29)	2	(1)$^\sharp$
Family	38	(47)	20	(13)*
Other social contacts	23	(28)	19	(12)
Emotions/Worries	40	(50)	31	(20)
Sports/Mobility	27	(33)	49	(32)$^\sharp$
Leisure activities	23	(28)	37	(24)*
Work	26	(32)	46	(30)$^\sharp$
Broad every day life	30	(37)	26	(17)
Other	20	(25)	19	(12)

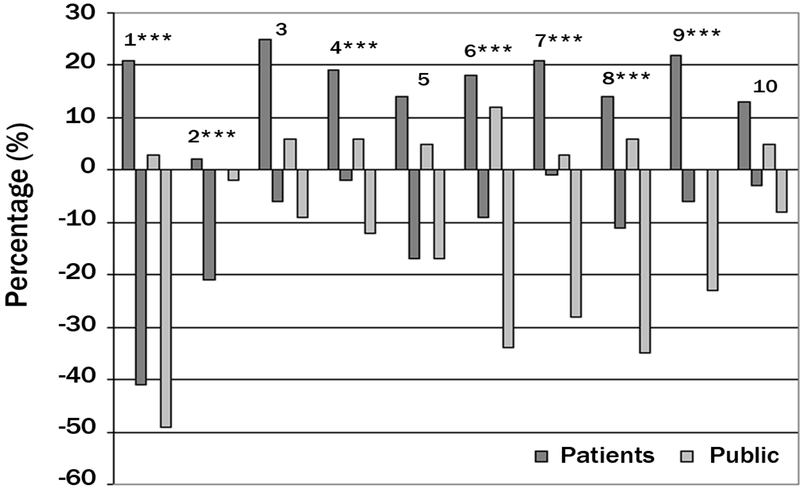
$^*p < 0.05,^\sharp p < 0.005$

resulting in a corrected level of $p < 0.005$.

The mean rank-order of EQ-5D dimensions was not different from the rank-order of self-named dimensions in patients ($mean = 3.47(1.24)$ vs. mean $= 3.26(1.19)$; $z = -1.64, p = 0.101$). By contrast, members of the public rated the EQ-5D dimensions as more important compared to the self-named dimensions (mean $= 3.30(0.99)$ vs. mean $= 4.01(1.42); z = -3.32, p = 0.001$). Table 4.5 shows the importance of each of the EQ-5D dimensions and self-named dimensions for patients and public; a lower number indicates greater importance. The mean reported is based on the rank-order of participants who named a dimension in this category. Consequently this mean is supported by only a sub-sample of participants (Table 4.4) making it impossible to statistically test the difference between the two respondent groups. Overall, the public framed the aspects more negatively compared to the patients (median framing value of the public = -0.75 vs. median of the patients = 0.00; $\chi^2(1) = 23.92, p < 0.005$).

Figure 4.1 shows the negatively and positively framed aspects for each of the dimensions separately; neutral aspects were excluded. The χ^2 tests shows that for almost all dimensions patients framed their aspects more positively, except for the dimensions family, emotions & worries, and other. Again we used a corrected significance level of $p < 0.005$.

Figure 4.1 Percentage dimensions named by patients and public divided in positive and negative



The negatively and positively framed aspects for each of the dimensions separately; neutral aspects were excluded.

$***p < 0.005$

1 Physical inconvenience RA-related
2 Illness/inconvenience not RA-related
3 Family
4 Other social contacts
5 Emotions/Worries
6 Sports/Mobility
7 Leisure activities
8 Work
9 Broad every day life
10 Other

Table 4.5 Importance of EQ-5D dimensions and self-named dimensions for patients and public

	Mean(SD) Rank-Order[a]	
	Patient	Public
EuroQol Mobility	2.92(1.70)	3.62(1.99)
EuroQol Self-Care	3.56(2.16)	3.29(2.16)
EuroQol Usual activities	3.12(1.88)	3.19(1.77)
EuroQol Pain/Discomfort	3.98(2.27)	3.21(1.67)
EuroQol Anxiety/Depression	3.81(2.42)	3.19(2.19)
Physical inconvenience RA-related	3.44(1.72)	3.13(1.68)
Illness/inconvenience not RA-related	4.02(2.21)	4.00(NA)
Family	2.51(1.52)	3.31(1.70)
Other social contacts	3.79(1.85)	3.82(2.14)
Emotions/Worries	3.23(1.87)	3.63(2.42)
Sports/Mobility	3.32(1.77)	4.32(2.20)
Leisure activities	3.52(1.69)	5.20(2.05)
Work	3.97(2.24)	4.67(2.22)
Broad every day life	3.00(1.90)	5.06(1.95)
Other	3.64(1.98)	3.33(1.87)

[a] lower number indicates greater importance

2. Do patients with low acceptance of their illness differ from patients with high acceptance in which dimensions they find important or in the framing of the aspects named?

No differences were found between these patient groups in the frequency of dimensions named. For patients with low acceptance the mean rank-order of EQ-5D dimensions (3.47 (1.27)) was similar to the mean rank-order of own named dimensions (3.35(1.12)), ($z = -0.51; p = 0.61$). However, patients with high acceptance tended to rank the self-named dimensions slightly more important compared to the EQ-5D dimensions ($mean = 3.22(1.27)$ vs. $mean = 3.48(1.23); z = -1.82, p = 0.07$). Aspects named by the patients with high acceptance were framed more positively (median = 0.27) than those named by the patients with low acceptance (median $= -0.20), (\chi^2(1) = 6.57, p < 0.01)$. However the framing within each dimension did not differ.

4.4 Discussion

Most studies show that patients assign higher health state utilities compared to members of the public.[16] We studied explanations for this difference. Support was found for the focusing illusion of the public. The findings underpin the hypothesis that the public is focused on life domains influenced by a health state: the public concentrated on not being able to perform work, leisure activities, and sports anymore. Alternatively, these findings could also be explained by adaptation of patients. Patients in this study were less concerned about not being able to work or play sports. A possible explanation is that they have learned to live without these activities and have found new life goals.

The overall framing of aspects important to patients was neutral, whereas a positive frame might have been expected. In the discussion about whose utilities should be used it is argued that patients overestimate their well-being, leading to a positive bias.[107] Patients might be unrealistically optimistic about their situation[108] and avoid complaining about their situation.[107] However, our results tentatively show that patients do not positively exaggerate their life domains. This is in line with Riis et al[107] who found no evidence that patients overestimate their mood. Possibly patients have a better picture about life in a certain health state, resulting in more accurate estimations. However, we have to take into account that the results are limited to patients with rheumatoid arthritis. It could be that these patients are less able to adapt to their illness. Rheumatoid arthritis is a disabling and often progressive disease, with varying symptoms and which is characterized by pain and deformity of the joints. There is evidence that pain is a symptom to which people only adapt to in a limited way.[109, 110]

To compare the effect of acceptance, patients with high acceptance were compared to patients with low acceptance. For all research questions a difference was seen between patients with high and low acceptance, but the effect was not strong. Patients who had better accepted their illness framed their aspects slightly more positively than patients who had a lower acceptance.

In contrast to the hypothesis that patients think more about their life in general we found that 77% of the patients still named one or more aspects of physical (in)convenience related to RA. This can be explained by the context of the study. Patients were invited to participate in a study investigating quality of life of patients with RA and answered several questions related to RA and their health before answering the open ended question about their most important aspects. It may be that the context of the study, the introduction, and the previous questions made information about RA easily accessible in the minds of the patients.[111,112] Along the same line, the answers given by members of the public may have been guided by the domains provided in the health state description.

Another explanation for the difference between patients and members of the public was the sparseness of the health state descriptions, such as in the EQ-5D. When compared to the self-named dimensions we found that according to the public the EQ-5D dimensions were ranked as more important. This suggests that the EQ-5D dimensions indeed give a good reflection of the most important dimensions to be included in a health state description according to the public. In contrast, patients found their self-named dimensions equally important as the EQ-5D dimensions. For patients the EQ-5D dimensions might not be complete, other dimensions are also important in creating comprehensive health state descriptions. A tentative suggestion can be made for adding information about family. As illustrated in Table 4, this dimension might be expected to be more important than the EQ-5D dimensions. Our findings suggest that patients find an EQ-5D health state description too sparse and information about family should be added. More research among patients is necessary to further investigate this finding.

Our findings were based on an open-ended question asking participants to name the most important aspects without referring to the predefined EQ-5D dimensions. Naturally, had we asked patients and the public which dimensions they found important in addition to the five EQ-5D dimensions, we might have obtained different results. We were afraid that if we referred to the EQ-5D dimensions, participants would have focussed on naming similar dimensions. Further, it was decided not to code the aspects into the predefined EQ-5D dimensions but to use a new coding system. EQ-5D dimensions can be interpreted in a very broad as well as in a narrow way. For instance, one can suggest that aspects related to family and friends are mentioned by the dimension “daily activities” whereas someone else might think that family and friends are not per se part of an activity. To determine if the self-named aspects correspond to the EQ-5D dimensions, more information would be necessary about how the participants interpreted the EQ-5D dimensions. For instance, broad everyday life might be expected to come under EQ-5D self-care or usual activities, but most often aspects named in this dimension were hard to distinguish. Often participants mentioned aspects such as “Generally I am able to do what other people would do on a typical day.” This can come under the EQ-5D dimension self-care but also under the dimension usual activities. When asked, participants did not make a clear distinction.

In our sample 62% of members of the public had experience with an illness. Experience with an illness may have influenced our findings.[17] Yet from a societal perspective it is suggested that a representative sample of the general public should be assessed for cost-utility analyses. This does not exclude people with current or previous health problems. When we compared the number of participants with significant health problems with the number of the Dutch population[113] we did not find major differences. Our sample seems to be a good representation of the Dutch population.

Patients and members of the public differed on some demographic characteristics. A major difference between the patients and the public was that in the sample of patients more participants were married. However, both members of the public and patients named family an equal number of times as an important domain. Another difference was seen in that more patients were unable to work, compared to members of the public. Post hoc analyses on work status showed that people who were able to work more often named aspects related to this dimension. This suggests that activities that have taken up most time in one's previous week come to mind more easily. However, participants unable to work, mostly patients, seem to have changed their focus. Participants able to work mentioned not being able to work as an important limitation, whereas participants who actually were not able to work did not mention work at all. Finally patients had fewer years of education than members of the public, but no effect was found on aspects named by education level.

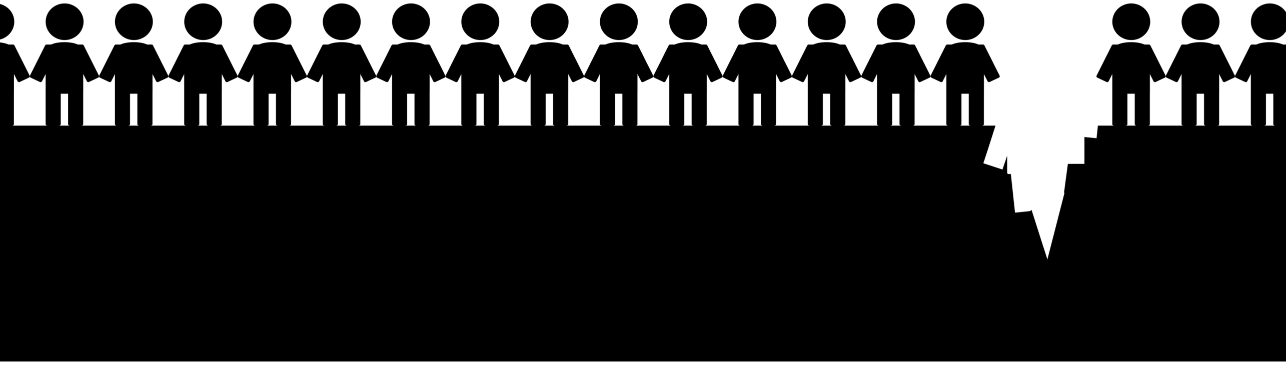
In this study patients were asked to think about their own health whereas members of the public were asked to imagine a health state description. Although this is in line with research in cost-utility analysis, it leads to differences in perception of health state severity and to variance between patients and public. In this study the average health state of patients was less severe than the health state description imagined by members of the public based on UK- EQ-5D tariff scores. However, given that the majority of patients were not able to work or had problems with sports/mobility but did not focus on these dimensions of life, we expect the influence of the severity of the health state to be minor.

In conclusion, regarding the EQ-5D health state descriptions, our findings indicate that patients find an EQ-5D health state description sparse. More research among patients is necessary. Regarding focusing illusion we found that the public is focused on life domains that are negatively influenced by the health state whereas patients are focused both on the negative and the positive aspects of their lives. Consequently, patients' picture of life in a health state might give a better reflection of reality. To what extent patients had accepted their illness did not have a strong influence on what aspects patients found important.

5

Health state utilities from patients, partners and the public: In search of an explanation for the gap between the imagined and the real

Stiggelbout A.M., Peeters Y., and de Jong Z. Health state utilities from patients, partners and the public: In search of an explanation for the gap between the imagined and the real. Submitted for publication



Abstract ***Objectives:****Many studies have found health state valuations of patients to be higher than valuations of members of the general public, and valuations of professionals to be intermediate. This difference has generally been ascribed to experience with the state, whereby vicarious experience of professionals leads to the intermediate values. No studies have compared utilities of patients with those of their partners, who have a similar but more "insider" vicarious experience. The aim of this study was to compare values of patients, their partners, and members of the public. Since the studies that have found higher values for patients have generally asked patients to value their own experienced health state rather than a scenario, we also wished to assess whether the difference maybe due to a scenario-effect, and whether 'enriching' a scenario by adding individualized attributes reduces the differences between the groups.* ***Methods:*** *Data were collected using semi-structured interviews among 127 patients with rheumatoid arthritis, 71 partners of these patients, and 64 members of the public. Participants indicated which aspects (attributes) were important to them when thinking about their life having (or imagining having) rheumatoid arthritis. These aspects were used to "enrich" the EQ-5D scenarios. Participants were asked to value in a Time TradeOff 1) their own experienced/imagined health; 2) a standard EQ-5D scenario, which (untold to them) represented their own health state; and 3) a standard EQ-5D scenario of their own health state (identified as such), enriched with individual attributes.* ***Results:*** *For all three states ratings were significantly higher for patients than for the general public. Ratings for partners were in between (and not all significantly different). No differences were seen between the three types of health states to be values.* ***Conclusion:*** *Differences between patients and the public are not likely to be due to sparseness of the scenarios, but may be due to a focusing illusion, enhanced by the negative framing of scenarios. It could be argued that for societal level decision making utilities from partners are more well-informed than those of the public, and suffer from less biases than those of patients.*

5.1 Introduction

Health state utilities play an important role in health care decision-making and health economics. The most important applications of utilities are in expected utility decision analysis and in cost-utility analysis. In such analyses, the expected utility for each possible strategy is calculated by combining the utilities for all possible health states with the probabilities of these states occurring. Utilities can be used at various levels of health-care decision-making, and the level determines whether they should be assessed from the general public or from patients. Gold et al.[12] have recommended the use of society's preferences in cost-utility analyses from a societal perspective, for macro-level decision-making, and utilities thus should be assessed from a representative sample of fully informed members of the general public. For guideline development, the meso-level, and for individual patient decision-making, the micro-level, they have recommended the use of utilities obtained from actual patients, assuming that the latter have a better understanding of the impact a health state can have. They have qualified their recommendations, however, basing their reasoning and their recommendations on the assumption that health state utilities in general differ between patients and the general public. In many instances, members of the public who are asked to imagine experiencing health states assign lower utilities to those states than do patients who are actually experiencing these states. The observation of the differences in utilities between respondent groups has therefore led to the question of whose utilities are most 'well-informed', and thereby most valid for use at the different levels of decision-making.[15] The answer to this question will depend on the reason for the differences. Are they due to better understanding by the patients of what the health states are really like? Or are differences due to cognitive mechanisms operating in patients during or after treatment, such as justification processes? Or, worse, are they simply artifacts of the method, caused by, for example, different objects of valuation? In most studies, patients are instructed to value "their health in the past week" or "their current health". "Health" may be specified further, referring to physical, psychological, and social well-being.[114] Members of the general public will mostly not be experiencing the health states to be valued. The respondents are therefore requested to value a health state description, or scenario, of the state. Generally, health state classification systems such as the EQ-5D,[89] or the Health Utilities Index, are used for this purpose. They give a limited description of the health state, usually in terms of functioning on some salient dimensions. The EQ-5D scenarios describe specific levels of impairment in mobility, self-care, and daily activities, and specific levels of pain/discomfort and anxiety/depression. Respondents are asked to imagine spending the rest of their life in the health state described in that scenario.

Relatively little research has been devoted to the development of these health state scenarios, despite early research in the field of utility assessment that showed the impact of framing of the scenarios on the elicited utilities.[91] Jansen et al.[20]found that patients undergoing radiotherapy assigned higher values to their “own health” during radiotherapy than to a scenario that aimed to describe their actual radiotherapy experience. The authors ascribed this difference to the scenario not matching the patients’ experience (’non-corresponding description’). Similar results were revealed in a meta-analysis of utilities assigned to prostate cancer.[46] Patients with prostate cancer rated a description of their health lower than their experienced health state. In a study using the EQ-5D health classification system, Insinga and Fryback[23] also found a difference between diseased respondents’ valuations of their current health (“however they conceived of it”) and their valuations of the EQ-5D scenario that by coincidence happened to describe their own health. They argued that ’non-corresponding description’, which they termed ’lack of scope’, best explained the different values that respondents assigned to their own health compared to the EQ-5D description of their health. They concluded that the scenarios of the EQ-5D are too sparse to describe certain health states. Further, the scenarios are often framed in negative terms. This leads to a focus on the negative impact of the health state, which might cause healthy subjects to overestimate the negative impact of a disease. Indeed focusing illusion has been suggested by many as an important explanation for the difference between health state utilities assigned by patients and public.[25] When the public imagines life in a certain health state they will concentrate on the difference between their current health state and the health state to be valued. Life domains influenced by the health state receive disproportional attention, whereas domains that will not be affected are ignored. The public concentrates on the negative aspects of an illness while patients give utilities in light of their experience in general.

Possibly EQ-5D descriptions should be enriched to create more resemblance between ratings of “own health” (self-ratings) and ratings of scenarios describing “own health” (self-identified EQ-5D ratings). In cost-utility research enriched EQ-5D descriptions have already been used to explore preferences for different medication types. Medication-related attributes added to the EQ-5D description induced differences in preferences between treatments.[90] In an earlier study we assessed the impact of enriching an EQ-5D scenario of patients' own health, but found it did not make it more similar to the patients' valuation of their own health.[115] Given the above it may be expected however, that the effect should particularly be seen in members of the general public. The inconclusive findings in patients therefore led us to the question whether enriching a scenario would make valuations of members of the general public more similar to those of patients. Furthermore, to obtain more in-depth insight into the problems of a possible better understanding that patients have[20] and of focusing illusion[25] we decided to include an additional population: partners of patients. To our knowledge no studies have assessed the valuations of this group. This is surprising since professionals have repeatedly been found to assign values that are intermediate between those of patients and of the public,[116] and partners can be expected to assign such intermediate values too, since they have vicarious experience, like professionals, and probably have a better understanding of the day to day adaptation to poor health than professionals. The aim of this study was to assess the difference in utilities between patients, their partners, and the general public, and to further investigate to what extent this may be due to an effect of using a scenario. We aimed to disentangle the effect of scenario vs. own health by eliciting valuations for three objects of valuation (in our case: rheumatoid arthritis health states) from each of the three groups, two groups who had to imagine living with a disease (members of the public and partners) and the third actually living with the disease (patients). Respondents valued their own imagined/experienced health state, a standard EQ-5D description of this health state, and an enriched EQ-5D description of this health state. We hypothesized that the value patients assigned to their own health would be highest, followed by that assigned by partners to their life with (imagined) RA, and that the lowest value would be assigned by the public to the standard EQ-5D state. We further hypothesized that adding individualized attributes to the EQ-5D (enriching it) would lead to higher valuations, and thereby lead to smaller differences between the valuations of the own experienced health and the scenario. Based on the findings of Smith et al.[49] and Llewelyn-Thomas et al.[91] we chose to enrich the standard EQ-5D description with individual patient attributes instead of giving more detailed but standard information.

5.2 Methods

5.2.1 Participants and procedures

The participant sample consisted of patients with rheumatoid arthritis (RA) aged 18 to 76 years old who had visited their treating rheumatologist in the past six months. From the database of the Leiden University Medical Center (LUMC), three hundred patients were randomly selected. In the selection method, we oversampled males to get an equal distribution of males and females, since RA is more prevalent in women. Medical records of the selected patients were reviewed to confirm the diagnosis of rheumatoid arthritis (RA) and identify co-morbid conditions. Of the 300 identified patients 50, patients had not been diagnosed with RA, and 7 had severe co-morbid conditions. The remaining 243 eligible patients were mailed information about the survey, including an informed consent form and asked to return the signed consent if they were willing to participate. Those who did not return the form within 3 weeks were telephoned as a reminder. Data were collected using self-report questionnaires and a semi-structured interview.

The partner sample consisted of partners of patients with RA (not of the above patients). We recruited partners by announcements at the outpatient clinic of the Department of Rheumatology of the LUMC and in the Newsletter of the Rheumatology Patient Society (In Beweging). Moreover, RA consultants and nurses at the outpatient clinic of the LUMC approached partners of patients to participate in the study. Partners contacted the researchers by phone or e-mail to make an appointment, preferably at the LUMC. All partners were telephoned by the research assistant to verify that the partner had no RA, was aged between 18 and 76 years old, that their spouse had been diagnosed with RA and did not participate as a patient in our study.

Members of the public were recruited through announcements in local newspapers distributed in Leiden and the surrounding mid-west region of the Netherlands. To meet inclusion criteria participants had to be between 35 and 76 years old. Those who had RA or whose partner had RA were excluded. The individuals who were invited to participate were mailed information about the survey, including an informed consent form. The medical ethics committee of the Leiden University Medical Center approved the study protocol.

5.3 The interview

Data were collected using semistructured interviews and self-administered questionnaires. Face-to-face interviews were performed by one of three trained interviewers. For patients and partners the interview took place at their preferred location: at home, at our department (an office area of the LUMC), or at work. Members of the general public were requested to visit our department. Patients who were interviewed in the hospital were not hospitalized at the time of the interview.

The interview started with the valuation of each participant's own experienced health of the previous week. The subsequent part of the interview was slightly different for patients, partners, and members of the public. All groups received the same questions but answered them from a different point of view. Patients were asked to answer questions about their own health during the previous week, partners were asked to imagine that they had their partner's (the RA patient's) health of the previous week, and members of the public read a detailed health state description of RA (see appendix C) and were asked to imagine that they had the RA state described, and to thus imagine their previous week, having this health state. At any time point during the interview the RA health state description could be reread.

Next, participants answered the EQ-5D questionnaire, five items on mobility, self-care, usual activities, pain/discomfort and anxiety/depression, on a three point scale: no problems, some problems, and no function at all, or in the case of pain, extreme pain.[89] They were asked to indicate the level of problems on each of the EQ-5D dimensions for their own health, their partner's health, or the health state description, respectively. This state is from now on referred to as "own standard EQ-5D" state, even though strictly speaking it was only their own state for the RA patients.

After this EQ-5D questionnaire two filler questionnaires followed, i.e. the Patient Satisfaction Questionnaire[92] and the Rosenberg Self-Esteem Scale,[93] to distract participants' attention from the answers they had given on the EQ-5D questionnaire. In the next part, participants were asked to value five standard EQ-5D states. Five of these EQ-5D states were retrieved from previous research with patients with RA, covering the full utility range from 0 to 1 according to the UK tariff.[11] A description of these health states can be found in appendix B. Unknown to the participants, the sixth health state was their own standard EQ-5D state of the previous week as indicated in the EQ-5D questionnaire. The computer retrieved the answers of the participant earlier in the interview and created the own standard EQ-5D state for this participant. All standard EQ-5D states were randomly presented except for the participant's own standard EQ-5D state, which was always presented as the fifth state. Participants were not informed that it was their own standard EQ-5D health state. If one of the five pre-selected EQ-5D states happened to be the same as the own standard EQ-5D state, this state was replaced automatically with the EQ-5D state that should have been presented in the sixth place.

Following this valuation task respondents were asked to name aspects important to their health during the last week, and partners and members of the public were asked the same questions but imagining that they had the RA state of their partner or as given in the description. An open-ended question was used to indicate these dimensions, with a maximum of five dimensions. The interviewer entered these attributes in the computer. It was impossible to add a full description of each attribute; consequently a keyword was used. The interviewer and the participant together created suitable keywords for each of the individual attributes. Only keywords on which the participant agreed were used. These individual attributes were then combined with the own standard EQ-5D state of previous week to create an own enriched EQ-5D state. On the computer screen, the description of the own enriched EQ-5D state was shown with the individual attributes represented beneath the five standard attributes. It was made clear to the participants that the order in which the attributes were presented was arbitrary and that it was up to the participants how important the attributes were to them. Furthermore, participants were told that the description as stated on the computer fitted their own health state. If this was not clear, the interviewer explained how this description was created and made sure that the participant understood that it was his or her own (experienced or imagined) health state.

All health states were valued using a time tradeoff (TTO). Participants rated how many years (x) of their remaining life expectancy (y), derived from Dutch life expectancy tables,[117] they were willing to trade to obtain perfect health. Utility was calculated as $\frac{y-x}{y}$. The computer program Ci3[94] was used to elicit the utilities based on a ping-pong search procedure. On the computer screen a short description of perfect health and the health state to be valued were presented. Perfect health was described as full well-being, physically, psychologically and regarding social activities.

5.4 Results

5.4.1 Participants

Of the 243 patients selected, 132 patients gave their approval to be interviewed (54%). No differences in age and time since diagnosis between responders and non-responders were found. Of the responders, one patient with emotional problems and two patients who were not able to speak and understand Dutch were excluded. Data of 2 participants created multivariate outliers and were excluded from further analyses (Mahalanobis distance, $F(3) = 31.07$ and F $(3) = 18.05$). The interviews took place at the hospital ($N = 82$), at the patients' home ($N = 44$) or at work ($N = 1$). Of the 74 partners of RA patients who originally gave their approval to be interviewed 73 were actually interviewed. One participant could not be reached. The interviews took place at the hospital ($N = 45$), at the participants' home ($N = 26$) or elsewhere ($N = 2$). Of the 69 members of the general public who responded two were excluded: one of the excluded participants had RA and the other one had a partner who had RA. Two participants decided to cancel the interview, both due to a tight time schedule. All 65 interviews took place at the hospital. Table 5.1 presents the demographic information of the participants.

5.4.2 Valuations of the three health states

Table 5.2 shows the means and standard deviations of the three health state valuations for the three groups. As expected, scores were highest for patients, lowest for the public and in between for partners. Using MANOVA, we found statistically significant differences between the three participant groups for all three health states. Since the higher utilities of the patients and partners may have simply been due to better health of the patients (better than the health state description presented to the public), we redid the analysis, using the tariff of the EQ-5D states scored by the respondents as covariate. Estimated marginal means are shown in Table 5.3. Using MANCOVA, we found an overall effect for respondent group ($F(6, 504) = 2.2, p = 0.04$). Between-subjects analysis showed a highly statistically significant difference between the groups for the enriched state, and a borderline effect for own (imagined) health and the standard EQ-5D state. Post hoc pairwise comparisons, adjusted for multiple comparisons, showed significant differences between patient and public for the own (imagined or experienced) health ($p = 0.02$), the standard EQ-5D state ($p = 0.03$), and the enriched EQ-5D state ($p < 0.001$). Differences between patient

Table 5.1 Participant Characteristics

	Patients ($N = 127$)	Partners($N = 72$)	Public($N = 65$)
Age (*mean* ± *s.d.*)	58 ± 11	56 ± 11 56 ± 11	
Gender			
Female	61(48%)	17 (23%)	50 (77%)
Education*			
Nine years or less	38(30%)	13 (18%)	5 (8%)
Between 10 and 12 years	62(49%)	21 (30%)	29 (45%)
13 years of more	24(19%)	36 (52%)	31 (48%)
Children			
Yes	107(84%)	58 (80%)	49 (75%)
Marital status			
Married/ Spouse	111(87%)	72 (100%)	37 (57%)
Divorced/ Widow	10 (8%)	-	13(20%)
Single	6 (5%)	-	8 (12%)
Tariff own standard EQ-5D state (mean± s.d.)	0.63 ± 0.28	$0.51 + 0.30$	$0.53 + 0.14$

*Numbers do not add up to total N due to missing data.

Table 5.2 Means and standard deviations of the valuation of the different health states

	Patients ($N = 122$)		Partners ($N = 71$)		Public ($N = 64$)		p-value
	Mean	SD	Mean	SD	Mean	SD	
Own (imagined/experienced) health	0.79	0.23	0.69	0.27	0.66	0.27	0.002
Standard EQ-5D state	0.81	0.25	0.70	0.27	0.67	0.29	0.001
Enriched EQ-5D state	0.83	0.22	0.71	0.26	0.66	0.28	<0.001

and partner and between partner and public were of borderline significance for the enriched EQ-5D state only (p = 0.10and 0.08, respectively). Table 5.3 shows the means and standard deviations of the 3 health state valuations for the 3 groups.

Contrary to our expectation, repeated measurement ANOVA showed no effect of health state valued (own vs. standard EQ-5D vs. enriched EQ-5D), nor an interaction between group (patients, partners, public) and state valued. Differences were minimal and own state, if anything, was lower than the other two states, not higher. In a comparison as commonly described in the literature, the valuation of the standard EQ-5D state of the public from Table 5.1 would have been compared to the patient's own health (0.67 vs 0.79). We indeed found the expected difference. However, we had expected to bridge this gap by enriching the scenario, which did not occur (enriched state: 0.66).

5.5 Discussion

Aim of this study was to assess the difference in utilities between patients, partners, and the general public, and to further investigate to what extent differences may be due to an effect of valuing a scenario vs. valuing one's own health. To our knowledge this is the first study assessing utilities in partners of patients. As we had expected, patients assigned the highest value to all states, followed by partners and the general public. However, we did not see a difference by object of valuation, i.e., whether participants imagined themselves living in a state described as a 'sparse' EQ-5D scenario, an enriched EQ-5D scenario, or their own health (patients), the health of their partner with RA (partners), or a detailed description of an RA state (public). Particularly for the general public we had expected that enriching an EQ-5D scenario would increase the valuations, bringing these closer to those of the patients. In this group, however, the three valuations were approximately the same, and 0.12-0.15 lower than those of patients. The findings in partners and the public are a confirmation of our findings in patients,13 that the generally found difference between patients and the public is not due to the scenarios being too sparse, a suggestion that had been made both by Jansen et al.[20] and Insinga and Fryback.[23] It may still however be the case that they lack positive aspects, as also argued by the latter. From a qualitative analysis of the open ended question that was used to elicit attributes to enrich the scenarios it appeared that the public was focused on life domains that were negatively influenced by RA: they concentrated on not being able to perform work, leisure activities and sports anymore. Thus, differences between

Table 5.3 Estimated marginal means and standard errors of the valuation of the different health states, controlling for the tariff of the EQ-5D description as scored by the participant

	Patients ($N = 122$)		Partners ($N = 71$)		Public ($N = 64$)		p-value
	Mean	SD	Mean	SD	Mean	SD	
Own (imagined/experienced) health	0.77	0.021	0.72	0.027	0.68^{1}	0.029	0.06
Standard EQ-5D state	0.78	0.021	0.74	0.027	0.70^{1}	0.028	0.08
Enriched EQ-5D state	0.80	0.020	0.75^{2}	0.026	$0.68^{1\ 3}$	0.027	0.002

[1] significantly different from valuation of patient ($p < 0.05$)
[2] significantly different from valuation of patient ($p < 0.05$)
[3] significantly different from valuation of partner ($p < 0.05$)

patients and the public are not likely to be due to sparseness of the scenarios, but may be due to a focusing illusion, enhanced by the negative framing of scenarios.

In the discussion about the difference between patients and public it has also been suggested that patients overestimate their well-being, leading to a positive bias.[107] Patients might be unrealistically optimistic about their situation,[108] or even suffer from self-deception: “people with disabilities . . . deceiving themselves about how close to normal they are”. However, the mentioned qualitative analysis did not indicate patients to positively exaggerate their life domains. Further, their utilities for their own health were not higher than those for the EQ-5D scenario that, unknown to them, represented their own situation, and were not close to 1. The finding of partners’ valuations being intermediate, points to an effect of (vicarious) experience with the disease, leading to a better ability to anticipate adaptation than the public. An important conclusion of research in behavioral economics that will also apply to utility assessment is that humans are poor in predicting how they will value a situation once it is experienced, so-called “poor hedonic forecasting”.[26] Processes of adaptation, e.g., appear difficult to anticipate. It has been suggested that members of the public use a transition heuristic when valuing health, i.e., they focus on the time of entering a poor health state and do not consider adaptation to that state. This will result in these respondents undervaluing states. As stated above, members of the public indeed focused on attributes that were negatively influenced by a health state. Further, the public ranked the EQ-5D dimensions as more important than the self-rated dimensions, whereas patients found both groups of attributes equally important. This indeed points to a transition heuristic or a focusing illusion.[25] Patients in this study were less concerned about not being able to work or play sports. Possibly they have learned to live without these activities and have found new life goals. Gold et al.[12] have argued that “techniques that create a better understanding in the general public of the experience of differering health states will be highly useful in strengthening this field” . Using adaptation exercises as proposed by Damschroder[38] may serve as such to overcome the poor hedonic forecasting of humans. Our findings in partners indicate that there is indeed something to gain here.

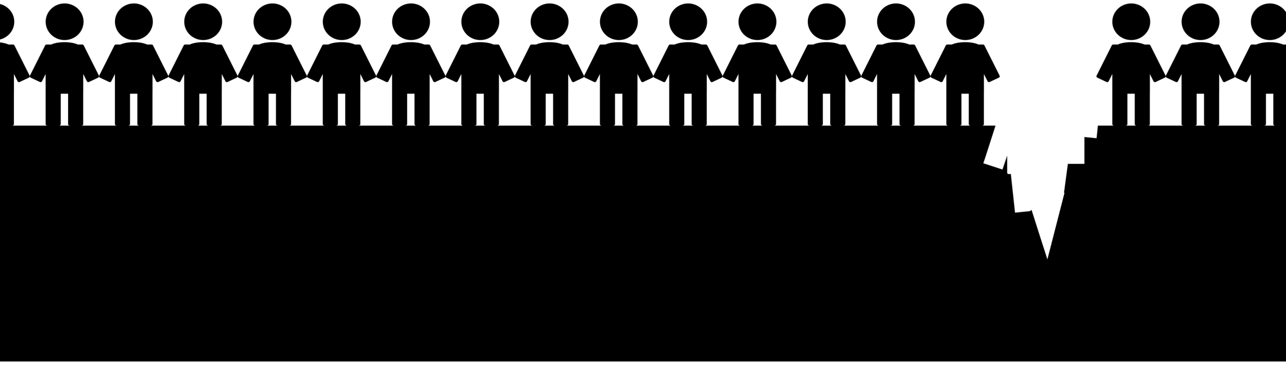
Other explanations often given for the gap between patients and members of the public are loss aversion and status quo bias. Good health is considered a gain to patients, and poor health a loss to healthy subjects. Losses weigh more strongly than gains, and thereby utilities are lower for healthy subjects. Related to this is status quo bias, or the endowment effect, which shows that people value goods more highly once they own them. One may expect this bias to be stronger for respondents valuing their own health (which they ’own’) than for respondents valuing hypothetical states (which they are asked to imagine ’owning’), leading to higher utilities in patients. Both loss aversion and an endowment effect might also explain why patients gave somewhat, although not significantly, higher values to the enriched than to the standard EQ-5D scenario (0.83 vs. 0.81), since for the latter they were not made aware that it was their own health state. If differences between groups were entirely due to loss aversion, however, partners’ valuations should have been similar to those of the general public, whereas they were in between those of patients and the public.

In the end it remains a matter of principle, or even a philosophical issue, which mechanisms to allow in the process of utility elicitation for societal decision making, and which to avoid. Even if patients have a more realistic picture of their life, they will still suffer from loss aversion, a mechanism that one may not wish to include in utilities used for cost-utility analysis, or from other cognitive mechanisms operating during the tradeoff task, such as justification processes, which, we would argue, should not be reflected either. As regards the failure to anticipate adaptation, Gilbert has argued that this may be a Darwinian mechanism “An organism aware of its ability to construct its own satisfaction might well loose its preferences for one outcome over another and become happily extinct” (p. 635).[26] Nevertheless one may argue that for allocating health care resources one does not want to incorporate all adaptation, since health care ultimately is aimed at improving health and preventing ill health. But clearly an improvement over current practice would be to incorporate more positive aspects in the scenarios used to elicit utilities from the public. An issue for debate in our opinion would be the use of partners of patients, since they by their vicarious experience they have a more realistic view of the health states than the public, but they do not suffer from some of the mechanisms described above that operate in patients.

6

Effect of adaptive abilities on utilities, direct or mediated by mental health?

Peeters, Y., Ranchor, A.V., Vliet Vlieland, T.P.M., & Stiggelbout, A.M. (2010). Effect of adaptive abilities on utilities, direct or mediated by mental health? Health Qual.Life Outcomes., 8, 130.



Abstract ***Objective:*** *In cost-utility analyses gain in health can be measured using health state utilities. Health state utilities can be elicited from members of the public or from patients. Utilities given by patients tend to be higher than utilities given by members of the public. This difference is often suggested to be explained by adaptation, but this has not yet been investigated in patients. Here, we investigate if, besides health related quality of life (HRQL), persons' ability to adapt can explain health state utilities. Both the direct effect of persons' adaptive abilities on health state utilities and the indirect effect, where HRQL mediates the effect of ability to adapt, are examined.* ***Methods:*** *In total 125 patients with rheumatoid arthritis were interviewed. Participants gave valuations of their own health on a visual analogue scale (VAS) and time trade-off (TTO). To estimate persons' ability to adapt, patients filled in questionnaires measuring Self-esteem, Mastery, and Optimism. Finally they completed the SF-36 measuring HRQL. Regression analyses were used to investigate the direct and mediated effect of ability to adapt on health state utilities.* ***Results:*** *Persons' ability to adapt did not add considerably to the explanation of health state utilities above HRQL. In the TTO no additional variance was explained by adaptive abilities* ($\Delta R^2 = 0.00, \beta = 0.02$)*, in the VAS a minor proportion of the variance was explained by adaptive abilities* ($\Delta R^2 = 0.05, \beta = 0.33$)*. The effect of adaptation on health state utilities seems to be mediated by the mental health domain of quality of life.* ***Conclusions:*** *Patients with stronger adaptive abilities, based on their optimism, mastery and self-esteem, may more easily enhance their mental health after being diagnosed with a chronic illness, which leads to higher health state utilities.*

6.1 Introduction

In health care, decisions are made about treatment at the level of individual patients, of patient groups (guideline development), and at the societal level.[15] Decisions about guideline development and decisions at the societal level are often guided by cost-utility analyses. In these analyses the gain in health obtained by treatment is compared with the costs that have to be made in order to obtain this gain.[12] To assess the value of this gain, cost-utility analyses make use of health state valuations, i.e. health state utilities.

A health state utility is a preference for a particular health state compared with perfect health and immediate death. Utilities can be seen as a global valuation of health related quality of life (HRQL)[118] and can be expected to show a strong relationship with health status. Nevertheless, only between 18% and 43% of the variance in health state utilities can be explained by HRQL. Most of the variance still remains unexplained.[119]

Health state utilities can be elicited from members of the public and from patients. Members of the public tend to give lower health state valuations, compared to patients.[100] This discrepancy in health state valuations has, among others, been suggested to be explained by the failure of members of the public to anticipate on their ability to adapt. Patients adapt to the physical and psychological challenges of their illness.[120] When health state valuations are elicited from patients, some of the variance in health state utilities might be explained by this adaptation.[18, 20, 36]

Tentative support has been found for the effect of adaptation on health state valuations. Members of the public who were made aware of their ability to adapt gave higher valuations on a person trade-off (PTO) and on a visual analogue scale (VAS) measuring quality of life,[24, 38] but not on the time trade-off (TTO) nor on the standard gamble (SG).[39] Whether health state utilities given by patients are actually correlated with adaptation has not been topic of study yet.

Adaptation can be defined as a response that diminishes or remains the same despite constant or increasing stimulus levels.[31] The outcome of adaptation can be measured by change over time, such as change in well-being[107] or life satisfaction.[121, 122] If researchers aim to gain more insight in the process of adaptation itself, adaptation can be conceptualised through certain coping-strategies.[123, 124] These coping-strategies are, among others, enabled by personal resources.

By studying adaptation Taylor[32] developed the Cognitive Adaptation Theory (CAT) which is based on cognitive interviews with chronically ill persons. This theory is one of the dominant theories in health psychology and has often been used to empirically test adaptation. Research using this theory suggests that psychological adjustment to an illness occurs around four themes; a search for meaning in the experience, an attempt to regain mastery over the event and over one's life more generally, an effort to enhance one's self-esteem, and the ability to find positive illusions, i.e. optimism. These concepts as suggested in the CAT are further described below.

After a threatening event, people often cannot find a sense of meaning in the experience and lose their feelings of mastery and of self-esteem. Most people manage to re-establish these over time. According to Taylor, this re-establishment is based on so-called positive illusions. People develop unrealistic beliefs that make it possible to regain control over the event and over one's life and to regain self-esteem.[32] Although positive illusions may create unrealistic and maybe 'false' ideas, these illusions have been found to be important resources.[45]

Previous studies have shown that patients who score high on indicators of CAT have better psychological functioning,[125–128] they are less anxious and depressed, report more vitality and have a better mental functioning.[126, 129, 130] Moreover, patients with a higher score on indicators of CAT reported better physical functioning,[126, 127] they showed fewer new coronary events or hospital admissions[125, 130] and lived longer.[131] It thus appears that patients who have higher self-esteem, mastery, and optimism, and who find a meaning in the experience have better abilities to adapt.

No standard method is available for investigating the ability to adapt based on CAT. Studies have used different indicators and methods for their analyses. For instance, studies have included indicators measuring optimism, mastery and self-esteem, but often exclude finding meaning. To our knowledge, only in two studies the effect of finding meaning was included.[131, 132] The rationale to exclude benefit finding was described by Major et al.[133] and Chan et al.[127] Both research groups suggest that mastery, self-esteem, and optimism are stable personality traits representing a persons' resilience, whereas finding meaning might be seen as a process facilitated by these personality traits.

Apart from this variety of indicators of CAT included to measure adaptive abilities, studies have also used different ways to measure these indicators. Some studies have analysed the effects of the different indicators separately,[128, 134] some have created a scale taking the indicators together,[130–132, 135] and again others have investigated each indicator separately as well as an aggregate scale of the indicators together.[125–127] The latter studies revealed that besides the effect of the aggregate scale, often only one of the indicators had an effect on the outcome measurement. Since the overall results of these studies show different 'single' indicators to reveal an effect, indicators of persons' abilities to adapt cannot be simplified to one single indicator. Exploring the results of these studies further, it seems that significant effects have mostly been seen in studies using an aggregate scale. Therefore, in the present study persons' ability to adapt is constructed with an aggregate scale based on mastery, self-esteem and optimism.

The first aim of this study was to investigate if above HRQL, persons' adaptive abilities explain health state utilities. That is:

- Do adaptive abilities account for the unexplained variance in health state utilities above the variance explained by HRQL?

Another possibility is that adaptation, in this study measured through persons' ability to adapt, has an indirect effect on utilities, through HRQL. As described above, adaptive abilities does affect psychological and physical functioning.[130] This would fit the hierarchical model of Spilker and Revicki,[136] in which three levels of quality of life are distinguished that have mutual impact on each other. The hierarchy of this model ranged from a global level such as a health state utility, to HRQL domains, and to specific determinants of domains such as personality characteristics,[136] which may include adaptive abilities. Thus, the second aim of this study was to investigate if adaptive abilities affect health state utilities via HRQL domains.

- Is the relation between adaptive abilities and health state utilities mediated[137] by HRQL domains?

Since we investigated psychological adaptive abilities we assume from a theoretical point of view that only mental health can mediate this relation.

6.2 Methods

6.2.1 Participants and design

We chose to study our research questions in a sample of patients with rheumatoid arthritis (RA). RA concerns a chronic disease with a wide spectrum of manifestations, for which adaptation is relevant, since no cure is available. From the database of the Leiden University Medical Centre, 300 people who were between 18 and 76 years old and had visited their treating rheumatologist in the previous six months were randomly selected. In total 1054 patients had visited their rheumatologist in the past six months. These patients were randomly numbered. First, 400 numbers had been drawn (using the software Excell) as a selection for a different study.[138] Of the remaining 654 patients 90 patients had to be excluded due to age restrictions, and 10 were excluded because they had refused to participate in a similar study.[139] Next, to get equal male/female distribution, 150 male patients and 150 female patients were randomly selected to participate in the current study. Based on the medical records, 50 people who had not been diagnosed with RA, and seven with severe co-morbid conditions were excluded. The remaining 243 eligible people received information about the survey by mail, including an informed consent form. Patients who did not return the informed consent form within three weeks were called as a reminder. Data were collected using self-report questionnaires and a semi-structured interview. The medical ethics committee of the Leiden University Medical Centre approved the study protocol.

6.2.2 The interview

Face-to-face interviews were performed by three trained interviewers following a strict interview protocol. The interviews took place at the persons' preferred location; at home, in the hospital, or at work. A full description of the interview can be found elsewhere.[115] Here, only the part of the interview used to gather the information necessary for this study is described.

At the beginning of the interview, people valued their health of the previous week using a visual analogue scale (VAS) and a time tradeoff (TTO). Next people completed three questionnaires: the EQ-5D questionnaire,[105] two scales of the Patient Satisfaction Questionnaire[140] and, the Rosenberg Self-Esteem Scale.[93] In this study only the information retrieved by the Rosenberg Self-Esteem Scale will be used. After the interview, people were asked to complete a questionnaire at home to lessen the burden. Among others this questionnaire included the Life Orientation Test,[141] the Mastery scale of Pearlin and Schooler,[142] and the MOS 36-item Short-From Health Survey (SF-36).[143]

6.2.3 Instruments

The Visual Analogue Scale (VAS) The VAS is a 100 mm horizontal line ranging from death to perfect health. Perfect health was described as full well-being in physical, psychological, and social functioning. Utility for the own health state of last week was elicited by asking respondents to place a mark between death and perfect health. The Time tradeoff (TTO) The computer program Ci3[94] was used to elicit the TTO utilities based on a ping-pong search procedure. On the computer screen a short description of perfect health and a description of the patient's own health state of the previous week were presented. Perfect health was again described as full well-being in physical, psychological and social functioning. People rated how many years (x) of their remaining life expectancy (y), derived from Dutch life expectancy tables,[117] they were willing to trade to obtain perfect health. Life expectancy was used as the time frame since it was shown to be more meaningful to the participant[144] and to lead to less loss aversion.[145] Utility was calculated as $\frac{y-x}{y}$.

6.2.4 Indicators for persons' adaptive abilities

Personal Control The Mastery List of Pearlin and Schooler[142] measures the extent to which people feel they are in control of their lives. People indicated their agreement with seven items such as "I can do about anything I really set my mind to do", on a five-point Likert scale ranging from 'totally disagree' to 'totally agree'. Total score ranged from $7-35$, with a higher score indicating more control. Good internal consistency ($\alpha = .58 - .70$) was reported previously among patients with a chronic illness.[130] Self-Esteem With the Rosenberg self-esteem scale[93] the positive or negative valuation people have toward themselves was measured. People rated how much they agreed with 10 statements such as "I feel I have a number of good qualities", on a four-point Likert scale. The total score of the scale ranges from 0-30, with a higher score indicating higher self-esteem. Among patients with a chronic illness good internal consistency ($\alpha = 0.83 - 0.90$) and test-retest reliability ($\rho = 0.71$) were reported previously.[130, 146]

Optimism The Revised Life Orientation Test (R-LOT)[147] consists of three items measuring pessimism, three items measuring optimism and four filler items. Items such as "In uncertain times, I usually expect the best", were scored on a five-point scale ranging from 'strongly disagree' to 'strongly agree'. The total score, ranging from 0-24, was calculated after recoding items measuring pessimism. A higher score indicates more optimism. The R-LOT previously revealed good internal consistency ($\alpha = 0.74 - 0.89$) and test-retest reliability ($\rho = 0.67$) among patients with a chronic illness.[130, 146]

Health related quality of life HRQL was measured with the SF-36.[143] The SF-36 comprises eight multi-item dimensions which can be summed into a physical and a mental component score (SF-36 PCS and SF-36 MCS). Scores in each component range from 0-100, with higher scores indicating better HRQL.

6.2.5 Data analysis

Prior to the main analyses, all variables were examined for uni- and multivariate outliers, missing data, linearity and normality. Missing data were excluded listwise. Principal component analysis was performed to check if the constructs 'personal control', 'self-esteem' and 'optimism' could be combined in one scale. The number of factors were decided upon by an eigenvalue > 1 and the scree plot. If the constructs measured one underlying factor, the standardized total scores of the separate constructs were summed and used as one scale measuring adaptive abilities. To further analyze the reliability of this scale Cronbach's alpha was calculated.

Hierarchical linear regression was conducted to assess if adaptive abilities could explain the variance in utilities above that explained by HRQL. To control for HRQL, the total scores on the PCS and MCS were entered first. In the next step the adaptive abilities were added. Separate analyses were performed for the VAS and TTO. Mediation analyses were performed as suggested by Baron & Kenny.[148] First we investigated if adaptive abilities affected mental health; second, the relation of mental health with health state utilities was investigated; third we investigated the direct effect of adaptive abilities on health state utilities without controlling for mental health, and finally we checked if after controlling for mental health the direct effect of adaptive abilities and health state utilities decreased (partial mediation) or even became zero (full mediation).[148] When partial mediation was shown, the Sobel test statistic[149] was used to test the strength of the mediation.

6.3 Results

Of the 243 people selected, 132 people gave their approval to be interviewed (54%). No differences in age and time since diagnosis between responders and non-responders were found. Of the responders, one person with emotional problems, and two persons who were not able to speak and understand Dutch were not invited for the interview. Of the interviewed patients four were excluded; three people could not finish the interview due to cognitive or concentration problems, and one person returned the questionnaire after more then a month. All variables met the assumptions for linearity and normality, except for health state utility measured by the TTO ($skewness = -1.36, SE = 0.22$).

Table 6.1 Characteristics of people with RA included in this study

	patients with RA ($N = 125$)
Age ($Mean(range) \pm SD$)	58(29-75)±10.80
Gender	
Female	60(48%)
Education[a]	
Nine years or less	38(31%)
Between 10 and 12 years	62(49%)
13 years of more	24(19%)
Children	
Yes	105(84%)
Marital status	
Married	110(88%)
Divorced/ Widow	9 (7%)
Single	6 (5%)
Time since diagnosis (years) ($Mean(range) \pm SD$)	13(2 - 47)±9.26
Health state Utilities	
VAS ($Mean(range) \pm SD$)	66.14 (14 - 100)± 19.15
TTO ($Mean(range) \pm SD$)	0.77 (0 - 1)± 0.25
Health status	
SF-36 PCS ($Mean(range) \pm SD$)	36.46 (12-58)± 10.66
SF-36 MCS ($Mean(range) \pm SD$)	52.36 (24-67)± 9.66

[a] Numbers do not add up to 125 due to missing data.

The interviews were administered by three trained interviewers (following a strict script), and took place at the LUMC ($N = 83$), at the respondent's home ($N = 41$) or at work ($N = 1$). People were not hospitalized at the time of the interview. Persons interviewed at home had on average more health problems than persons interviewed in the LUMC based on the SF-36 PCS score. No interviewer effect was found on the answers patients gave. Table 6.1 presents the demographic information of the 125 people who were included.

6.3.1 Creating a scale measuring persons' ability to adapt

Principal component analysis of the three indicators of persons' ability to adapt (Personal control, Self-esteem, and Optimism) could be aggregated to one compo-

Table 6.2 Pearson correlations of study variables

	TTO	VAS
Persons' ability to adapt	$0.33^{\sharp}$	$0.65^{\sharp}$
SF-36 PCS	$0.30^{\sharp}$	$0.57^{\sharp}$
SF-36 MCS	$0.33^{\sharp}$	$0.43^{\sharp}$

$^{\sharp}p < 0.01$

nent. This component explained 73% of the variance, the component loadings for self-esteem, personal control and optimism ranged from 0.81 – 0.88. With reliability analysis the scale measuring persons' ability to adapt showed good internal consistency, Cronbach's $\alpha = 0.80$.

6.3.2 Predicting utilities

Before hierarchical regression analyses, the associations between the utility measures and demographic characteristics (time since diagnosis, gender, age, having a partner, having children, and education) and the study variables (PCS, MCS, and persons' ability to adapt) were checked with Pearson correlations. The demographic characteristics showed low to no correlation with the TTO and VAS (all $\rho < 0.20$). All study variables showed moderate to strong intercorrelation (Table 6.2).

6.3.3 Adaptive ability as direct predictor of TTO and the VAS, over and above HRQL

Table 6.3 presents the relationships of HRQL and persons' ability to adapt with utilities measured by the TTO and VAS, using a two-step hierarchical regression analysis. HRQL explained 19% of the TTO and 49% of the VAS. After correcting for HRQL, persons' ability to adapt did not predict additional variance in the TTO. On the VAS 5% additional variance was explained by persons' ability to adapt.

Although persons' ability to adapt had no direct effect on health state utilities over and above the HRQL domains, it might have had an effect on HRQL domains that in turn affect health state utilities (mediation). Therefore this mediation effect was examined next. Firstly, it was found that persons' ability to adapt affected mental health, after correction of physical health ($\Delta R^2 = 0.46, p < 0.01$). Secondly, mental health was related to health state utilities ($\Delta R^2 = 0.11, p < 0.01$ for the TTO and $\Delta R^2 = 0.18, p < 0.01$ for the VAS). Third, without correcting for mental

Table 6.3 Hierarchical regression analyses direct influence of adaptive abilities on TTO and the VAS above HRQL

Predictors		ΔR^2	B	β	
TTO					
$N = 123$	Step 1	0.192			$p < 0.001$
	SF-36 PCS		0.006	0.265	$p < 0.001$
	SF-36 MCS		0.009	0.331	$p < 0.001$
	Step 2	0.000			$p = 0.886$
	SF-36 PCS		0.006	0.260	$p = 0.006$
	SF-36 MCS		0.009	0.319	$p = 0.006$
Persons' ability to adapt			0.000	0.018	$p = 0.886$
VAS					
$N = 123$	Step 1	0.487			$p < 0.001$
	SF-36 PCS		0.956	0.529	$p < 0.001$
	SF-36 MCS		0.848	0.420	$p < 0.001$
	Step 2	0.048			$p = 0.001$
	SF-36 PCS		0.761	0.421	$p < 0.001$
	SF-36 MCS		0.432	0.214	$p = 0.014$
Persons' ability to adapt			0.441	0.325	$p = 0.001$

health, persons' ability to adapt ($\Delta R^2 = 0.06, p < 0.01$) did have a direct effect on health state utilities measured with the TTO and with the VAS ($\Delta R^2 = 0.20, p < 0.001$). Finally, we found that the effect of persons' ability to adapt on both utility measurements decreased after controlling for mental health. As can be seen from Table 6.3 (explained previously) persons' ability to adapt was completely mediated by mental health when health state utility was measured with TTO. The explained variance of VAS by persons' ability to adapt on VAS decreased from 20% to 5% when mental health was added, which was a significant change (Sobel test statistic[149] $= 5.45, p < 0.01$), indicating partial mediation.

6.4 Discussion

In discussion sections of papers and in theoretical manuscripts, the difference in health state utilities between people with a chronic illness and the public is often explained by adaptation.[15,107] The results of this study show that adaptive abilities are indeed related to utilities, but that this effect is fully mediated by mental health for the TTO, and partly mediated for the VAS. It seems that in people with a chronic illness a stronger ability to adapt may lead to better mental health, which in turn leads to higher health state utilities. The suggested relation between adaptation and health state utilities given by people with a chronic illness does not occur directly, but appears to be mediated by mental health. Admittedly, this conclusion has to be made with caution since not adaptation but adaptive abilities are studied here.

Adaptive abilities explained 46% of the variance in mental health, which in turn explained between $11 - 18\%$ of the variance in health state utilities after correction for physical health. Arnold et al.,[150] already suggested such a mediation effect. They found that people with a chronic illness do not differ from healthy people in global quality of life and that global quality of life is mostly explained by mental functioning. Based on these findings they argued that people with a chronic illness psychologically adapt, causing a recovery of their mental health, which leads to recovery of global quality of life. The cross-sectional design of this study limits the points described above. From this study no conclusions can be drawn about the causal relationship between persons' ability to adapt, HRQL, and health state utilities. Nevertheless, causal relations between persons' ability to adapt and HRQL have been described previously.[128,130] Future longitudinal research is necessary to further investigate this causal relationship.

The index based on CAT to measure persons' ability to adapt, has been used in several studies but has not yet been validated. Given the number of studies using such a scale based on the CAT, validation is pressingly needed. Further, this index has been suggested to reflect stable personality traits, which might not change over time.[133] If adaptive abilities are indeed stable over time, then health state utilities of members of the public might be influenced in a similar way. Yet since members of the public find it difficult to anticipate on their ability to adapt[24] we still would expect a less substantial effect of adaptive abilities on HRQL and health state utilities in this population.

HRQL predicted 20% of the variance in the TTO, and 49% of the variance in the VAS. These results are comparable with previous findings concerning the relationship between HRQL and health state utilities.[151] The smaller amount of variance explained in the TTO compared to the VAS might be caused by the trading process. In this trading process, a series of information processing activities and construction of subjective values for dimensions are developed, making the variance in TTO-scores difficult to explain. Another explanation may lie in the cognitive nature of the TTO. Campbell[152] suggested that quality of life can be assessed with cognitive or with affective measurements. Cognitive measurements depend on a more intellectual process while affective measurements depend on subjective feelings. The TTO might be seen as a more cognitive measurement, the VAS as a more affective measurement. After a life event, the affective component of well-being appears to be more impaired than the cognitive component, which means that this component is sensitive to change and the cognitive component is more stable.[153] Finally, a more methodological explanation for the smaller amount of variance explained in the TTO might be that the TTO was skewed. When a dependent variable is skewed a smaller effect size might be anticipated.[154]

This study included patients with RA who had been diagnosed on average 13 years before. First, it can be questioned if patients still need to adapt to their illness so many years after diagnosis. It seems evident that adaptation takes place in the initial phase of the illness. However, the disabling, often progressive and uncontrollable characteristics of RA might result in adaptive processes, even after so many years. The results of this study indicate that adaptive abilities indirectly explain health state utilities, so this result might become more distinct when examining patients in the initial phase of their illness. Secondly, RA is a chronic illness characterized by pain and deformity of the joints, leading to physical limitations. There is evidence suggesting that people do not adapt to unpredictable stressors such as pain.[109] On the other hand, patients with RA might be able to adapt to other aspects of their illness such as the physical limitations by learning new ways to perform activities and they might learn to accept their pain.[155] More research is necessary to investigate the effect of adaptive abilities on health state valuations in other patient groups.

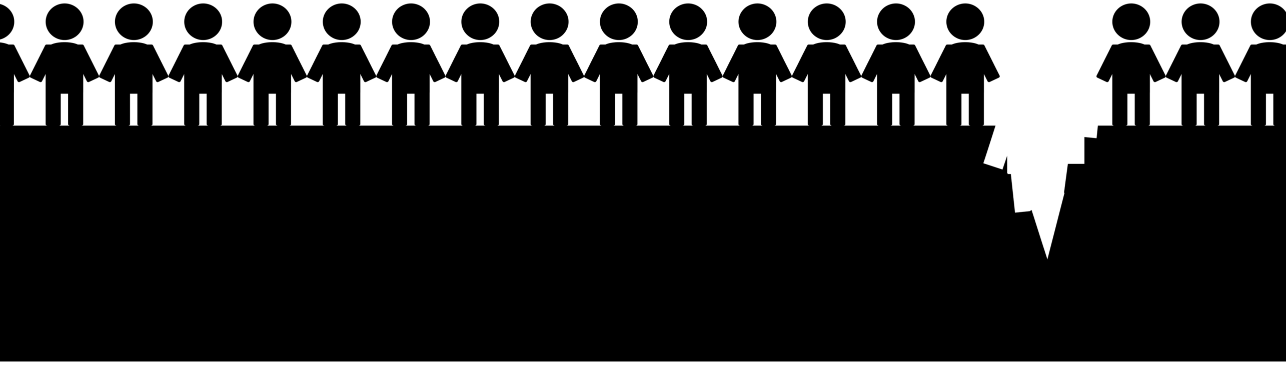
6.5 Conclusion

In conclusion, the results in this study seem to indicate that adaptive abilities indirectly explain health state utilities. Assuming that these adaptive abilities induce adaptation, then cost-utility analyses could partly be founded on utilities shaped by adaptation. Such utilities will result in less room for improvement between the patient's own health condition and perfect health, leading to a lack of justification to treat an illness.[36] Based on this challenge, one could argue that members of the public should provide valuations instead, but these respondents are limited in their knowledge and experience compared to patients, and perhaps anticipate insufficiently to adaptation. The results of this study call for a discussion about if and how adaptation should be compensated for in cost-utility analyses, but first longitudinal research is necessary on the relation between health state utilities and adaptation, before decisions about compensations for adaptation can be made.

7

The influence of time and adaptation on health state valuations in patients with Spinal Cord Injury

Peeters Y., Putter H., Snoek G.J., Sluis T.A.R., Smit C.A.J., Post M.W.M., and Stiggelbout A.M. The influence of time and adaptation on health state valuations in patients with Spinal Cord Injury. Under Revision (Medical Decision Making)



Abstract ***Objective:****One of the explanations for the difference between health state utilities elicited from patients and the public is adaptation. The influence of adaptation on utilities was investigated in patients with spinal cord injury.* ***Methods:*** *Face-to-face interviews were held at three time points (T1 after admission to the Rehabilitation Centre, T2 during active rehabilitation, T3 at least half a year after discharge). At T1, 60 patients were interviewed, 10 patients withdrew at T2 and T3. At all time points patients were asked to value their own health and to value a health state description of rheumatoid arthritis, on a Time Trade-Off (TTO) and Visual Analogue Scale (VAS), and the Barthel Index and the Adjustment Ladder were filled out. Main analyses were performed using Mixed Linear Models taking the time-dependent covariates (Barthel Index and Adjustment) into account.* ***Results:*** *Valuations given on the TTO for the own health changed over time, even after correction for gain in independence* $(F(2, 59) = 8.86, p < 0.001)$. *This change over time was related to adjustment. Both a main effect for adjustment* $(F(1, 87) = 10.05; p = 0.002)$ *and interaction effect between adjustment and time* $(F(1, 41) = 4.10; p = 0.024)$ *were seen for valuations elicited with the TTO. Valuations given for the own health on the VAS did not significantly change over time, nor did the valuations for the hypothetical health state.* ***Conclusion:*** *The effect of psychological adaptation on health state valuations can not be ruled out, but might be less prevalent than previously assumed.*

7.1 Introduction

In cost utility analyses costs of treatment are compared with benefits in health. These benefits in health generated by treatment are measured by health state utilities, which can be given by members of the public or by patients who are actually experiencing the health state. Health state utilities given by patients have generally been found to be higher than utilities given by members of the public.[100] Although this difference is well described in the literature, the origin remains unclear. To decide whose valuations are most valid it is important to understand which underlying mechanisms causes this difference.

One of the explanations for the difference between health state utilities given by patients and members of the public is adaptation. Patients will adapt to their illness whereas members of the public fail to anticipate on this ability to adapt.[24,38] For example Riis et al[107] studied hemodialysis patients and found that these patients reported similar moods as healthy controls which suggests that these patients have adapted to their illness. Moreover, the healthy controls gave significantly lower estimations of their moods imagining living with hemodialyses than the patients did.

Adaptation can be defined as a response that diminishes or remains the same despite constant or increasing stimulus level.[31] Adaptation is a generic term which represents various processes that may take place simultaneously. For instance adaptation can refer to the process in which patients learn new skills to deal with their handicap in activities of daily living.[156] For example patients with spinal cord injury learn how to handle their wheelchair, or how to empty their bladder, but most patients will never return to normal functioning. Nevertheless, patients with SCI report a rather high satisfaction with life.[157] This life satisfaction can partly be explained by rehabilitation but more important predictors are social and psychological functioning.[158] For patients with SCI psychological adaptation thereby appears at least as important as their physical adaptation. Members of the public might anticipate on the ability to learn new skills, but they might fail to anticipate to the psychological adaptation processes, which often operate unconsciously. In the following paragraphs some of these psychological adaptation processes will be discussed.

Psychological adaptation processes can occur through cognitive coping strategies such as positive reframing; patients reframe their situation to see it in a more positive light. Instead of focusing on the losses of having an illness patients might focus on the value of dealing with it.[159] Patients have described that by dealing with an illness they feel stronger, more self-secure, and experience improved relationships with others.[160] Besides positive reframing it is suggested that patients infuse ordinary events with a positive meaning. Tentative support indicates that when patients experience a stressful event they start searching for positive aspects in their lives compensating the negative experience.[159] Furthermore patients might change their goals and adjust their future expectations. A patient who develops weak sight might have previously aimed to become a pilot in an emergency helicopter. After developing the sight problems the patient discovers his ability to listen and instead becomes a counselor at an emergency line. By changing career the patient might forget how important flying was to him and instead learns that reassuring people gives him more fulfillments.

Strongly related to these coping strategies is the so called “psychological immune system”. The psychological immune system has been proposed to consist of defense mechanisms that weakens strong emotions over time.[26] In contrast to coping strategies the psychological immune system is suggested to fail when operating consciously. This restriction will be explained by the following example, based on one of the defense mechanisms used by the psychological immune system. Often people rationalize about their situation in such a way that a negative event might be actually quite positive to them. A patient diagnosed with cancer might rationalize that it is a good thing that he developed cancer instead of his spouse since this spouse already has diabetes. But if a friend reminds the person of his own kidney deficit, which the person conveniently had forgotten, then the repair is undone. Similar to rationalization people might justify previous decisions, such as occurs in preferences for medical treatment.[15,161] Besides rationalization and justification, several other mechanisms are used by the psychological immune system, among others self-enhancement and motivated reasoning.[26]

In studies investigating the gap between health state valuations of patients and of members of the public, the psychological processes of adaptation of patients are often provided as explanation but seldom have been the topic of study. In a study of Jansen and colleagues[26] they indirectly were. The authors assessed changes in utilities over time in a sample of breast cancer patients who received radiotherapy. Patients were asked to rate their own experienced health, a health state scenario of radiotherapy, and a health state scenario of chemotherapy at two time points, before and in the period in which they received radiotherapy. The authors expected patients’ valuations for a radiotherapy scenario to change when patients actually received this therapy. By experiencing radiotherapy patients might have learned that they are able to adapt to it. Change in the health state valuation is seen as an indication for this adaptation process. The results of this study showed however, that patients' ratings for the radiotherapy scenario were similar before and while experiencing this health state, but patients' ratings for their own health during radiotherapy were higher than those for the scenario. This could indicate that patients infused their own situation with positive meaning, or could point to more unconscious adaptation, both of which would not work for the scenario. An alternative explanation provided by the authors was that the experienced radiotherapy was not as severe as was described in the scenario, despite the evidence-based development of the latter.

These authors further assessed valuation shift, a change in valuation of hypothetical health states caused by a change in the own experienced health.[15,162] By adapting to poor health, patients would be able to anticipate on their ability to adapt when they rate hypothetical health states and assign higher ratings. Therefore, besides the radiotherapy scenario, the authors asked patients to rate a scenario of a hypothetical health state, chemotherapy, before and during radiotherapy. No valuation shift was seen for the chemotherapy scenario. The authors suggested that the stability of the valuations before and after radiotherapy might have been influenced by the short term and relative minor side effects of this health state (7 weeks of radiotherapy). Patients might not have adapted sufficiently. Further, it also can be questioned if patients adapt to temporary states[162] The authors noted that in the literature ratings for health state scenarios often remained stable over time whereas most change was shown in studies in which patients rate their own experienced health. It can be questioned if patients are able to project their ability to adapt to a health state (scenario) that they do not "own".[161]

In contrast to Jansen et al.[161]other studies did find evidence for valuation shift.[40,163] Among others, Dolan showed that patients with poor health assigned higher valuations to various EQ-5D scenarios than did patients with good health.[40] The discrepancy between Dolans' findings and the findings of Jansen et al[161] might be explained by the duration and severity of the illness or health state. The influence of adaptation on (hypothetical) health state valuations might only be expected after a certain time, not for temporary states, and, may depend on the severity of the illness. We therefore wished to study these issues in a serious condition with little or no prospect for cure, for which adaptation is a major issue.

To understand the influence of adaptation in valuations given by patients, longitudinal research in a study sample with patients with Spinal Cord Injury (SCI) was performed. SCI leads to permanent long term changes in physical functioning to which people adapt; physically as well as psychologically.[164] First it is examined if valuations for their own experienced health change over time apart from the change expected by the rehabilitation process. Obviously patients with SCI will rehabilitate in the first phase after injury, i.e. they learn new skills to perform activities of daily living independently, which can be expected to result in an improvement in health state valuations. Therefore we will adjust the health state valuations for the improvement in independence. Secondly we examined if the change in health state valuations can be related to adaptation. Finally the experience with a chronic illness on patients' valuations of a hypothetical health state unrelated to SCI will be studied. Does experience with a chronic illness lead to valuation shift?

7.2 Methods

7.2.1 Participants and procedures

Six rehabilitation centers (RC) specialized in spinal cord injury in the Netherlands were involved in inclusion of patients. Patients between 18 and 75 years old who were able to speak and understand Dutch, with acute SCI causing functional losses and problems with daily activities were approached by their treating physician or psychologist. Patients were approached to participate in this study in the first few weeks of admission. Patients with minor functional losses (neither problems with walking ability nor problems with bladder or bowel functions) and patients with severe emotional or cognitive problems were excluded. Eligible patients who gave consent to be interviewed were contacted by one of the interviewers. The interview was scheduled to fit into the patients' rehabilitation schedule.

Patients were interviewed at three times. The first interview took place as soon as possible after admission aimed at at most within the first 4 weeks, except for patients who had not started active rehabilitation. The second interview took place during active rehabilitation aimed at at least two weeks before discharge. The third interview took place at least half a year after discharge. Since the rehabilitation period of patients with tetraplegia is generally longer, the time between interviews differed for patients with paraplegia and patients with tetraplegia. For patients with paraplegia the second interview was aimed at three month after the first interview and the third interview about one year after the first. For patients with tetraplegia the second interview was aimed at about six months after the first interview and the third interview approximately 18 months after the first interview. The medical ethics committee of the LUMC and the local ethics committees approved the study protocol.

7.2.2 The interview

Face-to-face interviews were performed by two trained interviewers following a strict interview protocol. The first and second interview took place at the RC, the third interview took place at home or during an out-patient clinic visit to the RC. The interview was in outline the same at all three time points. In the introduction of the interview, people were instructed that in questions about their health they should not only take illness in account but also the limitations caused by their injury. At the start of the interview, patients answered an open-ended question about their life in the past period. In the first interview we focused on the incident causing the injury and the period between the incident and admission to the RC. This open-ended question was followed by several demographic questions. Next patients gave a valuation of their own health of the previous week using a visual analogue scale (VAS) and a time tradeoff (TTO), followed by a valuation of a rheumatoid arthritis (RA) health state (Appendix C). When patients valued their own health they were asked to imagine that their health would not improve due to rehabilitation and would remain the same as it was in the previous week. In the last part of the interview several questionnaires were interviewer-administrated including the Barthel Index[165] and the adjustment ladder.[166]

7.2.3 Assessments

The Time Tradeoff (TTO)

TTO utilities were elicited with interviewer help using a time-line and board on which descriptions were placed of perfect health and respectively the patients' own health of previous week or the RA health state. Perfect health was described as full well-being in physical, psychological and social functioning. Patients rated how many years (x) of their remaining life expectancy (y), derived from Dutch life expectancy tables for their gender and age category,[117] they were willing to trade to obtain perfect health. The indifference point was searched through the bisection method. Given the severity of SCI and the emotional status of the patients, the lowest tradeoff was set at three months living in perfect health. Utility was calculated as $\frac{y-x}{y}$.

The Visual Analog Scale (VAS)

The VAS is a 100 mm horizontal line ranging from death to perfect health. Perfect health was again described as full well-being in physical, psychological and social functioning. Utility for the own health state of the previous week and for the RA health state were elicited by placing a mark between death and perfect health.

Barthel Index

The Barthel Index (BI)[165] is a measure of performance and rates the degree of independence of help in activities of daily living. The index consists of ten items with a total score between 0 and 20. A higher score indicates more independence. The BI had been examined in a Dutch sample of patients with SCI and found to be a reliable and valid measure which can be used in an interview.[167]

Adjustment ladder

The adjustment ladder is a horizontal ladder ranging from 1, worst possible adjustment to 10, best possible adjustment. Patients named the number which indicated their current overall adjustment.[166] The question was translated in Dutch using a forward and backward translation procedure. Patients were allowed to choose between whole numbers.

7.2.4 Data Analysis

Change over time of utilities for the own health was examined with linear mixed models taking time-dependent covariates into account. For both the TTO and the VAS models were fitted with time included as fixed and random factor. To correct for gain in independence the BI score was included as fixed factor. This implies that the health state valuations given at the three time points were corrected for the BI score at the corresponding time, before we examined the change over time. Using the main effect of time (or the post hoc univariate test when an interaction effect was present), the overall change over the three time points was examined. To examine the change more precisely, the corrected health state valuations (Estimated Marginal Means (EMM)) at T1 and T2 and at T2 and T3 were compared pairwise. If a significant effect of time was present, the model was extended by including the adjustment ladder, as fixed factor. From the significance level of the main effect for adjustment we examined if the change in utilities could be related to conscious adjustment. If conscious adjustment measured by the adjustment ladder totally explained the change in health state valuations over time, we expected the main effect of time to diminish.

Next we studied if experience with SCI changed patients' valuations of a hypothetical health state, in a similar way as for the own health, but not corrected for gain in independence. Mixed linear models for TTO and VAS were fitted with time included as fixed and random factor. Overall change and change between the time points was assessed through the main effect of time and post hoc pairwise comparisons of the utilities. If the effect of time was significant the model was extended with adjustment.

In each model fixed factors were tested on main effect and on interaction with time. Distributions of residuals were checked on normality. Missing values were estimated based on answers given to the remaining questions, given at different time points and given by other patients. In all models T1 was used as reference point.

7.3 Results

In total 74 patients met our inclusion criteria and were approached by their physiatrist or psychologist to take part. Of these patients 13 refused due to personal reasons and one patient was excluded since he found it impossible to answer the TTO given his religion. In total 60 (81%) patients agreed to participate and were interviewed at T1. Of this sample 10 patients had to be excluded at T2; four patients withdrew for personal reasons, one patient could not be contacted, one patient was excluded due to an infection which made active rehabilitation impossible, and four patients were not interviewed because the time between the first interview and discharge had been less than a month. Of the 50 patients who were interviewed at T2 six were interviewed after discharge. For this time point the answers of these six patients were recoded into missing since their adaptation process might have been influenced by being at home. At T3 again 50 patients were interviewed, the four patients who had been excluded for the second interview due to the short time frame were added again, and another four patients had to be excluded (two withdrew due to major pain, one had passed away, and one was not interviewed due to logistic reasons). Table 7.1 and Table 7.2 show the demographic characteristics and mean valuations of these patients.

Table 7.1 Characteristics of people with RA included in this study

	Time 1 ($N = 60$) After admission	Time 2 ($N = 50$)[a] During active rehabilitation	Time 3 ($N = 50$)[a] At least six month after discharge
Gender			
Female	19(32%)	17(34%)	17(34%)
Marital status			
Married	38(63%)	30(60%)	28(56%)
Divorced/ Widow	10 (17%)	10(20%)	12(24%)
Single	12(20%)	10(20%)	10 (20%)
Children			
Yes	36(60%)	30(60%)	29(58%)
Education			
Nine years or less	18(30%)	14(28%)	13(26%)
Between 10 - 12 years	25(42%)	20(40%)	21(42%)
13 years of more	17(28%)	16(32%)	16(32%)
Type injury			
Incomplete paraplegia	19(32%)[b]	15(30%)[b]	18(36%)
Complete paraplegia	19(32%)[b]	19(38%)[b]	19(38%)
Incomplete tetraplegia	20(33%)[b]	14(28%)[b]	11(22%)
Complete tetraplegia	2(3%)[b]	2(4%)[b]	2(4%)
AIS - score[c]			
A	20(33%)	N.A.	N.A.
B	9(15%)	N.A.	N.A.
C	13(22%)	N.A.	N.A.
D	15(25%)	N.A.	N.A.
Cause of injury			
Accident	34(57%)	27(54%)	28(56%)
Illness	14(23%)	12(24%)	12(24%)
Surgery	12(20%)	11(22%)	10(20%)
Help answering VAS[d]			
No help	41(68%)[b]	35(70%)	41(82%)
Help needed	18(30%)	15(30%)	9(18%)

[a] The study sample time point two and three is not exact the same, see Results section; [b] Numbers do not add up to the total number of participants due to missing data; [c] Based on the scores at admission to the RC; [d] Some participants with problems in their upper limps were helped when answering the VAS; N.A. = No information Available

Table 7.2 Continuous characteristics

	Time 1 ($N = 60$) After admission rehabilitation	Time 2 ($N = 50$)[a] During active after discharge	Time 3 ($N = 50$)[a] At least six month
	Mean (SD)	Mean (SD)	Mean (SD)
Age	46(16)	47(16)	46(15)
Weeks after incident			
Paraplegia	11(5.0)		
Tetraplegia	12(5.3)		
Weeks after admission			
Paraplegia	6(3.0)		
Tetraplegia	5(2.5)		
Weeks after Time 1			
Paraplegia		12(5.3)	56(6.1)
Tetraplegia		24(5.7)	85(14.2)
Weeks after discharge			
Paraplegia			30(21.4)
Tetraplegia			39(24.7)
TTO own health state	0.47(0.33)	0.63(0.31)	0.68(0.28)
TTO RA health state	0.54 (0.35)	0.58(0.29)	0.59(0.27)
VAS own health state	61(17.9)	67(15.6)	65(19.3)
VAS RA health state	54(23.5)	54(19.0)	53(21.9)
Barthel Index	9.1(5.9)	12.3(5.4)	13.3(5.2)
Adjustment Ladder	6.2(2.0)	6.9(1.7)	7.3(1.7)

[a] The study sample time point two and three is not exact the same, see Results section.

Table 7.3 Linear mixed model of TTO including time, Barthel Index and Adjustment

	β	SE	p
Intercept	0.18	0.105	0.09
Barthel Index	0.01	0.004	0.03
Adjustment	0.03	0.015	0.04
Time (T2 vs. T1)	0.32	0.131	0.02
Time (T3 vs. T1)	-0.21	0.166	0.21
Time · Adjustment (T2 vs. T1)	-0.03	0.019	0.15
Time · Adjustment (T3 vs. T1)	0.05	0.023	0.04

Valuations on T1 are used as reference point.

7.3.1 Change in health state valuations for the own health and the impact of adaptation

TTO

Results of the mixed linear model examining the effect of time after correction for BI showed both a main effect for BI ($F(1,120) = 9.44, \beta = 0.01(SE = 0.01), p = 0.003$) and a main effect for time ($F(2,59) = 8.86, p < 0.001; beta(T1vsT2) = -0.13(SE = 0.04), p = 0.001; \beta(T1vsT3) = 0.16(SE = 0.05), p = 0.001$). Post hoc pairwise analyses showed that the increase between T2 ($EMM = 0.63, SE = 0.04$) and T3 ($EMM = 0.66, SE = 0.04$) was not significant. Given the significant main effect of time the model was extended with adjustment as fixed factor. Both the main effect for adjustment ($F(1,87) = 10.05; p = 0.002$), and the interaction between adjustment and time ($F(1,41) = 4.10; p = 0.024$) were significantly related to TTO. Table 7.3 shows the results of the linear mixed model including time, BI, and adjustment. The effect of adjustment at T3 was stronger ($\beta = 0.05(SE = 0.02), t(92) = 2.05, p = 0.044$) than the effect of adjustment at T1. After inclusion of adjustment as fixed factor the effect of time remains significant ($F(2,57) = 7.37, p = 0.001$).

VAS

The linear mixed model examining the effect of time on VAS of the own health state corrected for the BI shows only a main effect of BI ($F(1,101) = 13.80, p < 0.001$). After correction for BI no main effect of time was found ($F(2,58) = 1.36(p = 0.27)$). The corrected VAS valuations do not differ significantly between the three time points ($T1 = 63, SE = 2.3; T2 = 66, SE = 2.1; T3 = 63, SE = 2.7$).

7.3.2 Change in patients' valuations of the RA health state

TTO

Using mixed linear model with time as fixed and random effect, no main effect of time on TTO valuations for the hypothetical health state was found ($F(2, 47) = 0.35, p = 0.70$). The estimated marginal means were T1 = 0.54 ($SE = 0.05$), $T2 = 0.57 (SE = 0.04)$ and $T3 = 0.58 (SE = 0.04)$.

VAS

Mixed linear models for VAS valuations of the hypothetical health state with time as random and fixed effect revealed no main effect of time either ($F(2, 48) = 0.08, p = 0.93$). At the three time points the estimated marginal means were $T1 = 54 (SE = 3.1), T2 = 54, (SE = 2.8)$ and $T3 = 53 (SE = 3.1)$.

7.4 Discussion

In studies investigating the gap between health state valuation of patients and of members of the public psychological adaptation is often provided as explanation but seldom has been studied. In this paper provisional support for the effect of adaptation on health state valuations has been found. Health state valuations for the own health given on a TTO changed over time even after correction for the gain in independence. This increase could partly be explained by conscious adjustment. However, even after correction for conscious adjustment the valuations on the TTO still increased.

Strongest increase in the valuations for the own health given on the TTO (after correction for improvement in independence) appeared during in patient rehabilitation. Patients adapt to their injury most in the initial phase after they were admitted to the RC. In the period after discharge when most patients returned home, not much improvement was reported. This is in line with previous findings. Van Koppenhagen et al.[164] found that in the first period during inpatient rehabilitation life satisfaction of patients with SCI changed with the strongest change in the first three months. Whereas in the second year after injury when patients are discharged from the rehabilitation centre only minor[168] or no change[169] in life satisfaction is reported. Moreover patients with injuries developed several years ago reported even a moderate decline over a period of nine years.[170]

The increase in valuations for the own health on the TTO could partly be explained by self-reported, conscious, adjustment. After correction for conscious adjustment the valuations still increased over time. Adaptation measured by a single estimation of the overall adjustment by patients themselves, as used in this study, does not explain the increase in valuations sufficiently. Several explanations can be suggested for this finding. First, adaptation measured by the adjustment ladder might include physical as well as psychological adjustment. However it can be questioned if patients are able to reflect on their unconscious adaptation that is enabled through their psychological immune system.[26] Patients might be aware of their adaptation, but not of the underlying unconscious processes creating it. Secondly, by using a single item estimator with an open description of adjustment we can only speculate which aspects of adjustment patients have taken in account. Third, the change can be related to social aspects or prevalence of secondary problems.[157] Although, this assumption only holds if social aspects and prevalence of secondary problems are also related to the time since injury.

The effect of adaptation on health state valuations for the own health on the TTO was not equal at all time points. Interestingly compared to the effect of adaptation on the valuation at T3, the valuation for the own health at T2 was almost not influenced by adaptation. Possibly patients found it difficult to estimate their overall adjustment at T2 because they were thinking about their discharge. In the period before discharge patients are occupied with the last steps before they are going home, such as arranging devices necessary for their daily activities. Further, they are aware that they soon have to change the relative secure surroundings of the RC for their home where no immediate help is available. During the interview patients therefore often made remarks that they thought they were adjusted to their SCI while being in the RC, but they were not sure about their adjustment when returning home.

In contrast to the findings on the TTO the valuations given on the VAS did not change over time, whereas actually change on the VAS would have been expected.[171] This finding might provide insight in the effect of adaptation on health state valuations. In the TTO patients focus on time, comparing their quality of life with length of life. Some psychological adaptation processes are time dependent and might influence on how much time a patient is willing to trade. For example if patients are able to change their future expectations and goals they will have something “new” to live for, resulting in fewer years they are willing to trade. Schwartz et al.[171]found such a relation between life goals and willingness to trade.

Nevertheless the finding that health state valuations of the own health on the VAS did not increase is noteworthy. The VAS can be described as an affective measure which is suggested to be more sensitive to change than cognitive measures.[172] However this suggested sensitivity of the VAS is based on patients diagnosed with cancer or with RA, whereas in this study patients with SCI were interviewed in the initial phase after injury. SCI is a serious condition with little or no prospect for cure. It can be questioned if findings based on patients with cancer or RA can be generalized to our study sample.

To examine valuation shift patients were asked to value a RA health state. Experience with SCI did not change how patients valued this hypothetical health state. The valuations for the RA health state on the TTO and VAS did not increase over time. These findings are similar to those described by Jansen et al.[20] In contrast to what was expected, the long-term adjustment to SCI did not result in an increase in valuations of a hypothetical health state. Probably patients are not able to project their ability to adapt on a hypothetical health state. This failure to anticipate on ability to adapt has been described before in members of the public.[24,39] This finding also gives support for previous findings which showed that health state valuations of patients and members of the public are more similar when both groups give valuations to hypothetical health states.[100]

From previous research it followed that the effect of adaptation on health state valuations could best be examined in patients with a severe chronic illness.[20,40] Therefore patients with SCI were included, since SCI is a serious condition with little or no prospect for cure. This entailed some disadvantages, however. The inpatient rehabilitation process of this study sample was unpredictable. Where some patients had a rehabilitation period of several months or even a year, other patients were allowed to go home after only a few weeks. Since this variation was only partly related to the level of injury it was not always possible to keep track of discharge. As a result the time point T2 was less consistent than we aimed for. To minimize this variation in the time, T2 interviews that took place after discharge were coded as missing. Despite this variation in time the findings described in this study seem substantial. Including only patients who were interviewed at the aimed time points did not change our findings.

Finally, we want to focus on an unexpected secondary finding. TTO valuations were lower than VAS valuations in the first interview Table 7.2. In general, VAS valuations have been found to be lower than TTO valuations. Only for very severe health states valuations on the VAS have been found to be higher than TTO valuations.[173] It is proposed that for severe health states patients find it unthinkable to live in it for several years, whereas this focus on time is less prominent in the VAS. This phenomenon, entitled "maximum endurable time"[174] is reflected in comments of patients related to the length of time. For example one patient said: "not 21 years like this, sometimes I hope when I close my eyes that I will not wake up again". Although patients were allowed to adjust the number of expected years in the TTO when they found the described situation unbearable, only two patients chose to do so.

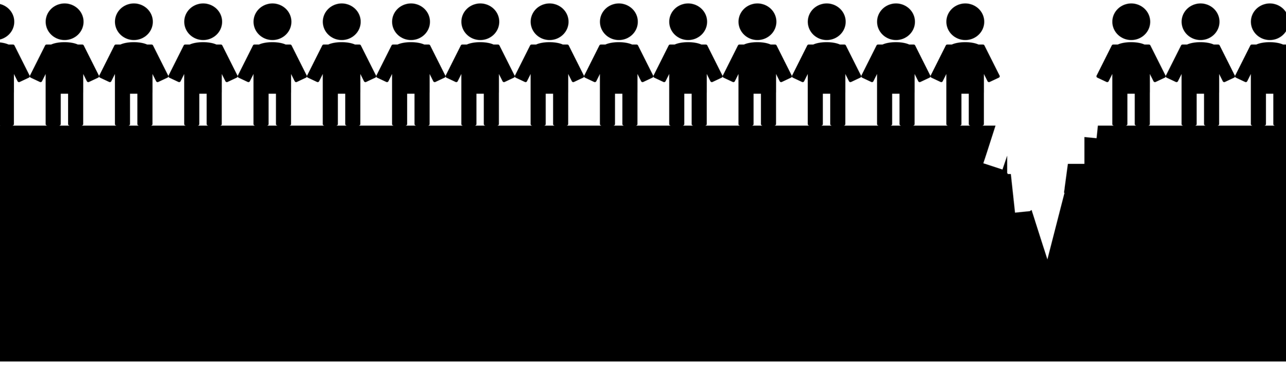
7.5 Conclusion

Valuations for the own health state given on the TTO seem to be influenced by psychological adaptation over time, but for the VAS no such effect was found. Experience with a chronic illness did not result in a change in valuation of a hypothetical health state. The effect of psychological adaptation on health state valuations can not be ruled out, but it might be less prevalent than previously assumed.

8

After adversity strikes: Predictions, recollections and reality among people experiencing the onset of adverse circumstances

Peeters Y., Smith D.M., Loewenstein G., Ubel P.A. After adversity strikes: Predictions, recollections and reality among people experiencing the onset of adverse circumstances. Conditionally Accepted (Journal of Happiness Studies)



Abstract *Numerous studies on affective forecasting have demonstrated that people frequently underestimate their ability to adapt to adverse circumstances. But to date, these studies have not assessed people's affective forecasts early in the experience of these new circumstances. We present two longitudinal studies of people experiencing new adversities. In the first study 54 patients experiencing new limb amputations were recruited to participate in a mailed survey. Patients assessed their well-being, functioning and general health (1) two weeks after discharge from the hospital and (2) three months later. At the first time point patients also predicted their well-being, functioning and general health at three months. In the second study 55 patients experiencing new colostomies were recruited and received mailed surveys at three time points; (1) at baseline (within one week after leaving the hospital), (2) one month after baseline, and (3) seven months after baseline. Again we assessed their actual and predicted well-being, functioning and general health. In both studies the actual change was compared to the change expected by patients. Across both studies, patients expected to significantly improve on all three domains but reported little actual improvement. Together, these studies demonstrated that people with new disabilities overestimate hedonic adaptation-they expect their overall well-being to improve more than it actually does.*

8.1 Introduction

In order to make good decisions, people need to imagine how their future well being will or will not be affected by their choices. For example, if confronted with a choice of whether to move across the country to accept a better paying job, a person would need to predict how her happiness will be affected by the increased income, the new location, leaving friends behind, etc. To a large extent, making the best choice in this type of circumstance depends on making accurate predictions. However, numerous studies on affective forecasting have demonstrated that people frequently mispredict their long-term emotional responses to events.[26, 175, 176] People typically overestimate the duration and affective impact of negative life events, assuming that major events will have enduring emotional consequences, while underestimating their ability to adapt to such circumstances.[26, 176] For instance, people imagine that chronic illness and disability will have a sustained impact on their wellbeing, whereas people experiencing such problems often report high levels of wellbeing.[18, 177]

The general thrust of research on affective forecasting implies that when people experience new adversities, they underestimate the speed and thoroughness of hedonic adaptation-feeling miserable because of their new circumstances, they imagine that such misery will be deeper and longer lasting than it will actually be.[178] However, research on affective forecasting has not, yet, investigated people's beliefs about adaptation when they are early in the experience of a new adversity. Instead, many studies have been cross-sectional,[26, 177, 179] comparing people's naive predictions to the reported experiences of people in the circumstance in question. But this type of cross-sectional design does not allow us to determine whether people who are newly experiencing adversity fail to properly consider hedonic adaptation in forecasting their own happiness; they might be miserable now, but do they expect to remain miserable?

Other studies have employed longitudinal designs, but do not capture people's experience and predictions early in the course of adapting to the adversity, and then compare those to experiences after adaptation has had a chance occur. For instance,[26] assessed people's predictions of their long-term emotional reaction to the outcome of a political election and found that people who supported the candidate who lost the election expected to experience stronger negative emotions than they actually did. Similarly, patients waiting for a renal transplant expected to experience a greater increase in well being than they actually did.[180] These prospective designs established that immune neglect and hedonic adaptation are powerful phenomena.

But they did not provide an opportunity to see what people predict early in the process of adapting to adversity.

The aim of this study was to investigate if “newbies”-people early in the experience of a new adversity-are prone to the same kind of forecasting errors as have been demonstrated in previous research. Using a longitudinal design, we followed recently disabled patients to compare their predictions (about how much they would adapt to their condition) to their actual experience of adaptation over time. Based on previous research in affective forecasting, we examined three competing hypotheses-that people new to disabilities would a) underestimate their ability to adapt over time, resulting in predictions of well being that are biased low, b) accurately predict adaptation, and c) overestimate adaptation. We elaborate on each of these hypotheses in the following paragraphs.

There are reasons to think that these newbies will underestimate adaptation. First, as reviewed above, such underestimation is a wide spread phenomenon, having been demonstrated for short-term minor events like the outcomes of football games,[27] more significant phenomenon like people’s beliefs about how long they will be influenced by a move to a different climate,[181] and serious chronic adversities, like spinal cord injuries and divorce.[37] Second, early in the experience of a new adversity, many people experience strong negative emotions. It is plausible that it would be difficult for them to therefore imagine themselves with weaker emotions, due to what Loewenstein calls a hot/cold empathy gap.[182]

On the other hand, there are members to think that newbies may accurately predict adaptation. First, having begun to experience the new adversity, they may already have new insight into the speed and thoroughness of hedonic adaptation. With their psychological immune systems already in high gear, they may be more able to imagine the long term trajectory of their emotions. Second, under normal circumstances people typically believe that happiness will increase in the short run-that they will be happier several months or years from now than they are now. People expect positive events in their own future even when there is no supportive evidence for it.[108] Similarly, people with chronic or terminal illnesses maintain positive beliefs about their future health despite their health problems.[45]

Finally, there are reasons to think that newbies will actually overestimate adaptation. Some adversities, like new health problems, may create realistic hope for improvement in health related domains, and people might mistakenly assume that these improvements will be accompanied by similar improvements in well-being.[180] For instance, people undergoing below-the-knee amputations must recover from arduous surgeries, and must then undergo taxing physical therapy regimes. While these people cannot expect to get their lower legs back, they can expect to experience improvement in physical functioning in the months following their amputation. Will they overgeneralize from their beliefs about physical functioning, and therefore mispredict how much their overall quality of life will also improve?

8.2 Study 1: Predicting physical functioning, general health and well-being after amputation surgery

8.2.1 Overview

In study 1, we report on a longitudinal survey of patients undergoing limb amputations, in which we assessed their physical function, general health and well-being by mailing surveys to them at baseline (two weeks after discharge) and three months later. At baseline, we also asked patients to predict what their physical functioning, general health and well-being would be three months later. With this design, we were able to assess the accuracy of people's predictions across these three domains.

8.2.2 Participants

We recruited patients at the University of Michigan Medical Center who underwent a major single limb amputation. We excluded people who had had previous limb amputations, were suffering from dementia, were terminally ill, or could not understand written English. We contacted 69 patients while still in the hospital recovering from the surgery. 54 (78%) agreed to participate in our longitudinal study. Participants were paid $40 for each completed survey.

8.2.3 Study measurements

Well-being: We assessed life satisfaction by asking patients how much they agreed with the statement "I am satisfied with my life," on a scale ranging from 1 (strongly disagree) to 7 (strongly agree).[183] We also asked patients how often they felt "calm and peaceful", "energetic" and "depressed", on a scale ranging from 1 (none of the time) to 5 (all of the time).[143] We then created a composite measure of well-being by averaging scores across these four measures (Cronbach's $\alpha = 0.75$). Physical

Table 8.1 Patient Characteristics of patients who had amputation surgery

Three month survey	($N = 37$)
Age (Mean (SD))	55 (12)
Gender	
Female	17(49%)
Race	
Non White	5(14%)
Marital status	
Married	15(43%)
Divorced/ Widow	14 (40%)
Single	5 (14%)
Cause amputation	
Something sudden	10 (29%)

functioning: We assessed three aspects of physical functioning: (1) “satisfaction with current level of physical functioning” on a scale ranging from 1 (very dissatisfied) to 5 (very satisfied); (2) “engagement in social activities outside the home such as visiting friends, neighbors and relatives”, on a seven point scale ranging from 1 (never) to 7 (very frequently); and (3) “social activities inside the house such as talking on the phone, having someone over for a visit”, on the same seven point scale. Cronbach’s α for a composite of these three measures was 0.53. General health: To assess general health, we utilized the first item of the MOS 36-item short form health survey which assesses self-reported general health on a scale ranging from 1 (poor health) to 5 (excellent health).[143]

8.2.4 Results

41 (76%) patients returned the first written survey and the three month survey. The most common reasons for non response were moving, death, and voluntary withdrawal. Four patients had additional amputations during the survey period and were excluded. The demographic characteristics of the 37 patients included in this study are shown in Table 8.1.

Table 8.2 shows patients’ baseline ratings for well-being, physical functioning, and general health, their predictions for how these three domains would change at three months, and their actual outcomes at the three month time point. Students’ t-tests were conducted to compare these ratings. As can be seen, patients’ self-reported well-being did not increase over time ($t(31) = 0.05, p = 0.96$, effect size

= 0.01) whereas they expected a significant improvement of approximately seven points ($t(31) = 3.64, p = 0.001$, effect size = 0.64). Also, on functioning and on general health they reported no significant improvement (functioning: $t(34) = 0.42, p =$ 68, effect size = 0.07; general health: $t(34) = -1.72, p = 0.10$, effect size = 0.29) even though they had expected to improve (functioning: $t(34) = 4.17, p < 0.001$, effect size = 0.70; general health: $t(34) = 2.97, p = 0.005$, effect size= 0.50).

In general, patients anticipated significant improvement across all three domains, but did not experience any significant improvements (and in fact experienced a decline in self-reported health of borderline statistical significance). Rather than underestimate adaptation, then, these patients overestimated it-they anticipated an improvement in well-being that did not arise.

8.2.5 Discussion

Rather than underestimate adaptation, the patients in study 1 overestimated how much their well-being, physical functioning, and general health would improve in the months following their amputation. In the introduction, we discussed several factors that could cause people to overestimate adaptation to adversity. We suggested that people might overgeneralize when making predictions: anticipating that they would experience improvements in physical functioning they might, therefore, overestimate how much their sense of well-being would also improve. In Study 1, however, such an overgeneralization does not account for such mispredictions, because these patients did not, by patient self-report, experience significant improvements in general health or physical functioning over this time period. Instead, they overestimated how much these domains would improve too.

Why did people with new amputations overestimate improvement across all three domains? One possibility is that the baseline measure, completed several weeks after the amputation, took place after significant adaptation had already occurred. They might have experienced several weeks of significant improvement, and mistakenly assumed that they would continue to experience similar improvements. To further complicate matters, patients with amputations are often plagued by many other chronic, even progressive, illnesses, like vascular disease and diabetes. Having begun to recover from their amputations, they may mistakenly imagine their health improving over the next three months, while overlooking the likelihood that they will experience new medical problems. Indeed, four patients were removed from our analyses because they required additional amputations during the three month follow-up period. Patients focused too narrowly on the likelihood that their recently

Table 8.2 Predicted and actual valuations of patients with Amputation Surgery

	N	Baseline actual Mean (SD)	3 Month prediction Mean (SD)	3 Month actual Mean (SD)	$B - P^a$ t-value	$B - 3M^b$ t-value	$P - 3M^c$ t-value
Well-Being	32	56.2 (19.6)	63.4 (17.3)	56.1(21.5)	$3.64^\sharp$	.05	$2.28^\sharp$
Functioning	35	53.0 (18.1)	65.5 (16.4)	54.2 (19.8)	$4.17^\sharp$	.42	$3.39^\sharp$
General Health	35	41.4 (26.4)	50.0 (25.0)	35.0 (25.2)	$2.70^\sharp$	-1.72^*	$3.75^\sharp$

Ratings of patients with amputation surgery for their actual and predicted well-being, functioning and general health 6 weeks and 3 months after surgery. Actual change is compared with expected change using students' t-test.

[a] Baseline Actual - 3 Month Prediction

[b] Baseline Actual - 3 Month Actual

[c] 3 Month Prediction - 3 Month Actual

$^\sharp p < .001, ^* p =< .10$

amputated limb would improve, they might have underestimated the chance that other problems would arise.

We address some of these complicating issues in our second study including people with a different health problem: patients undergoing surgery to have a colostomy. First, we not only asked these patients to predict how their lives would change after the surgery, but also asked them to reflect back on how their lives had in fact changed at later time points. Using this method, of both assessing predictions and recollections, we can more thoroughly test whether people have theories about how these different life domains ought to change over time, and whether these theories are accurate.[42] Thus, for example, a patient might assume at baseline that both his health and well-being will improve in the next six months. If his health then declines however, due potentially to unforeseeable events, our method allows us to test with the recall measure whether he recognizes this decline in health or, instead, whether his theory about how his health has changed will trump his actual experience. Second, we mailed our first survey within several days to one week after patients were discharged from the hospital following their surgery, thus capturing earlier experiences and predictions than we captured in Study 1.

8.3 Study 2: Predicting and recalling well-being, physical functioning, and general health after colostomy surgery

8.3.1 Overview

In Study 2, we recruited patients undergoing colostomy surgery at the University of Michigan Medical Center. We assessed their physical functioning, general health and well-being at three time points: (1) baseline (within one week of leaving the hospital), (2) one month after baseline, and (3) seven months after baseline. At baseline, we also asked people to make predictions about their lives one month later. And at one month, we had them make predictions about their lives at the seven month period, while also asking them to recall their physical functioning, health and well-being at baseline. Finally, at seven months, we asked people to recall how they stood on these three domains at the one month time point.

Table 8.3 Patient Characteristics of Patients who had Colostomy Surgery

	One month after release ($N = 55$)	Six month after release ($N = 34$)
Age (Mean (SD))	51 (14)	54(14)
Gender		
Female	21(39%)	15(44%)
Race		
Non White	7(13%)	6(18%)
Marital status		
Married	37(68%)	24(71%)
Divorced/ Widow	9 (16%)	6(17%)
Single	9 (16%)	4(12%)
Colostomy/Ileostomy supposed reversed		
Yes	25(48%)	14(44%)
Cause colostomy/ileostomy		
Inflammatory bowel disease	18 (33%)	12 (35%)
Familial adenomatous polypsis	2 (4%)	1 (3%)
Cancer	21 (38%)	15 (44%)
Trauma/accident	2 (4%)	1 (3%)
Spinal Cord Injury	2 (4%)	1 (3%)
Other cause	14 (26%)	9 (27%)
More than one reason listed	4 (7%)	5 (15%)

8.3.2 Participants

107 patients at the University of Michigan Medical Center who had either a colostomy surgery were recruited shortly after their surgery. Out of these 107 patients 11 patients were excluded because they could not speak English or had poor health. Participants were paid $40 for each completed survey.

In total 76 (79%) of the 96 patients agreed to participate and returned the first survey by mail. Of these 76 patients, 3 had their colostomy reversed between the first and second measurement and 14 between the second and third measurement. Only patients who did not have their colostomy reversed during the study period were included in analyses. Table 8.3 presents the demographic information of the remaining patients.

8.3.3 Study design and measurements

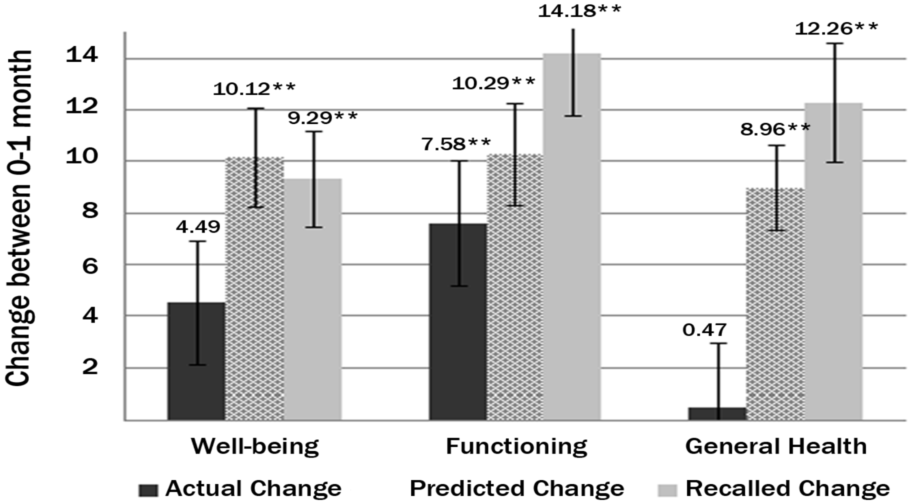
General Health and Physical functioning (Cronbach's $\alpha = 0.58$) were measured in the same way as in Study 1. We added a Quality of Life rating to the measures of well-being (on a scale ranging from 0 to 100, where 0 represents the worst imaginable quality of life and 100 represent the best imaginable quality of life). $(Cronbach's\alpha = 0.75)$. When people receive colostomy surgery, the colostomy can be intended to be permanent or temporary, creating two subgroups of colostomy patients with different ultimate outcomes.[184] All of the following analyses focused on patients who still had their colostomy at the time of assessment, even though some still expected to get their colostomy reversed in the future. We also included this variable-permanent versus temporary colostomy-in analyses and checked for any interactive effects. We did not find any significant or near-significant ($p < 0.10$) interactions, and therefore combined the data across these two groups of patients.

8.3.4 Results

Figure 8.1 illustrates changes in well-being, physical functioning and general health from baseline to one month, and contrasts these actual changes with predicted changes (how much people thought at baseline that these domains would change over that time), and recalled changes (how much people thought, at one month, that those domains had changed). Students' $t - tests$ were conducted to test the significance of these changes.

From baseline to one month, people's overall well-being increased by approximately four points, ($t(37) = 1.86; p = 0.07$; effect size = 0.27), an almost statistically significant improvement, but one that paled in comparison to people's expectations (with people predicting approximately a ten point increase ($t(37) = 5.26; p < 0.001$; effect size = 0.66)) and also compared to their recollections (with people recalling approximately a nine point increase ($t(37) = 4.96; p < 0.001$; effect size = 0.48)). A similar pattern emerges for the other domains. The patients did experience significant improvement in physical functioning, ($t(50) = 3.10; p = 0.003$; effect size = 0.39), approximately what they predicted , ($t(50) = 1.27; p > 0.05$; effect size = 0.14), but significantly less than what they recalled ($t(50) = 2.23; p = 0.03$; effect size = 0.32). This pattern was even more dramatic for measures of general health, which did not change significantly from baseline ($t(52) = 0.19; p > 0.05$; effect size = 0.02) despite people both predicting that it would change ($t(52) = 5.39; p < 0.001$; effect size = 0.34) and remembering that it had changed ($t(52) =$

Figure 8.1 Actual, predicted and recalled mean change in well-being, functioning and general health between 0 and 1 month



Actual, predicted and recalled change on well-being, functioning and general health reported by patients with colostomy surgery within one week of leaving the hospital and 1 month.

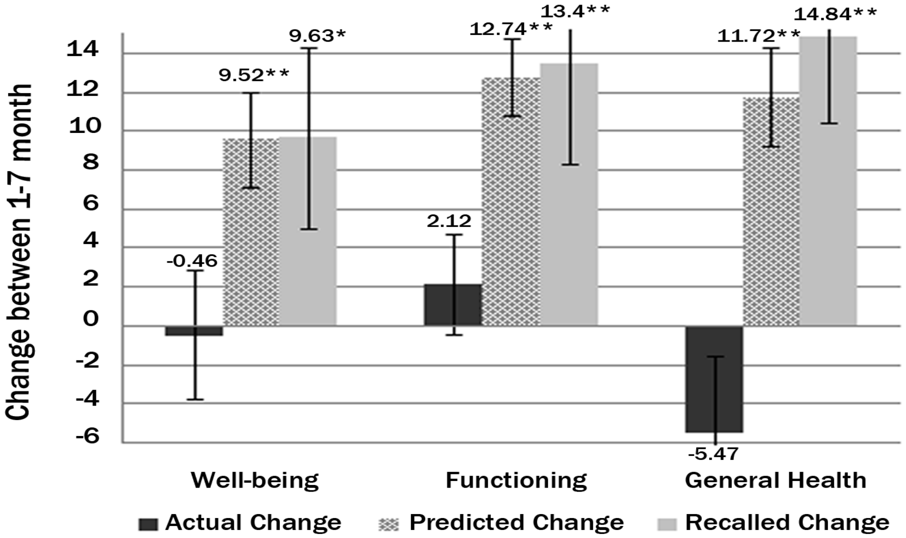
⊤

⊥95% Confidence Interval **p < .001, *$p \leq .05$

$5.34, p < 0.001$; effect size = 0.50).

Figure 8.2 illustrates the actual changes patients experience from one to six months, as well as their beliefs about these changes. Again students't-tests were conducted to test the significance of these changes. For space reasons, and because they were substantively similar to the baseline/one month comparisons, we briefly summarize these results. Once again, the data demonstrate striking disparities between actual experience and belief. And once again, the main error people make is to expect (and remember) more improvement than they actually experience.

Figure 8.2 Actual, predicted and recalled mean change in well-being, functioning and general health between 1 and 7 months after release



Actual, predicted and recalled change on well-being, functioning and general health reported by patients with colostomy surgery between one month and 7 month after leaving the hospital.

⊤
⊥95% Confidence Interval
* $p < .001$, *$p \leq .05$

8.4 Discussion

Across two very different health conditions, we discovered that people newly experiencing a serious adversity overestimated their own hedonic adaptation; they expected their overall sense of well-being to improve more than it actually did. In addition, they overestimated how much their general health and physical functioning would improve over the same time period. Finally, when asked to recall changes over these same time periods, people “remembered” experiencing substantial improvements in all three domains; their recollections, like their expectations, indicated substantial overestimatation of adaptation. The patients’ apparent belief that they would quickly thrive in the face of adversity stands in contrast to prior research, which has shown that people underestimate their ability to adapt to a wide range of adverse circumstances.

Why did the mispredictions in this study run in the opposite direction of those found in so many other studies? One possible factor that could account for the differences is that, in contrast to other studies that have elicited estimates of adaptation before individuals experienced adverse events, our study assessed people early in their experience of the new circumstance. This raises the possibility that mispredictions differ depending on whether one is viewing circumstances completely from the inside or partially from the outside. When healthy people imagine life with a colostomy, for example, they recognize that life with normal bowel function is better than life with a colostomy, and theorize that these differences must therefore significantly influence overall well-being.[185] By contrast, people with a new colostomy, when imagining their well-being over the next six months, are imagining life from the inside. They are still imagining themselves as someone with a colostomy, and might therefore tap into different theories about how their well-being will change over time, theories about how emotions, in general, change over short periods of life, or theories about the likelihood that early improvements in physical function or well-being will persist, and will have large, positive effects on overall well being. As described in the introduction, people expect positive events in their own future even when there is no supportive evidence for it.[42]

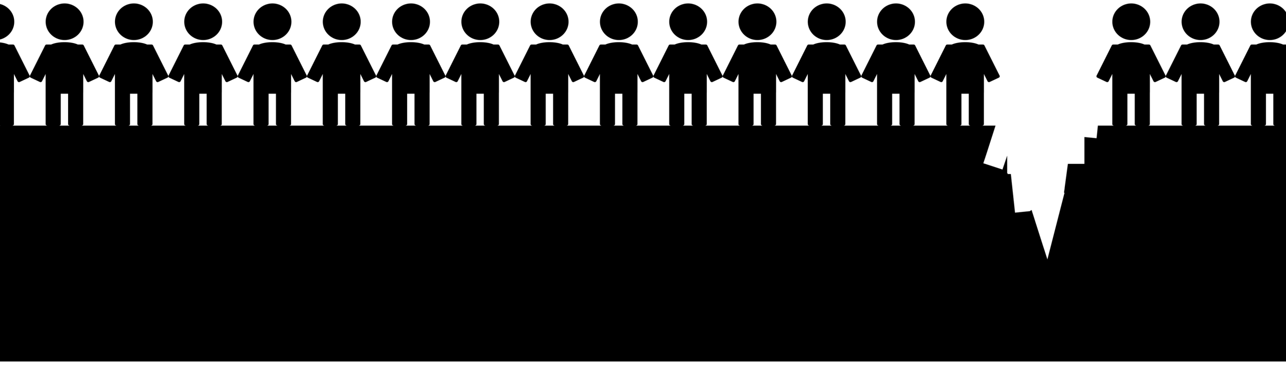
In the studies described, patients predicted on average that their general health would improve over time, and yet they did not as a group report such improvements. It is possible that the patients in our studies simply did not experience the kind of health improvements that they expected to. These mispredictions could have contributed to their affective forecasting errors. But we favor an alternative explanation-that the lack of improvement in general health seen in our studies reflects the subjective nature of our health measures, which relied on patient self-report. For example, new amputees, recently home from a stay in the hospital, may have considered their health to be relatively good compared to what it had been immediately after their operation. One month later in our follow up survey, patients might have reported a decline in health even though their objective health was stable, because they now judged their health relative to different standards. Our data cannot determine whether this kind of scale recalibration occurred. But in support of this theory, the patients in study 2 demonstrated recall bias not only in measures of self-reported well-being, but also in measures of self-reported health. Our health measures, in other words, behaved similarly to our measures of well-being.

Our findings add nuance to the story researchers have been developing about hedonic adaptation. In early studies, researchers established the surprising frequency and intensity of adaptation. People's emotions were shown to be relatively resistant to even substantial changes in their circumstances[33] and people often underestimate the extent of their hedonic adaptation .[178] More recently, researchers have uncovered more subtle findings about adaptation and affective forecasting. Adaptation is not as universal as experts once believed, nor as complete.[186] Individual differences, too, have been shown to influence people's ability to adapt to specific circumstances.[187] Our research adds yet another twist to the plot. We have shown that, at least in some circumstances, people shift from underestimating adaptation to overestimating it. Future research is needed to elucidate when people are prone to making these different kinds of mispredictions.

9

Abandoning the language of “response shift”a plea for conceptual clarity in distinguishing scale recalibration from true changes in quality of life

Ubel, P.A., Peeters, Y., & Smith, D.M. (2010). Abandoning the language of “response shift”: a plea for conceptual clarity in distinguishing scale recalibration from true changes in quality of life. Quality of Life Research, 19, 465-471.



Abstract *Quality of life researchers have been studying ”response shift” for a decade now, in an effort to clarify how best to measure QoL over time and across changing circumstances. However, we contend that this line of research has been impeded by conceptual confusion created by the term ”response shift”, that lumps together sources of measurement error (e.g., scale recalibration) with true causes of changing QoL (e.g., hedonic adaptation). We propose abandoning the term response shift, in favor of less ambiguous terms, like scale recalibration and adaptation.*

9.1 Introduction

In 1999, Social Science and Medicine published a series of articles on the methodological importance of understanding response shift in quality of life research.[41, 188, 188–192] Since that time, researchers have published more than 100 studies exploring various aspects of response shift.[193] As a consequence of this work, quality of life researchers have becoming increasingly aware of issues relevant to the measurement of quality of life over time. Response shift experts have drawn researchers' attention to sources of bias in quality of life estimates. They have shed light on important mysteries relevant to understanding the experience of people with chronic illness and disability.[194] And they have focused researchers on the challenge of explaining why people with disabilities often provide quality of life reports that seem to belie their objective circumstances.[18]

But it is time to abandon the term "response shift." As we explain below, the term response shift is currently being used to lump together distinct phenomena that often have very different implications for the accuracy of quality of life measurement. Moreover, the specific term "response shift" has misleading connotations. The term suggests that the high quality of life reported by many people with chronic illness and disability are measurement artifacts-their "responses" have "shifted"-and that such people are not really experiencing high quality of life. We think such connotations, even if not originally intended, are misleading. A major goal of QoL measurement is to discern to what extent changes in QoL reports over time represent true changes in QoL and to what extent they reflect measurement error. Unfortunately, the response shift literature often fails to make this important distinction. At times, the term response shift is equated with measurement error, while at other times, the term is used to characterize a mechanism by which people's true QoL changes. In this paper, we respectfully contend that the field of response shift research has been characterized from its outset by conceptual confusion. We propose that QoL researchers should abandon the term response shift and focus instead on developing ways to disentangle measurement error-specifically scale recalibration-from true changes in QoL.

9.2 Two examples of response shift: hypothetical case studies

To illustrate our central concern about the term response shift, consider the following two hypothetical case studies. We use these oversimplified cases to elucidate our distinction between measurement error and true change.

Happiness after paraplegia

The first case involves a man who develops paraplegia as a result of an accident. Early on, he is emotionally devastated by his disability. But over time, he begins to recover. In part, his recovery is aided with the help of physical therapists and occupational therapists, who teach him how to engage in important activities, like getting in and out of a wheelchair and driving a car. Yet despite these advances, several months after the accident, he remains unhappy.

But after many months, his mood improves. He no longer spends time focusing on what he cannot do. Instead, he shifts his attention toward new goals, such as participating in wheelchair basketball tournaments. He even gets more involved in church, reestablishing a spiritual life that had slipped away from him before his accident. Even though his physical function is stable by all objective measures, his emotional response to his disability has abated. Eventually, his mood is close to what it was prior to the accident. Indeed, for the purposes of our discussion, let us assume that emotion researchers have videotaped this man's facial expressions over time and confirmed that in the early months after his injury, he rarely smiled and frequently frowned, but over time his face revealed an increasing percentage of time experiencing positive rather than negative emotions.[195] Let us suppose that QoL researchers have been surveying this man over time. As part of a multi-dimensional QoL scale, they ask him to provide a global report of his overall QoL, on a 0-100 scale. Several months after his accident, he rates his QoL as 36 out of 100, a very low score. By eighteen months, his self-reported QoL has risen to 67 out of 100.

A person with chronic pain who experiences kidney stones

In our second case, we ask you to imagine a thirty-year-old woman who suffered a leg wound while serving in the armed forces and has experienced chronic leg pain ever since. The pain interferes with her sleep and makes it difficult for her to concentrate. She rates the pain on average as being 7 out of 10 most days. Then, this woman experiences kidney stones, with pain significantly more intense than anything she has previously experienced, 10 out of 10. Indeed, she cannot believe that she thought her leg pain qualified for a score as high as 7 out of 10.

After her kidney stones are treated, her life is unchanged. Her leg pain continues unabated. A researcher monitoring her facial expressions records that she exhibits just as many grimaces of pain now as she did before her kidney stones. Her leg pain, in other words, is exactly the same. But now, her interpretation of this 0 to 10 pain scale has changed. She has a very different idea about what a pain score of 10 means. Therefore, when asked to rate her leg pain now, she replies that it averages about 5 out of 10.

9.2.1 Viewing these case studies through the lens of response shift

These two hypothetical studies carry quite different implications for the measurement of subjective experiences like QoL and pain. In the first case, QoL measurement appears to have accomplished exactly what researchers want it to accomplish-it has captured the true change in QoL that this person experienced when he emotionally adapted to his chronic disability.[35] In the second case, however, a person with stable pain reported a change in her pain score. And yet her pain had not changed. In this case, the pain measure failed to provide us with a valid way of comparing this person's pain over time.

In the first case, a person emotionally adapted to a chronic disability and thus reported a change in QoL. In the second case, a person recalibrated the pain scale and thus, despite experiencing stable pain, reported a changing pain score. Two very different phenomena. And yet both stand as examples of response shift. Admittedly, these two cases are not only hypothetical, but also relatively simplistic, lacking the complexity of actual patient trajectories. In addition, both cases make reference to unidimensional measures (of QoL and pain, respectively), whereas many actual measures are multidimensional. Nevertheless, these cases are meant to illustrate our main point: that response shift lumps together quite distinct, and potentially distinguishable, phenomena, ones that seem better off separate than lumped together. (We expand on this argument below.) In addition, the issues we raise with these hypothetical cases are not limited to unidimensional outcome measures. Indeed, many multidimensional measures contain items within that are susceptible to the same phenomena illustrated above, phenomena like adaptation and recalibration. Our concern is that these two cases came to be lumped together under one rubric, of ”response shift.” To understand how this lumping has occurred, we need to look more closely at how experts have defined response shift.

9.3 Defining response shift

Sprangers and Schwartz define response shift as: A change in the meaning of one’s self evaluation of a target construct as a result of: (a) a change in the respondent’s internal standards of measurement (scale recalibration, in psychometric terms); (b) a change in the respondent’s values (i.e. the importance of component domains constituting the target construct); or (c) a redefinition of the target construct (i.e. reconceptualization).[41] They pack a great deal of meaning into this description, so it is worth unpacking their definition.

9.3.1 Scale recalibration

The first component of response shift, scale recalibration, occurs when changing circumstances cause people to change how they interpret a subjective response scale. In our hypothetical example, the woman with chronic leg pain exhibited scale recalibration-her experience of kidney stones changed how she interpreted the 0 to 10 pain scale. Scale recalibration is a threat to the validity of self-reports. For example, imagine an 85-year-old man who rates his health as 90 out of 100 on a scale where 100 is defined as “perfect health.” Imagine a 35-year-old man who also rates his health as 90 out of 100. How confident would you be that the two men mean the same thing by 90 out of 100? In a survey of people 50 or older in the United States, we found evidence for scale recalibration in how people interpret the phrase “perfect health.”[177] We discovered that as people get older, they redefine what perfect health means. This makes it difficult to compare subjective health ratings across people of different age groups. By similar logic, it is possible that people who experience chronic illness or disability will redefine what it means to have “high” levels of happiness or “7 out of 10” quality of life.

9.3.2 Change in values

Quality of life reports can also be influenced by how people’s values change in response to their circumstances. For example, imagine that soon after developing paraplegia, a person is despondent, because he has to give up physical activities that mattered a great deal to him. But over time, his values change. He places a higher value on intellectual pursuits than on physical activity. And with that change in values, comes a change in his self-reported quality of life.

Changing values can be a mechanism by which people emotionally adapt to illness or disability.[196] For example, a study of patients with prostate cancer who had experienced treatment side effects showed that those men who were able to change their values, by shifting what was important in their lives, were able to maintain a high QoL in the face of treatment complications, whereas those men who did not change their values experienced a decline.[197]

Note that by changing values, people exhibit response shift, but this cause of response shift does not necessarily invalidate QoL measures. By contrast, scale recalibration is by definition a threat to the validity of QoL measures. A person who recalibrates a scale makes it hard for researchers to compare one self-report to another. But a person who changes her values has not necessarily invalidated her QoL reports. Instead, changing values in this way can simply be a mechanism by which people gain true changes in QoL.

9.3.3 Reconceptualization

A third component of response shift is reconceptualization of the construct being measured. Reconceptualization is a challenge for QoL measurement, because the meaning of QoL is broad, and therefore can be interpreted differently by different people and, of even more concern, can be interpreted differently by the same person at two points in time. For example, prior to experiencing an illness, a person may evaluate his QoL primarily on affective or emotional grounds. By this interpretation, he would provide a high QoL rating only if the frequency and intensity of positive affect in his moment to moment life significantly outweighed the frequency and intensity of negative affect.[198] But in response to his illness, he may care less about his mood and place more importance on the frequency with which he pursues meaningful activities.[199] What does reconceptualization mean for the validity of QoL reports? The answer is not clear. If we want people to freely interpret QoL (or happiness, or well-being), then reconceptualization is not a threat. However, if we want them to judge their QoL at any given time using the exact same definition of QoL, then this is a threat, and we need either to find ways to prevent such reconceptualization or, instead, to develop methods to discover whether such reconceptualization is influencing people's self-reports.

For example, it is plausible that people adapt to chronic illness or disability in part by changing their values. Prior to an illness, a person might consider athletic activity to be an important component of his QoL. After being sick, he may think of QoL in different ways, no longer feeling that a vigorous workout is key to the good life. His life goals and values may migrate from physical recreation to intellectual pursuits. Indeed, this change in values may even lead him to redefine what it means to have good QoL (hence, showing the subtleness of the distinction between the second and third parts of Sprangers' definition of response shift).

9.4 Problems with current conceptualization response shift

Based on Sprangers' definition, response shift can occur through several mechanisms, some of which raise fundamental questions regarding the validity of self-reported QoL, but others which do not.

9.4.1 Connotation that response shift is always a threat to validity of self-reports

The way response shift is defined, a phenomenon like scale recalibration-a true source of measurement error-is lumped together with other phenomena that do not necessarily create measurement error. Unfortunately, the term "response shift" conjures connotations that more clearly resemble scale recalibration than the other phenomena. Consequently, some people have mistakenly assumed that response shift is always evidence that QoL measures are not valid. For instance, Brossart and colleagues describe response shift as a "threat to validity of outcome data".[200] Wilson writes about distinguishing "between true change (which here is called a "shift") and scale recalibration, concept redefinition or a change in values ("response shift")".[192] In other words, Wilson believes response shift is pseudo change, not true change. Similarly, in introducing medical researchers to response shift, Schwartz and Sprangers state that when response shift occurs: "answers to the same items by the same individual may not be as comparable as originally thought."

In a longitudinal study of people with multiple sclerosis, Schwartz and colleagues continue this line of reasoning: The apparent stability in these QoL outcomes over five years of follow up might be considered a gain from the perspective of optimal rehabilitation. Our data suggest that this "gain" [their quotation marks] may reflect recalibration and reconceptualization response shiftsĚThus, overall patients' QoL conceptualizations seemed to reduce an emphasis on physical functioning and increase an emphasis on psychological well-being.[201] This quote implies that the QoL of people with multiple sclerosis is a measurement artifact, because people have found happiness by changing their life goals. Clearly, to understand the QoL of people with chronic illness or disability, it is important not only to know what their overall quality of life is, but also to understand what they mean by quality of life. It is valuable to determine whether the things they value in their lives have been changed by their experiences. But to lump all of these together under one header, ”response shift,” and to then imply that high reported quality of life is not to be trusted, is not justifiable. The problem, once again, is that the term ”response shift” carries a specific connotation-that the self-reports of people with chronic illness and disability are misreports. Yet when people find happiness by shifting their values, their high self-reported QoL may simply reflect that they have a good QoL!

9.4.2 Identification of response shift with the “Then Test”

Why do people believe that response shift does not reflect true change? One possible reason, as we have suggested, is that the term response shift carries connotations that fit much better with scale recalibration than with the other two components of response shift. In fact, the term response shift was initially used in the educational literature in the 1970s and was specifically limited to the concept of scale recalibration.

There is another reason, however, why researchers often equate response shift with measurement error-because they have too often relied on the Then Test to determine whether a QoL self-report is being influenced by response shift.[193] In the Then Test, researchers collect at least three data points: (1) a baseline or “Time 1” measure of QoL; (2) a “Time 2” measure of QoL; and (3) a “Then” measure of QoL-a retrospective assessment at Time 2 of what one’s QoL was at Time 1. For example, imagine a patient with chronic 7/10 pain at Time 1, who receives a new treatment for his pain and at Time 2 reports experiencing only 5/10 pain. Without conducting a Then Test, many researchers would conclude from these data that this patient experienced a 2 out of 10 reduction in his pain. But suppose the Then Test reveals that this patient now judges his Time 1 pain as having been 9/10. This suggests that the patient has reinterpreted what the points on the pain scale mean. The patient is telling us that he experienced a 4 out of 10 reduction in his pain (from 9 to 5), according to his new interpretation of the pain scale.

In the next section, we will explain why we believe people are misinterpreting the Then Test. But for our present purposes, we merely want to establish that the vast majority of researchers, when studying a response shift, have focused on the Then Test and have therefore equated response shift with scale recalibration. Response shift means, for most people, the same thing as scale recalibration, and so when they find evidence of response shift, they assume that they have found evidence of measurement error.

9.4.3 Misinterpretation of the Then Test

As we have explained, response shift is primarily identified by use of the Then Test. This reliance on the Then Test is troubling, because most researchers misinterpret the data from this test, by assuming that if the Time 1 measure differs from the retrospective measure, then patients must be recalibrating the outcome scale. By making this false assumption, researchers have been downplaying the likelihood that the retrospective measure is being influenced by recall bias [20].

People's theories about how a life domain is supposed to have changed over time can bias their retrospective reports about those very domains. For example, Ross conducted a study in which he assessed people's study skills, by an objective measure of such skills. He also had people provide self-reports of their own study skills. He found at baseline that people had very good insights into their study skills- a person who was, say, 6/10 on a study skills score typically perceived himself as being a 6/10. Then, Ross followed these students after they took a course designed, purportedly, to improve their study skills. As it turns out, the course had no effect at all on people's study skills. A student who began with study skills of 6/10 would typically end up with study skills of 6/10 at the end of the course. When Ross reassessed these people's study skills, he demonstrated that the course had failed, and he also found that students still had accurate perceptions of their own study skills. A student who was still 6/10 typically perceived himself as being 6/10.[42]

But here is the catch: at the end of the course, Ross asked the students to recall what their study skills had been before they began the course. Students had a theory that this course would improve their study skills. So at the end of the course, a student who accurately reported himself as having study skills of 6/10 would typically “remember” that he began the course with study skills of only 3 or 4/10. The students, in other words, misremembered their previous study skills, because their memory was influenced by their theory about how their study skills should have changed over time.

In this example, students did not exhibit any scale recalibration. Their Time 1 and Time 2 self-reports were entirely accurate. Yet, response shift researchers, relying on the Then Test, could potentially look at these same data (in the absence of objective measures of study skills) and conclude that response shift had occurred-that people had recalibrated the study skills scale. Our research team has found a similar theory-driven recall bias affecting people's beliefs about how their happiness has changed over their life span.[43] We have also shown that recall bias influences patients' assessments of how much they have benefited from kidney transplantation.[202] The key point here is that the Then test is not able to distinguish between scale recalibration and recall bias, two distinct phenomena.

9.4.4 Lumping instead of splitting

The most fundamental problem with the term response shift is that it lumps together distinct phenomena, and in doing so makes it harder for researchers to disentangle scale recalibration from true change. Response shift theorists are correct to note that multiple phenomena could be simultaneously influencing people's subjective self-reports. A person's pain score over time may be influenced by both scale recalibration and true change; her retrospective report might simultaneously be influenced by both scale recalibration and recall bias. Leaders in response shift theory have advanced the field by drawing attention to those important phenomena. But for the field to reach its potential, the time has come to disentangle these phenomena. And we believe this disentangling will move forward more quickly if researchers adopt more precise language.

9.5 Where do we go from here?

QoL researchers want to know when their measures are reliable, valid, and comparable within people across time. When patients receive healthcare interventions, or experience changes in health, researchers want to know what it means when their QoL scores also change. Researchers also want to know whether these changes are real or instead reflect measurement bias. We think that conceptual confusion around the concept of response shift has impeded researchers in pursuing these important goals. We suggest the following to address this problem: "Abandon response shift"

To achieve this goal, researchers need to abandon the term "response shift." The term is simply too confusing to help researchers disentangle these very complicated issues. The language and definition of response shift impedes research by lumping together sources of measurement bias-scale recalibration-and other phenomena-like change in values-that may simply reflect mechanisms by which people experience true change.

9.5.1 Use precise language

In place of response shift, researchers should use more precise language to characterize the specific issues they are assessing. Scale recalibration, for example, has a more precise and narrow connotation than response shift. By definition, scale recalibration is an example of measurement bias. If scale recalibration has occurred, researchers can be confident that measurement bias exists, and therefore that the measures are not comparable over time. If response shift has occurred, however, researchers do not know whether QoL measures over time are comparable or not.

Other terms exist in the scientific literature that capture other relevant phenomena currently lumped together under the concept of response shift. For example, emotional adaptation, or hedonic adaptation, is a term used by psychologists to characterize true changes in subjective wellbeing or happiness over time, when people's emotional reaction to changing circumstances weakens.[31, 35] We predict that if researchers set out in advance to identify whether a specific change in QoL is a result of scale recalibration versus adaptation, they will be much more likely to generate useful research results than if they simply set out to determine whether the measurement reflects the more ambiguous concept of response shift.

9.5.2 Move beyond the Then Test

Researchers should not rely on the Then Test to reveal whether scale recalibration has occurred. The causes of Then Test discrepancies are too ambiguous to lay out solely at the feet of scale recalibration. Discrepancies between time 1 measures and recalls of time 1 could result from scale recalibration, but could also result from recall bias. In place of the Then Test, researchers should use other methods to test for scale recalibration. We presented one such method in this article, when we discussed how scale recalibration influences people's interpretation of ”perfect health” as a function of their age. We elaborate on other methods elsewhere.[112]

9.5.3 More careful review of response shift research

Journal editors and reviewers should look carefully when scrutinizing studies that explore the concept of response shift. They should look to see whether the study focuses solely on scale recalibration. When it does so, they should ask the researchers to rewrite the study using more precise terminology. If they note that the researchers are using the Then Test as the sole evidence of scale recalibration, they should make sure the authors indicate that other phenomena, like recall bias, could be influencing their results. And when the research focuses on how things like change in values can lead people to experience true increases in their quality of life, they should ask authors not to call this an example of response shift, but instead to use other appropriate terminology, such as adaptation or resilience.

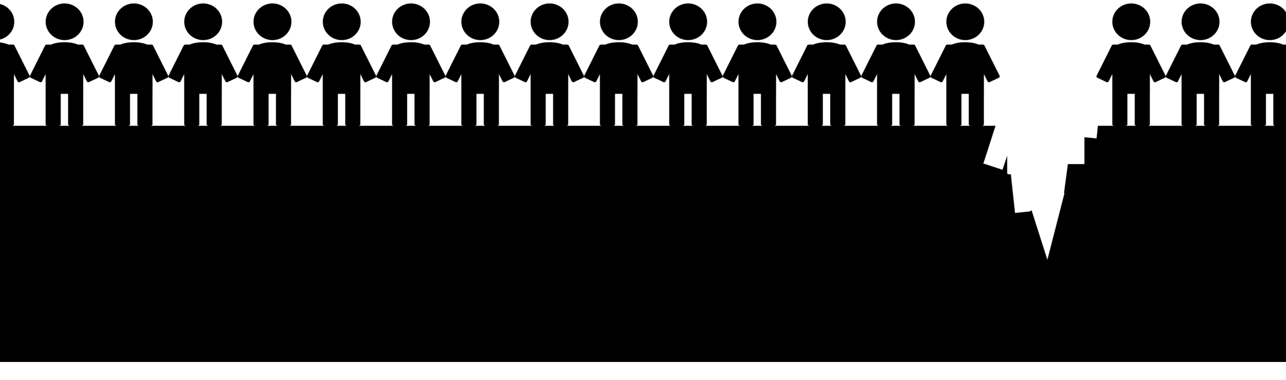
9.6 Concluding remarks

We expect that the ideas we present here will be controversial. Some scholars have made reputations for themselves by disseminating the concept of response shift. But the long-term success of QoL research depends on researchers striving for precision and clarity in their work. Concise, specific terminology is an important part of this research enterprise.

Response shift experts have done a worthy job of drawing scholars' attention to important issues like scale recalibration and adaptation. They should be commended for generating so much interest in these important topics. Now, however, it is time for QoL researchers to abandon the term response shift and focus their efforts on determining when they can trust the comparability of their QoL measures across time. We recognize that our article raises more questions than it answers. Future research should explore better ways to empirically disentangle some of the concepts we identify in this manuscript. But our goal in this article is not to show the field how to disentangle all these complex issues. Instead, we have set out to demonstrate that, by lumping distinct phenomena under a single term, response shift, we make it that much harder for researchers to begin to disentangle these distinct phenomena.

10

Summary & General Discussion



10.1 Summary

In medical care resources are scarce and choices have to be made about how these resources are distributed. To judge the optimal allocation of medical resources, cost-utility analyses can be used. In cost-utility analysis costs of a therapy are compared with the benefits in health. These benefits are estimated by health state utilities, preferences for certain health states compared to perfect health and death. Health state utilities can be elicited from patients or from members of the public. Whose' utilities are used does matter. Utilities of members of the public are found to be lower than health state utilities of patients. Several mechanisms causing this gap in health state utilities between patients and members of the public have been put forward as explanations.[17, 18, 203] The overall objective of this thesis is to examine some of these mechanisms, in order to gain insight in the validity of health state utilities estimated by patients or by members of the public.

Previous studies have described contrasting findings about the gap in health state utilities given by patients or members of the public.[16, 44] Chapter 2 presents a meta-analytical comparison of health state utilities given by patients or members of the public. This comparison is based on 30 eligible studies including 40 estimators, which were rated by two independent judges. Overall it was found that patients gave significantly higher valuations than nonpatients (Cohen's $d = 0.20, p < 0.01$). For each elicitation method separately this resulted in an unstandardized $d = 0.05 (p < 0.05)$ for the TTO, an unstandardized $d = 0.04 (p < 0.05)$for the VAS, and an unstandardized $d = 0.01 (p = 0.70)$ for the SG. Post hoc moderator analyses showed that the difference between patients and nonpatients became smaller when an illness label was provided or when patients were asked to give valuations for a health state description instead of for their own health. We can conclude that the respondent group used does influence health state utilities elicited.

In Chapter 3 the effect of the health state description on health state utilities among patients with RA was investigated. Patients were asked to rated their own health based on three different descriptions. First, patients gave a valuation of their own experienced health of the previous week, secondly, they were asked to value a standard EQ-5D description of their own health (which was untold to them), and third, patients valued a standard EQ-5D description of their own health (identified as such) enriched with individual attributes. All valuations were given on a TTO. We found no differences in health state valuations between the three health state descriptions, but an interaction effect showed that patients with better health appeared to give lower valuations to their own experienced health compared to the enriched EQ-5D health state description of their own health. Surprisingly, utilities for scenarios enriched with exclusively negative individual attributes were not lower than those for the own standard EQ-5D description. In conclusion, it remains unclear if disparities in valuations can be attributed to the EQ-5D description being too sparse.

Chapter 4 describes a study in which the mechanisms of focusing illusion and adaptation are examined. Patients with RA and members of the public imagining having RA were asked to name aspects that were important to them in their previous week. Secondly, all participants rated the importance of these self-named aspects and of the EQ-5D dimensions. Overall we found that members of the public tend to name more aspects in life domains which are negatively influenced by the disease. Patients named both positive and negative aspects, reflecting life in general as well as their illness. Members of the public rated their own named aspects as less important compared to the EQ-5D dimensions, whereas patients found both groups of aspects equally important.

In Chapter 5 we compared values of patients, their partners, and members of the public to examine the effect of vicarious experience. The inconclusive findings described in Chapter 3 led us to question whether enriching a scenario would make valuations of members of the public more similar to those of patients. Therefore we also whished to assess whether the difference in health state valuations maybe due to a scenario-effect, and whether 'enriching' a scenario reduces differences between the groups. Data were collected using semi-structured interviews, similar as the one used in chapter 3, among patients with RA, partners of patients and members of the public. For all three health state descriptions ratings were significantly higher for patients than for the general public. Ratings for partners were in between. Differences between patients and the public are not likely to be due to sparseness of the scenarios, but may be due to a focusing illusion, enhanced by the negative framing of scenarios.

Chapter 6 examines the effect of adaptive abilities on health state utilities. We investigated the direct effect of adaptive abilities on health state utilities and the effect mediated by mental HRQL. Adaptive abilities were conceptualized by combining the three constructs Self-esteem, Mastery, and Optimism, as suggested by Cognitive Adaptation Theory. In an interview, patients with RA gave health state valuations for their previous week on a VAS and TTO, and filled in questionnaires measuring Self-esteem, Mastery, Optimism, and HRQL. Persons' ability to adapt did not add considerably to the explanation of health state utilities above HRQL. In the TTO no additional variance was explained by adaptive abilities ($\Delta R^2 = .00, \beta = .02$), in the VAS a minor proportion of the variance was explained by adaptive abilities ($\Delta R^2 = .05, \beta = .33$). However, the effect of adaptation on health state utilities seemed to be mediated by the mental health domain of quality of life. Patients with stronger adaptive abilities may more easily enhance their mental health after being diagnosed with a chronic illness, which leads to higher health state utilities.

Chapter 7 further describes adaptation and examines valuation shift, in a longitudinal study among patients with Spinal Cord Injury (SCI). Health state valuations of patients with recent onset SCI were assessed at three points in time. At each time point patients gave valuations for their own health and for a hypothetical health state on a TTO and VAS. Furthermore patients rated their adjustment and filled in a questionnaire measuring their independence. Valuations given on the TTO for the own health did change over time, above improvement in independence. Furthermore we found that this change was partly related to change in self-rated adjustment. Health state valuations for the own health given on the VAS and corrected for independence did not change over time. No change was found either in valuations for the hypothetical health state description. We can conclude that the effect of psychological adaptation has an impact on the valuations of the own health based on the TTO in serious incurable disease. No support was found for the effect of valuation shift.

In Chapter 8 the effect of adaptation as well as the effect of implicit theories of stability and change were studied. Two longitudinal studies are presented: one study among patients experiencing new limb amputations, and one among patients experiencing new colostomies. In both studies patients estimated at two points in time their current well-being, functioning, and general health and predicted of their future well-being, functioning, and general health. In the second study patients not only predicted their well-being, functioning, and general health, but they also recalled previous performance on these domains. In both studies the actual change was compared to the change expected and recalled by patients. It appeared that patients expect a significant improvement and recall significant improvement on all three domains, but report little actual improvement. Apparently, patients expect stronger improvement than they actually experience: they overestimate their hedonic adaptation.

While examining the different mechanisms suggested in the literature causing the difference between health state valuations of patients and members of the public we were often challenged by ambiguous descriptions of the mechanisms. Chapter 9 describes such a conceptual confusions specific to the field of “response shift”. To clarify the conceptual confusion around “response shift” a suggestion is made about abandoning this term and instead to use more precise language. First, response shift confounds sources of measurement error, ’scale recalibration’, with true causes of change which in the language of response shift are described as ’change in values’ and ’reconceptualization’. Secondly ’change in values’ or ’reconceptualization’ might better be changed into conceptualizations such as adaptation, which is a common term used by psychologist.

10.2 General Discussion

The main objective of this thesis was to examine mechanisms that have been suggested to explain the gap between health state utilities given by patients and by members of the public. By examining these mechanisms more insight in the validity of health state utilities of patients and members of the public is gained. In the first part of this discussion the conclusions retrieved from the previous Chapters 2 - 9 are described. The next part consists of an evaluation of the results. Finally the policy implications, implications for patient decision making and suggestions of areas for future research are discussed.

10.2.1 Mechanisms underlying the gap between members of the public and patients

Lack of Scope The results of this thesis give only limited ground for the effect of Lack of Scope. Throughout the whole thesis exploratory findings tentatively confirmed an effect of Lack of Scope. That is, in Chapter 2 the discrepancy between health state utilities of patients and members of the public was shown to be smaller in studies in which a label was provided. In Chapter 3 the EQ-5D seemed to be too sparse for patients in better health, and in chapter 4 patients valued their own named aspects as important as EQ-5D dimensions. These findings were often based on post hoc analyses, however, and were not very strong. Lack of Scope was most thoroughly examined in Chapters 3 and 5, using different health state descriptions. In both chapters the effect of using different health state descriptions was minor.

Focusing illusion Support was found for a focusing illusion. Members of the public were more focused on life domains that are negatively influenced by a health state compared to patients (Chapter 4). Patients seemed not overly optimistic; they did not overlook the negative aspects of their health state (chapter 4).

Valuation shift Previous lack of support for valuation shift was suggested to be due to studies investigating rather mild, non-permanent health states.,[2040] In this thesis, patients with Spinal Cord Injury, a serious condition with little or no prospect for cure, were studied, but still no support for valuation shift was found (Chapter 7).

Adaptation The results of this thesis showed that an effect of adaptation cannot be ruled out. Adaptive abilities indirectly influenced health state utilities, (Chapter 6) health state utilities measured with a TTO did change over time due to adaptation over and above improvement in independence (Chapter 7). Further, based on aspects indicated as important to patients, patients seemed to have different interests compared to members of the public (Chapter 4). Interestingly, the influence of adaptation was also seen in partners (Chapter 5) and patients early in the experience of an illness (Chapter 8). Yet, this latter group tended to overestimate their ability to adapt.

Implicit theories of stability and change Patients early in the experience of an illness seemed to base their valuations on implicit theories of stability and change. They assumed that after surgery they will get better, and that this improvement will continue over time, leading to an underestimation when recalling previous health and an overestimation when making predictions of their future health (Chapter 8).

10.2.2 Evaluating the results

Health state utilities used in cost-utility analyses to compare the costs and benefits of interventions have generally been elicited from members of the public or patients.[15] Whose preferences are used does have influence, patients tend to give higher valuations compared to members of the public (Chapter 2). Based on the results described in this thesis most of the mechanisms generally suggested do not or only marginally explain the gap between patients and public, except for the mechanisms focusing illusion and adaptation. Clearly no conclusion can be drawn about mechanisms not examined. Most of the mechanisms described in the introduction were examined, but the intensity differed.

Focusing illusion is the tendency of members of the public to focus on the difference between their current situation and a situation under consideration.[27] In previous studies in the field of medical decision making no empirical evidence for this focusing illusion has been described. Studies have aimed to defocus members of the public through a defocusing task[24] but they were not able to reduce the focusing illusion. The study described here (chapter 4) in which evidence for focusing illusion in medical decision making was shown, was directed at describing focusing illusion. Health state valuations of members of the public seem to be biased due to focusing illusion, but reduction of this focusing illusion is challenging.

The second mechanism for which support has been found is adaptation. Adaptation is one of the most often suggested mechanisms causing the gap, especially when coping strategies and the constructs "change in values" and "reconceptualization" of response shift (Chapter 9) are taken into account. Damschroder et al.[38] found tentative support for the effect of adaptation when using an adaptation task. Members of the public who received an adaptation task gave higher health state valuations on PTO and VAS compared to members of the public who did not receive this task. However in a follow up study using the TTO and SG no such difference was found.[39]

It has been argued that the expected impact of adaptation on health state utilities is less prevalent than previously assumed. For instance, Lucas[186] described that patients do not adapt to onset of major disability. In this study, two nationally representative panel studies from Germany and from the United Kingdom were used. Based on these data it was shown that patients with onset major disability did not return to their previous levels of life satisfaction (prior to their injury). Even seven years after the injury the life satisfaction was as low as in the first year after onset of the disability. Surely this study reveals much interesting information but it is limited by the fact that the results were restricted to annual measurement points in time.

The effect of adaptation on health state utilities in this thesis is shown among patients with RA and patients with SCI. Anticipated adaptation is shown among patients who underwent amputation or colostomy surgery and partners of patients with RA. The effect of adaptation seems to generalize among different patient groups and even partners of patients. But even though the main conclusions are similar, some of the underlying results point to interesting differences. The adaptive abilities among patients with RA seemed to have more impact on the VAS (Chapter 6) whereas the change over time among patients with SCI was only prevalent in valuations elicited with the TTO (Chapter 7). These distinct findings might have been caused by the affective and cognitive nature of the measurements. The VAS can be seen as a more affective measurement whereas the TTO as a more cognitive measurement. Affective measurements have been found to be more sensitive to change,[172] leading to a stronger effect of adaptive abilities on the VAS in Chapter 6. Yet, cognitive measurements might be more strongly related to cognitive change, such as cognitive adaptation. The TTO in chapter 7 might have triggered patients with SCI to think about future life goals which may be influenced by cognitive adaptation. A different explanation for the distinct findings in Chapters 6 and 7 might be related to the severity of the health states under consideration. Adaptation and adaptive abilities might have a different impact on health states of varying severity levels.

The mechanisms causing the gap between health state utilities of patients and members of the public have been suggested based on cross fertilization between research fields such as the fields of HRQL research, medical decision making, and social and health psychology. However, the translation of a theory from a particular research field into a mechanism explaining the gap in health state utilities is challenging. Theories from different research fields may lead to similar mechanisms with only subtle differences. Given these subtle differences and often ambiguous descriptions, similar mechanisms are difficult to examine separately.

For instance, the distinction between the failure to anticipate on adaptation on the one hand and focusing illusion on the other can be ambiguous. Some of the findings in Chapter 4 on attributes of importance could be caused by a failure to anticipate on adaptation, or by a focusing illusion, or maybe by both. Ubel et al.[24] tried to describe how to distinguish between these mechanisms. According to these authors focusing illusion is defined as a failure to appreciate that not all life domains or life events will be equally affected by a given change in circumstances. Failure to anticipate on adaptation is defined as the failure to appreciate that one's emotional response to the given change in circumstances will diminish over time. However, they do not elaborate on the distinction between focusing illusion and other aspects of adaptation, such as shifting goals and priorities, which they had described as aspects of adaptation in previous papers.[18, 177]

Both in the definition of Ubel et al.[18] and according to Schkade and Kahneman,[181] focusing illusion is caused by overweighting of a subset of aspects of an entire object under consideration, for instance when only attention is drawn to a change in significant aspects of life. If we revert to this definition of focusing illusion, a failure to anticipate on a shift in priorities or goals then seem to be distinct from focusing illusion by the fact that it is not caused by an exaggerated focus on a subset of the object under consideration. Regarding the findings in Chapter 4 it seems that the distinction between focusing illusion and adaptation remains challenging due to the limited information we have. By statements as "I am not able to play soccer" no information is provided about the underlying grounds. If a member of the public makes such statement he or she might have been thinking about the differences between his or her current situation compared to the situation when having RA (focusing illusion) or might have thought about both situations in general but failed to imagine that instead of playing soccer fulfillment can be retrieved by watching soccer instead (adaptation).

10.2.3 Policy implications

Organizations developing guidelines on the use of new and existing treatments (NICE, the panel of the U.S. Public Health Service, CvZ) advise the use of a societal perspective in which health state utilities elicited from members of the public are preferred.[12, 14] However given the experience of patients it has incidentally been argued that patient valuations should be used instead.[12] The findings described in this thesis reveal that health state utilities of patients and public do differ (Chapter 2) which is mostly caused by adaptation (Chapter 4, 5, 6, 7) and focusing illusion (Chapters 4 and 5).

Whose utilities we should use depends on whose utilities are most valid. This validity depends on the effect that the underlying mechanisms have, that is, do they cause a true effect or do they cause an error in the health state utility elicited? Obviously utilities of members of the public that are shaped by focusing illusion are biased.[26, 37] If members of the public focus too narrowly on only negative consequences of the health state under consideration they overestimate the burden of the health state leading to too low health state utilities.

The question if patients' health state utilities shaped by adaptation are biased is more difficult to answer. Ubel et al.[37] assume that failure to anticipate on adaptation leads to misjudgments of members of the public. Yet, in the field of response shift, patients' adaptation is seen as a validity threat leading to misjudgments of patients,[41] assuming that adaptation is described as change in values and reconceptualisation (Chapter 9). According to Oort et al.[204] and Sprangers and Schwartz,[41] true change in Quality of Life (QL) is only reflected when it is corrected for change in values, reconceptualization, and scale recalibration. This implies that change in patients' QL due to adaptation is regarded as untrue change. However, given that by definition (World Health Organization) QL is based on an individual's perception, change as experienced by a patient must truly reflect change. That is, QL is a subjective concept based on individual perceptions and should include subjective change if this reflects true change from the patients' point of view. A change in perception does not have to harm the validity of QL measurement. Therefore, health state valuations shaped by adaptation do reflect true experienced health. It also seems necessary to use health state valuations shaped by adaptation when it comes to the fact that health state utilities should include all aspects of life that are affected by the health state. Patients' understanding of life in their current health state includes the influence of adaptation, especially as the influence of adaptation will differ between health states, thereby providing valid information about the perceived severity of the illness. In other words, if the ability to adapt differs between health states this does reveal important information about the severity of the illness, which should be accounted for in cost-utility analyses. Moreover, the health care system does put effort in developing physical rehabilitation and psychological interventions to help people to adjust to their illness, since this leads to an improved QL. We have to keep in mind that although patients do adapt, when given the choice they probably would not have wanted to adapt in the first place. It may even seem counterintuitive to take full adaptation into account. Gilbert et al.[26] argue that if people would take their ability to adapt into account, they would not engage in problem avoiding behavior and become "happily extinct". Although this suggestion is made for coping behavior of individual people, it can be transposed to cost-utility analyses. If people adapt to their illness anyway, we would not have to worry about health state utilities but only take the number of lives saved into account.

Further, health state utilities shaped by adaptation leave less space for a gain in benefit by treatment. The effort that patients put in adapting to their illness will cause effectiveness of treatments to decrease.[36] This creates an ethical problem; by using health state utilities shaped by adaptation the effort that patients make in adapting will lead to a decrease in availability of resources.

In summary, health state utilities shaped by adaptation are not biased, they do reflect true values, and more importantly they provide valuable information about the severity of health states. However, using utilities shaped by adaptation will lead to ethical challenges. That is, if people fully adapt, medical treatment would not be necessary anymore, or would not be cost-effective. Therefore adaptation should be taken into account in cost-utility analyses but only to a certain extent.

Concerning to what extent health state utilities shaped by adaptation should be used some suggestions from the literature may be relevant. Menzel et al.[205] suggested that disabled and chronically ill should be consulted on this topic. Another approach is that of "cost-value analysis".[10] In this approach two stages are used to construct a valuation. In stage 1, patients assign health-related utilities, and in stage 2, based on these health state utilities, the general public assigns weights to different gains in utility. A third approach might be asking patients to recall how their health state has changed over time.[37] However, given the findings described in Chapter 8 the recalled health valuation will be significantly different from the actual health valuation. Until now it remains unclear which valuation reflects the true value. Another approach examined in this thesis might be using health state utilities of patients given to unfamiliar health states. Given the lack of evidence for valuations shift, it might be concluded that utilities of patients imagining an unfamiliar health state are similar to valuations given by members of the public. Also examined in this thesis are health state utilities given by proxies. These valuations appeared to lie between valuations given by patients and members of the public (Chapter 5). Likewise in the meta-analyses we found health state valuations of proxies to lie between valuations given by members of the public and patients (Chapter 2). Partners do seem to incorporate (vicarious) experience. The use of partners' valuations might be an interesting new approach.

For now, no conclusive advice can be given on the extent to which health state utilities should incorporate health state valuations of patients, to be used in cost-utility analyses. We do conclude that health state utilities of members of the public are biased by focusing illusion and that adaptation should be incorporated to some extent, but not fully. Therefore the use of health state utilities of patients or maybe partners of patients should be open to consideration by organizations developing guidelines on the use of new and existing treatments.

10.2.4 Implications for patient decision making

In individual treatment decisions, understanding of the mechanism causing the gap in health state utilities between patients and members of the public is also important.[206] It seems that when people imagine living in a certain health state they fail to anticipate on adaptation and focus only on those life domains that are influenced by the illness. To be able to make treatment decisions, people have to make such affective forecasts of their life in a certain health state. Based on the found biases people might make the wrong decision.

People should be made aware of these biases prior to making treatment decisions. Admittedly, given the results of Chapter 8 this suggestion has to be made with caution. Patients just recovering from surgery overpredicted their ability to adapt. Before being able to inform patients about these biases we have to understand if this overprediction occurs only in patients just having received treatment or also in patients anticipating on their treatment. Secondly, variation exists among people's ability to adapt (Chapter 6). People who have to make treatment decisions might not only benefit from information about their failure to anticipate on adaptation and on their focusing illusion, but also from information on their personal ability to adapt. Finally and most importantly, we always have to keep in mind that patients do adapt but when given the choice they probably would not have wanted to adapt in the first place.

10.2.5 Future research

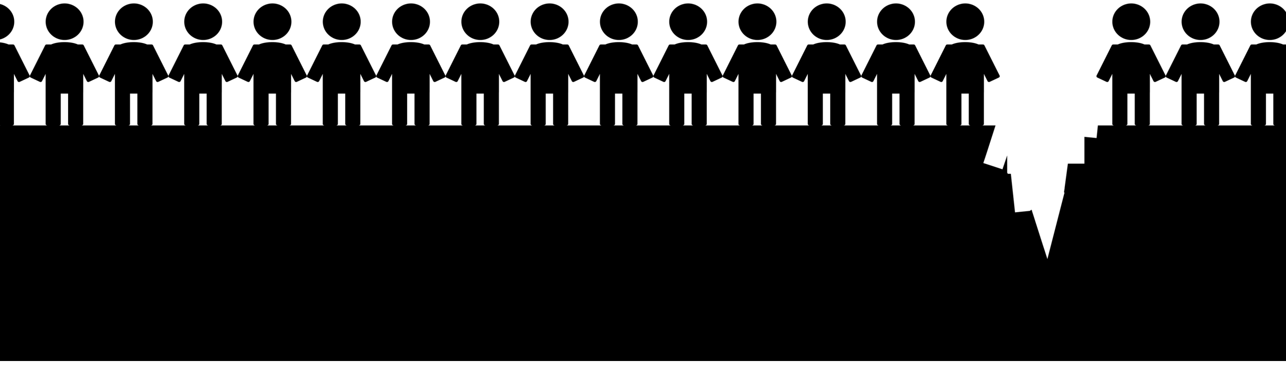
Even from the findings described in this thesis and several studies performed before[20, 38, 40, 107] the gap in health state utilities between patients and members of the public can still not be fully explained with certainty. More research can always be performed to further investigate mechanisms underlying this gap, but it can be questioned if this will lead to more conclusive information. Based on previous finings it seems that several mechanisms are layered on top of each other, although adaptation and focusing illusion seem to have the most influence on health state valuations.

Regarding policy, future research should concentrate on how the effect of adaptation and focusing illusion on health state utilities should be handled. Concerning the ethical dilemma of using health state utilities shaped by adaptation, Menzel et al.[205] suggested that disabled and chronically ill should be consulted. In the medical context future research should concentrate on how and which information about the biases should be provided to patients. However, first the ability to anticipate on adaptation should be investigated among people who are confronted with a treatment decision.

Finally it has to be emphasized that the findings described in this thesis are based on physical health problems. From recent research we know that that the gap in health state valuations between patients and members of the public might be different for mental disorders. Given the burden of mental disorders on society, more future research should focus on mental health.

11

Nederlandse Samenvatting



Bij beslissingen over het al dan niet vergoeden van behandelingen wordt vaak gebruikt gemaakt van kosteneffectiviteitsanalyses, waarin de kosten worden afgezet tegen de baten. Voor het bepalen van deze baten is het nodig om een waardering voor een bepaalde gezondheidstoestand te hebben. Een waardering kan worden gegeven op basis van een gezondheidsbeschrijving of van de eigen ervaren gezondheidstoestand en wordt meestal bepaald aan de hand van één van drie methoden; de Visueel Analoge Schaal (VAS) Time TradeOff (TTO) of de Standard Gamble (SG). De Visueel Analoge Schaal (VAS) is een lijn van 100 mm. die loopt van 0 (overleden of dood) tot 1 (perfecte gezondheid). Respondenten krijgen een omschrijving van een gezondheidstoestand en geven aan waar op deze lijn zij zich plaatsen als ze zich in de beschreven gezondheidstoestand zouden bevinden. De afstand tussen dood (0) en de plek die wordt aangeven is de waardering. De Time Trade-Off (TTO) is een methode waarin respondenten afwegen hoeveel jaar ze in perfecte gezondheid ze gelijk vinden aan een bepaald aantal jaar (in dit proefschrift wordt de levensverwachting gebruikt) in de beschreven gezondheidstoestand. Hiervoor worden respondenten allereerst gevraagd of ze liever x-jaar in perfecte gezondheid of x-jaar in de beschreven gezondheidstoestand willen leven - na deze periode zou de respondent plotseling overlijden. De meeste mensen kiezen dan voor x-jaar in perfecte gezondheid. Vervolgens wordt dezelfde vraag gesteld maar dan met $x/2$-jaar in perfecte gezondheid. Ook nu maakt de respondent een keuze op basis waarvan het aantal jaren in perfecte gezondheid weer wordt aangepast. De waardering is het aantal jaar in perfecte gezondheid gedeeld door x (het aantal jaar in de beschreven gezondheidstoestand). In de Standard Gamble maakt de respondent steeds een keuze uit leven in de beschreven gezondheidstoestand of een kans (p) op perfecte gezondheid maar ook een kans van $(1 - p)$ op overlijden. De kans op perfecte gezondheid wordt, net zoals de jaren in de TTO, gevarieerd totdat de respondent niet meer kan kiezen. Deze kans (p) is dan de waardering voor de beschreven gezondheidstoestand.

Dergelijke waarderingen, ook wel utiliteiten genoemd, worden meestal gegeven door de algemene bevolking op basis van een gezondheidsbeschrijving die vaak summier is en de negatieve kant van de gezondheidstoestand overbelicht. Daarom is gesuggereerd dat, in de plaats van de algemene bevolking, juist patiënten de waardering moeten geven. Patiënten hebben ervaring met de gezondheidstoestand en daardoor een beter beeld van de situatie. In de praktijk blijkt de algemene bevolking lagere utiliteiten toe te kennen aan hypothetische gezondheidstoestanden dan patiënten die ervaring hebben met die gezondheidstoestand. Het is daarom belangrijk inzicht te krijgen in de wijze waarop patiënten in vergelijking tot de algemene bevolking tot een oordeel komen over een bepaalde gezondheidstoestand. In dit proefschrift hebben we gekeken naar welke aspecten ten grond slag liggen aan het verschil in waarderen. Door hier inzicht in te krijgen, hoopten we een onderbouwde suggestie te kunnen doen wiens gezondheidswaarderingen zouden moeten worden gebruikt in kosteneffectiviteitsanalyses.

Omdat in de toch literatuur discussie blijft bestaan over het verschil in waarderen tussen patiënten en algemene bevolking en de hogere waarderingen van patiënten soms in twijfel getrokken worden, is in hoofdstuk 2 allereerst een beschrijving gegeven van een meta-analyse waarin we dit verschil nader onderzocht hebben. Op basis van 40 waarderingen voor gezondheidstoestanden van zowel patiënten als de algemene bevolking - beschreven in 30 artikelen - zagen we dat patiënten significant hogere waarderingen geven ten opzichte van de algemene bevolking. Wanneer we deze gegevens opsplitsen naar methode (VAS, TTO & SG), bleek dit verschil zowel in waarderingen gegeven op de VAS als op de TTO voor te komen, maar niet in waarderingen gegeven op de SG. Daarnaast zagen we dat het verschil tussen patiënten en algemene bevolking kleiner werd wanneer de te waarderen gezondheidstoestand was voorzien van de naam van de beschreven ziekte, of, wanneer patiënten en de algemene bevolking beide waarderingen gaven aan gezondheidsbeschrijvingen.

In hoofdstuk 3 onderzochten we het effect van de gezondheidbeschrijving op de waardering. Patiënten met reuma werden gevraagd hun eigen situatie drie keer te waarderen op de TTO aan de hand van drie verschillende omschrijvingen. Allereerst gaven ze een waardering van hun eigen gezondheidstoestand zoals ze deze hadden ervaren in de voorafgaande week, daarna van een EQ-5D beschrijving van hun eigen situatie (waarvan ze niet wisten dat het over hun eigen situatie ging) en tot slot een geïndividualiseerde EQ-5D beschrijving (waarvan ze wel wisten dat het over hun eigen situatie ging). De EQ-5D beschrijving was gebaseerd op antwoorden die patiënten eerder in het onderzoek gaven op de EQ-5D vragenlijst. De EQ-5D vragenlijst bestaat uit vijf vragen over de mate van problemen die patiënten hebben op respectievelijk mobiliteit, zelfzorg, dagelijkse activiteiten, pijn & andere klachten en angst & depressie. De mate van problemen op deze vijf domeinen vormt de EQ-5D beschrijving. De EQ-5D beschrijving werd geïndividualiseerd door patiënten te vragen welke aspecten belangrijk waren voor hun Kwaliteit van Leven (KvL) in de voorafgaande week. Deze aspecten, maximaal vijf, werden aan de EQ-5D omschrijving toegevoegd. Uit de resultaten van deze studie bleek dat de gezondheidsbeschrijving weinig invloed heeft op de waardering, alle beschrijvingen werden hetzelfde gewaardeerd. Wel zagen we een interactie effect; patiënten met een betere gezondheidstoestand gaven een lagere waardering aan hun eigen gezondheidstoestand zoals ze deze hadden ervaren in de voorafgaande week ten opzichte van de EQ-5D beschrijving en de geïndividualiseerde EQ-5D beschrijving. Mogelijk dachten deze patiënten aan kleine irritaties (bijvoorbeeld dat ze vanwege medicatie geen alcohol konden drinken op een feest) tijdens de waardering van hun eigen gezondheidstoestand van de voorafgaande week. Het denken aan dergelijke irritaties zal veel minder waarschijnlijk zijn tijdens de waardering van de EQ-5D, omdat zulke irritaties hier niet in terug te zien zijn.

In hoofdstuk 4 gaan we verder in op de aspecten die patiënten genoemd hebben tijdens het interview dat hierboven beschreven staat. Naast patiënten hebben we hetzelfde interview ook bij mensen uit de algemene bevolking afgenomen. Deze groep hebben we gevraagd welke aspecten zij belangrijk gevonden zouden hebben in de voorafgaande week wanneer ze reuma zouden hebben zoals beschreven (Appendix C). Daarnaast vroegen we alle deelnemers aan het onderzoek de eigen aspecten en de EQ-5D domeinen op volgorde van belangrijkheid te leggen. Uit de genoemde aspecten bleek dat de algemene bevolking gericht is op de negatieve invloed die het hebben van reuma zal hebben op hun leven (focusing illusie) terwijl patiënten zowel de negatieve als de positieve kant benoemen. Als we de eigen aspecten laten vergelijken met de EQ-5D dimensies, zien we dat de algemene bevolking de EQ-5D dimensies belangrijker vindt dan de eigen aspecten, terwijl patiënten de EQ-5D dimensies en hun eigen aspecten even belangrijk vinden.

Naar aanleiding van de bevindingen die beschreven staan in hoofdstuk 3 wilden we nagaan of de gezondheidsbeschrijving effect heeft op waarderingen van mensen uit de algemene bevolking of partners van patiënten met reuma (hoofdstuk 5). Zoals we hadden verwacht waren alle patiënt waarderingen hoger dan waarderingen van de algemene bevolking en lagen waarderingen van partners er tussen in. Er bleek echter ook hier geen effect van de gezondheidsbeschrijving op de waarderingen die mensen uit de algemene bevolking of partners van patiënten gaven. Op basis van de resultaten beschreven in dit en bovenstaande hoofdstukken lijkt het verschil in waarderen tussen patiënten en algemene bevolking niet veroorzaakt te worden door de beperkte gezondheidsbeschrijvingen.

In hoofdstuk 6 hebben we onderzocht of het vermogen van patiënten om zich aan hun ziekte aan te passen invloed heeft op de waardering die patiënten geven. Daarnaast hebben we onderzocht of het effect van dit vermogen om zich aan te passen op waarderingen via mentale Kwaliteit van Leven (KvL) verloopt. Dat wil zeggen, heeft het vermogen om aan te passen effect op mentale KvL dat dan weer de waarderingen beïnvloed, of is er een direct effect van het vermogen om aan te passen op waarderingen. De waardering die patiënten met reuma gaven aan hun eigen situatie van de voorafgaande week op zowel de TTO als op de VAS worden inderdaad beïnvloed door het vermogen om aan te passen. Dit effect bleek grotendeels via mentale KvL te verlopen. Het vermogen om zich aan een ziekte aan te passen heeft effect op de mentale KvL en dit heeft effect op de waardering die patiënten geven.

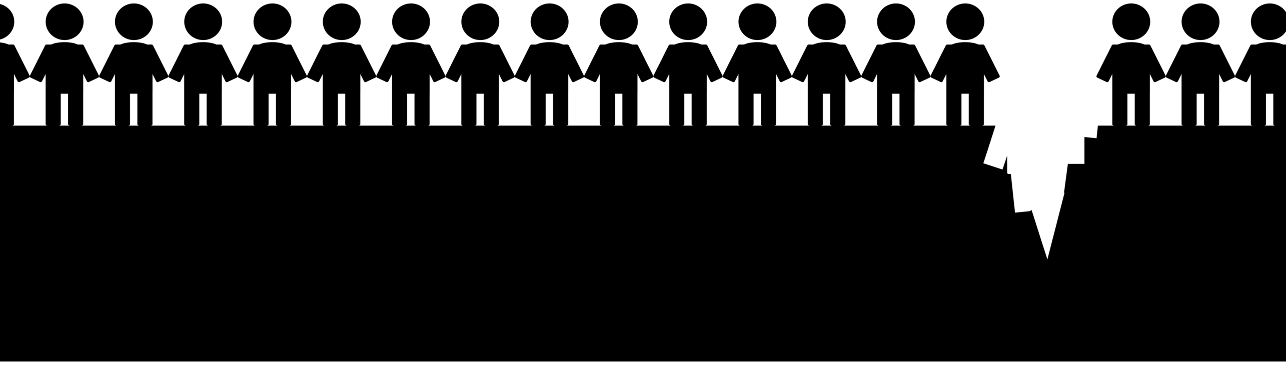
We hebben het effect van aanpassing op waarderingen niet alleen onderzocht onder patiënten met reuma maar ook onder patiënten met een dwarslaesie (Hoofdstuk 7). Patiënten met een dwarslaesie werden drie keer geïnterviewd, (1) zo snel mogelijk na opname in het revalidatie centrum, (2) gedurende actieve revalidatie, en (3) minstens een half jaar na ontslag uit het revalidatie centrum. Alle patiënten gaven een waardering voor hun eigen gezondheidstoestand en voor een hypothetische gezondheidstoestand (Appendix C). Daarnaast vulden we samen met patiënten vragenlijsten in, waaronder een vragenlijst die de onafhankelijkheid in dagelijkse activiteiten meet en een vragenlijst die de mate van aanpassing meet. De waardering op de TTO en de VAS die patiënten gaven aan hun eigen gezondheidstoestand steeg over de tijd. Op de VAS bleek dit effect geheel veroorzaakt te worden door de verbetering in onafhankelijkheid in dagelijkse activiteiten. Met andere woorden, patiënten revalideren en dat is terug te zien is in hun waardering voor hun eigen gezondheidstoestand op een VAS. Op de TTO wordt deze stijging slechts gedeeltelijk verklaard door verbetering in onafhankelijkheid in dagelijkse activiteiten. Aanpassing verklaart verdere stijging, daarnaast blijft een gedeelte onverklaard. De waarderingen op de VAS en TTO voor de hypothetische gezondheidstoestand blijken niet te veranderen over de tijd. In hoofdstuk 8 beschrijven we een studie waarin we hebben onderzocht of patiënten die nog maar weinig ervaring hebben met hun ziekte anticiperen op aanpassing. We geven in dit hoofdstuk een omschrijving van twee longitudinale studies. De eerste studie is uitgevoerd onder patiënten die recentelijk een amputatie hadden ondergaan, de tweede studie is uitgevoerd onder patiënten die recentelijk een stoma hadden gekregen. Patiënten gaven op twee momenten een schatting van hun welzijn, functioneren en algemene gezondheid. Daarnaast gaven ze in het eerste interview een schatting van hun verwachte waarderingen tijdens het tweede interview. In de tweede studie gaven patiënten ook nog in het tweede interview een schatting van hun waardering tijdens het eerste interview. Patiënten verwachtten en herinnerden een significante vooruitgang over de tijd, maar de werkelijke waarderingen waren niet vooruit gegaan. Mogelijkerwijs zijn deze inschattingsfouten te wijten aan onrealistisch optimisme over de toekomst. Aan de andere kant kan het ook veroorzaakt zijn door de manier waarop welzijn, functioneren en algemene gezondheid zijn gemeten. Misschien hebben patiënten in hun waardering rekening gehouden met de periode dat ze uit het ziekenhuis waren. Patiënten die een amputatie hadden ondergaan dachten waarschijnlijk in het eerste interview dat hun functioneren best goed was gegeven dat ze pas een week ontslagen waren uit het ziekenhuis, en in het tweede interview dat hun functioneren best goed was gegeven dat ze al weer een maand ontslagen waren uit het ziekenhuis.

Tijdens ons onderzoek merkten we dat de mechanismen die wij probeerden te onderzoeken niet altijd goed omschreven staan in de literatuur. In hoofdstuk 9 gaan we in op de conceptuele verwarring in de response shift-literatuur. Volgens de response-shift theorie wordt er onterecht gesproken van verandering in KvL wanneer (1) patiënten het concept KvL veranderen (bijvoorbeeld: een patiënt waardeert tijdens het eerste interview vooral lichamelijke gezondheid en op een later tijdstip psychische gezondheid); (2) wanneer de waardering van aspecten van KvL verschuift (bijvoorbeeld: een patiënt geeft eerst een lage waardering omdat hij of zij niet meer kan sporten, maar over de tijd wordt het sporten minder belangrijk en hierdoor stijgt de waardering van KvL) of; (3) wanneer patiënten de schaal waarop KvL gemeten wordt anders gaan interpreteren (bijvoorbeeld: zoals beschreven in hoofdstuk 8, mogelijk hebben de patiënten hier hun antwoorden gegeven in relatie tot wat ze verwachten voor de specifieke periode na ontslag uit het ziekenhuis). Wij vragen ons af of een verandering van het concept van KvL en van waardering van KvL niet vergelijkbaar zijn met aanpassing en daarmee valide veranderingen in KvL veroorzaken.

Uit onze studies zoals hierboven beschreven blijkt dat het verschil in waarderingen tussen patiënten en mensen uit de algemene bevolking niet zo zeer beïnvloed wordt door de vaak beknopte gezondheidsbeschrijving, maar grotendeels ontstaat door cognitieve processen binnen beide groepen. Enerzijds passen patiënten zich aan hun situatie aan, terwijl mensen uit de algemene bevolking niet op deze aanpassing anticiperen bij het geven van een waardering. Anderzijds is de algemene bevolking gericht op de negatieve aspecten van de gezondheidstoestand. Dit zou komen omdat zij zich focussen op het verschil tussen hun huidige toestand, meestal perfecte gezondheid, en de gezondheidstoestand die ze moeten beoordelen. Omdat beide groepen worden beïnvloed door deze cognitieve processen blijft het discutabel wiens waardering een beter weergave geeft van de ernst van een ziekte. In de discussie gaan we hier op in.

Het lijkt duidelijk dat wanneer de algemene bevolking zich te zeer richt op alle negatieve aspecten van een gezondheidstoestand hun waarderingen ook te negatief zullen zijn. Het is echter discutabel of waarderingen van patiënten te hoog worden door aanpassing. Ondanks het feit dat fysiek de gezondheidstoestand niet verandert over de tijd, kunnen patiënten wel een hogere waardering geven aan hun gezondheidstoestand na fysieke en psychische aanpassing. Wanneer een patiënt met een dwarslaesie zich aanpast door zich bijvoorbeeld te richten op nieuwe activiteiten, kan deze hierdoor gelukkiger worden. Moeten we in kosteneffectiviteitsanalyses waarderingen voor de "objectieve" gezondheidstoestand gebruiken, of mag dergelijke psychische aanpassing worden meegenomen? Na deze aanpassing gaat immers wel de kwaliteit van leven van de patiënt omhoog. Het meenemen van aanpassing in kosteneffectiviteitsanalyses zou echter kunnen zorgen voor een vermindert effect van behandel methoden. De effectiviteit van een behandeling wordt bepaald aan de hand van het verschil tussen de waardering voor en na behandeling. Wanneer een patiënt een 0.8 geeft aan zijn of haar gezondheidstoestand kan een behandeling maximaal een effect van 0.2 hebben, terwijl voor iemand uit de algemene bevolking die dezelfde gezondheidstoestand een 0.5 geeft de behandeling een effect van 0.5 kan hebben. Hierdoor ontstaat een ethisch probleem, het meenemen van aanpassing zou een vermindering van beschikbaarheid van behandel methoden kunnen veroorzaken. In een van onze studies hebben we gekeken naar waarderingen van partners van patiënten. Deze waarderingen liggen tussen die van patiënten en de algemene bevolking in. Mogelijk zijn 'partnerwaarderingen' een goed alternatief; partners hebben een beter beeld over het leven met een bepaalde gezondheidstoestand maar worden minder beïnvloed door de verschillende cognitieve processen in vergelijking met patiënten. Waar men zich bij patiënten kan afvragen of aanpassing daadwerkelijk plaats vindt of dat patiënten hun situatie beter voor spiegelen om er zo zelf mee om te kunnen gaan, lijkt dit bij partners minder van belang.

Appendix A - D



Appendix A

Study characteristics of the 30 included studies(Chapter 2)

Source	♯patients	♯non-patients	Unstandardized difference	Elicitation method	Kind Patients	Kind non-patients
Ashby & O'Hanlon (1994)	18	59	0.15^{I}	TTO	Women with breast cancer	Professionals
			0.13^{I}			General public
Badia et al. (1996)	103	360^{*}	0.04^{I}	VAS	Critical medical & surgical patients	Proxy
Badia et al.(1998)	120	360^{*}	0.09^{Γ}	VAS	Chronically ill	General public patients
Boyd et al. (1990)	40	99	0.09(NR)	VAS	Patients with colostomy	Professionals
			0.01(NR)	SG		
			0.23(NR)	VAS		General public
			0.09(NR)	SG		
Calhoun et al. (2004)	40	50	-0.01(NR)	TTO	Ovarian cancer patients	Professionals
			0.03(NR)	TTO		General public
Cappelli et al.(2001)	57	46	0.08^{Π}	VAS	Women with breast cancer	General public
			0.11^{Π}	SG		
Chen et al. (1998)	14	41	$-0.10^{\dagger}$	VAS	Patients with psoriasis	Professionals
			-0.02(NR)	VAS		General public

Source	Kind scenario	Illness label	Country
Ashby & O'Hanlon (1994)	6 breast cancer health state scenario's	Yes	UK
Badia et al. (1996)	15 EQ-5D health states	No	Spain
Badia et al.(1998)	15 EQ-5D health state scenario's	No	Spain
Boyd et al. (1990)	Scenario living with a colostomy	Yes	Canada
Calhoun et al. (2004)	6 chemotherapeutic-induced toxicity health states	Yes	USA
Cappelli et al. (2001)	6 health state descriptions related to breast cancer	Yes	Canada
Chen et al. (1998)	3 psoriasis health state descriptions	Yes	USA
Clarke et al. (1997)	3 Gaucher Disease scenario's	Yes	Canada & USA

Source	♯patients	♯non-patients	Unstandardized difference	Elicitation method	Kind Patients	Kind non-patients
Clarke et al. (1997)	50	39	-0.03^{I}	TTO	Chronic ill patients	General public
			-0.04^{I}	SG		
			-0.02^{I}	TTO	Gaucher disease patients	General public
			-0.01^{I}	SG		
Cunningham et al. (2000)	40	55	-0.05(NS)	TTO	Orthognathic patients	General public
			0.05(NS)	VAS		
			-0.06(NS)	SG		
De Wit et al.(2000)	152	103	0.19‡	TTO	Dialysis patients	General public
			0.04‡	SG		
Gabriel et al. (1999)	44	199	0.39‡	TTO	Women with osteoporotic fractures	General public
Goldberg (2006)	173	128	0.00(NR)	SG	Prostate cancer	General public
			-0.03 NR	SG	Diabetes	General public
Grann et al. (1999)	20	125	0.04(NR)	TTO	Breast cancer patients	Professionals
			0.12(NR)	VAS		
			0.01(NR)	TTO		
			0.06(NR)	VAS		
Hallan et al. (1999)	41	117	0.15(NR)	SG	Stroke survivors	General public
Happich & Lengerke (2005)	105	66	0.06†	TTO	Patients affected with tinnitus	General public
			0.19‡	VAS		
			0.08‡	SG		
Hayman (1997)	97	20	0.02(NR)	SG	Early-stage breast cancer	Professionals

Source	Kind scenario	Illness label	Country
Cunningham et al. (2000)	One description with pictures of dental problems	Yes	UK
De Wit et al.(2000)	3 EQ-5D scenario's	No	The Netherlands
Gabriel et al. (1999)	3 health state descriptions, disabling hip fracture,established osteoporosis and multiple vertebral fractures	Yes	USA
Goldberg (2006)	4 Erectile dysfunction descriptions	Yes	USA
Grann et al. (1999)	8 Cancer-related health states	Yes	USA
Hallan et al. (1999)	2 scenario's describing functional outcomes after stroke	Yes	Norway
Happich & Lengerke (2005)	Description of Tinnitus	Yes	Germany
Hayman (1997)	5 Breast-cancer related health states	Yes	USA

Source	♯patients	♯non-patients	Unstandardized difference	Elicitation method	Kind Patients	Kind non-patients
Hayman (2005)	120	210	-0.01 ‡	SG	Women with ductal carcinoma-in-situ	General public
Korfage et al. (2007)	53	52	-0.02(NS) 0.02^{I}	TTO VAS	Men with prostate cancer	General public
Molzahn (1997)	215	49	0.07‡	TTO	End stage renal disease	Professionals
Naraine et al. (2002)	28	30	0.02†	SG	Patients with haemophilia	General public
Postulart & Adang (2000)	22	54	0.25‡	VAS	Insulin-dependent diabetes mellitus with end-stage renal disease	General public
Prosser et al. (2003)	29	33	0.07‡	SG	MS patients	General public
Smith et al.2006	95	567	0.21‡	TTO	Colostomy patients	General public
Souchek et al. (2005)	66	64	0.02(NS) 0.00(NS) -0.07(NS)	TTO VAS SG	Patients with osteoarthritis	General public
Sackett & Torrance (1978)	29	246	0.132$^{\Pi}$	TTO	Dialysis patients	General public
Stewart (2005)	84	78	0.03(NR) 0.01(NR) 0.05(NR)	TTO VAS SG	Men with prostate cancer	General public
Suarez -Almazor & Conner-	50	147	0.02(NR) -0.00(NR) -0.05(NR)	TTO VAS SG	Patients with rheumatoid	Professionals

Source	Kind scenario	Illness label	Country
Hayman (2005)	8 breast cancer related health states	Yes	USA
Molzahn (1997)	Life on dialysis	Yes	Canada
Naraine et al. (2002)	7 haemophilia A health state descriptions	Yes	Canada
Postulart & Adang (2000)	Self vs. Description of Insulin-dependent diabetes mellitus and end-stage renal disease	Yes	The Netherlands
Prosser et al. (2003)	6 MS health state descriptions	No	USA
Smith et al. (2006)	Colostomy health state description	Yes	USA
Souchek et al. (2005)	2 EQ-5D health states	Yes	USA
Sackett & Torrance (1978)	5 Dialysis health state descriptions	Yes	Canada

Source	♯patients	♯non-patients	Unstandardized difference	Elicitation method	Kind Patients	Kind non-patients
Spady (2001)			-0.00(NR)	TTO		General public
			-0.02 (NR)	VAS		
			0.07 (NR)	SG		
Suarez-Almazor (2001a)	105	5	-0.16‡	VAS	Musculoskeletal diseases	Professionals
			-0.19‡	SG		
Van der Donk et al. (1995)	20	20	-0.11(NR)	TTO	Laryngeal cancer	Professionals
			-0.08(NR)	VAS		
			-0.23(NR)	SG		
			-0.11 (NR)	TTO	Laryngeal cancer	General public
			-0.03 (NR)	VAS		
			-0.13 (NR)	SG		
			-0.12 (NR)	TTO	Floor of the mouth cancer	Professionals
			-0.05 (NR)	VAS		
			-0.18 (NR)	SG		
			-0.12 (NR)	TTO	Floor of the mouth cancer	General public
			0.04 (NR)	VAS		
			-0.07 (NR)	SG		
Wells et al. (2004)	114	83	0.14‡	TTO	Cirrhotics	Professionals

Source	Kind scenario	Illness label	Country
Stewart (2005)	21 prostate cancer related health states	Yes	USA
Suarez -Almazor & Conner - Spady	2 scenario's of arthritis	Yes	Canada
Suarez-Almazor (2001)	Own experienced health of patients seen by Rheumatologist	Yes	Canada
Van der Donk et al. (1995)	4 health states after T3-laryngeal cancer	Yes	The Netherlands
Wells et al. (2004)	6 scenarios of cirrhosis	Yes	USA

[*] Original sample size, in analyses the sample sizes were divided by two since in both studies valuations of one group of non-patients were used.

NR Significance of the difference between respondent groups not reported

NS Difference between respondent groups is not significant

[†] Difference between respondent groups $p < 0.05$

[‡] Difference between respondent groups $p < 0.01$

[I] Difference between respondent groups varied for the different health states between not significant and $p < 0.05$

[Π] Difference between respondent groups varied for the different health states between not significant and $p < 0.01$

[Γ] Difference between respondent groups varied for the different health states between $p < 0.05$ and $p < 0.01$

Appendix B

Description of the five additional EQ-5D scenarios and RA-patients' valuations(Chapter 3)

EQ-5D dimension	Scenario A	Scenario B	Scenario C	Scenario D	Scenario F
Mobility	Some problems	Some problems	Some problems	Some problems	Some problems
Self-care	No problems	Some problems	No problems	No problems	No problems
Usual activities	Some problems	Unable to perform	Unable to perform	Some problems	No problems
Pain/discomfort	Extreme pain or other discomfort	Some pain or other discomfort	Some pain or other discomfort	Some pain or other discomfort	No pain or other discomfort
Anxiety/ depression	Moderately anxious /depressed	Moderately anxious /depressed	Not anxious /depressed	Moderately anxious /depressed	Not anxious /depressed
N	124	126	127	120	126
Mean (SD)	0.53 (0.30)	0.57 (0.30)	0.69 (0.25)	0.76 (0.24)	0.91 (0.14)

Appedix C

RA health state description

Imagine that your have been diagnosed with rheumatoid arthritis. This means that:

- some of your joints in your fingers and feet are swollen

- you have pain due to mild inflation in your tendons

- you have difficulty performing fine motor skills with your hands

- you have sore feet after walking more than half an hour

- you have some problems with your self-care, such as brushing hair, tying knots, washing your back, opening certain packages, and with cooking, e.g. with peeling potatoes

- you have some difficulty in using a computer mouse or keyboard, and with lifting items onto a shelf above your head at work

- you are not being able to run as a form of sports

Appendix D

Coding system chapter 4

Physical (in)convenience RA-related

Physical inconvenience that is obviously related to RA such as stiffness, medication or pain. Taking a rest, physical state, and fatigue were also included as well as undesirable effects of medication. Other aspects included were; the RA was improved, that momentarily the RA did not give problems, constraints or pain, or that it is possible to find solutions.

Illness/inconvenience not RA-related

Illness or inconvenience that is not RA-related, such as headache or kidney problems. Losing weight and medication or undesirable effects of medication are also included if these were not RA-related.

Family

Family or activities in which family was important, such as (grand)children or party of a son. Family members who are ill are also included.

Other social contacts

Social contacts such as friends, colleagues or acquaintances excluding family members.

Emotions/Worries

All emotions and worries and aspects in which the emotion or worry was most important. Emotions due to social contacts like conflict with a child and insomnia due to emotions, are included; as well as worries due to medication, the future or growing old. Hope, satisfaction, contentment, and rejoicing also belong to this category.

Sports/Mobility

Sport activities and mobility, such as playing soccer, swimming, sitting inside, constraints in walking and cycling around. Participants sometimes named aspects as walking or cycling as their sport, sometimes as (limitation) in their mobility and sometimes without referring to one of both. All aspects were included in this domain, therefore no distinction could be made between sports & mobility.

Leisure activities

Leisure activities, such as sewing or reading. When bicycling is named as a pleasant, relaxing and social activity it is included as leisure activity instead of a sport.

Work

All work related aspects. Housekeeping is also included.

Broad every day life

Performance of activities of daily living, such as I am able to do what I want, I only need to adjust my tempo and am able to do the things someone does in every day life.

Others

Dimensions that could not be included in any of the categories or that can be included in more than one category.

References

[1] GGD rivierenland. Bofvirus in aantal nederlandse studentensteden, 2010.

[2] InfoNu.nl. Arthrose: Nieuwe kraakbeentransplantatietechniek, 2010.

[3] M.F. Drummond, G.L. Stoddart, and G.W. Torrance. Critical assessment of economic evaluation. In *Methods for the Economic Evaluation of Helath Care Programmes*, number 3, pages 18–38. Oxford Medical publications, Oxford, 1990.

[4] A.M. Stiggelbout and J.C. de Haes. Patient preference for cancer therapy: an overview of measurement approaches. *Journal of Clinical Oncology*, 19(1):220–230, January 2001.

[5] C. Harries and A.M. Stiggelbout. Approaches to measuring patients' decision-making. In Ebrahim S.(eds.) Bowling, A., editor, *Handbook of Health Research Methods: Investigation, Measurement and Analysis.*, number 16, pages 362–393. Open University/McGraw Hill, Maidenhead, 2005.

[6] Morgenstern O. Neumann J.von. *Theory of games and economic behavior (First Edition ed.).* Princeton: Princeton University Press, 1947.

[7] G.W. Torrance. Social preferences for health states - empirical evaluation of 3 measurement techniques. *Socio-Economic Planning Sciences*, 10(3):129–136, 1976.

[8] P. Wakker and A.M. Stiggelbout. Explaining distortions in utility elicitation through the rank-dependent model for risky choices. *Medical Decision Making*, 15(2):180–186, April 1995.

[9] H. Bleichrodt. A new explanation for the difference between time tradeoff utilities and standard gamble utilities. *Health Economics*, 11(5):447–456, July 2002.

[10] E. Nord. Methods for quality adjustment of life years. *Social Science & Medicine*, 34(5):559–569, March 1992.

[11] P. Dolan. Modeling valuations for euroqol health states. *Medical Care*, 35(11):1095–1108, November 1997.

[12] M.R. Gold, J.E. Siegel, L.B. Russell, and M.C. Weinstein. *Cost-effectiveness in health and medicine.* Oxford University Press Inc., New York, 1996.

[13] Luijn JCF Wolff I. Riteco JA, Heij LJMD. Richtlijnen voor farmaco-economisch onderzoek. rapport uigebracht aan de minister van volksgezondheid, welzijn en sport. amstelveen. Technical report, College voor Zorgverzekeringen, 1999.

[14] National Institue for Clinical Excellence. Guide to the methods of techonology appraisal. Technical report, NICE, 2004.

[15] A.M. Stiggelbout and E. Vogel-Voogt. Health state utilities: A framework for studying the gap between the imagined and the real. *Value in Health*, 11(1):76–87, January 2008.

[16] G.A. De Wit, J.J.V. Busschbach, and F.T. De Charro. Sensitivity and perspective in the valuation of health status: Whose values count? *Health Economics*, 9(2):109–126, March 2000.

[17] P. Dolan. Whose preferences count? *Medical Decision Making*, 19(4):482–486, October 1999.

[18] P.A. Ubel, G. Loewenstein, and C. Jepson. Whose quality of life? a commentary exploring discrepancies between health state evaluations of patients and the general public. *Quality of Life Research*, 12(6):599–607, September 2003.

[19] D. Kahneman and A.B. Krueger. Developments in the measurement of subjective well-being. *Journal of Economic Perspectives*, 20(1):3–24, 2006.

[20] S.J. Jansen, A.M. Stiggelbout, P.P. Wakker, M.A. Nooij, E.M. Noordijk, and J. Kievit. Unstable preferences: a shift in valuation or an effect of the elicitation procedure? *Medical Decision Making*, 20(1):62–71, January 2000.

[21] J. Horsman, W. Furlong, D. Feeny, and G. Torrance. The health utilities index (hui(r)): concepts, measurement properties and applications. *Health and Quality of Life Outcomes*, 1 (54)(1):–, October 2003.

[22] P. Dolan, C. Gudex, P. Kind, and A. Williams. A social tariff for euroqol: Results from a uk general population survey. Technical report, York, York, 1995.

[23] R.P. Insinga and D.G. Fryback. Understanding differences between self-ratings and population ratings for health in the euroqol. *Quality of Life Research*, 12(6):611–619, September 2003.

[24] P.A. Ubel, G. Loewenstein, and C. Jepson. Disability and sunshine: can hedonic predictions be improved by drawing attention to focusing illusions or emotional adaptation? *Journal of Experimental Psychology Applied*, 11(2):111–123, June 2005.

[25] P.A. Ubel, G. Loewenstein, J. Hershey, J. Baron, T. Mohr, D.A. Asch, and C. Jepson. Do nonpatients underestimate the quality of life associated with chronic health conditions because of a focusing illusion? *Medical Decision Making*, 21(3):190–199, May 2001.

[26] D.T. Gilbert, E.C. Pinel, T.D. Wilson, S.J. Blumberg, and T.P. Wheatley. Immune neglect: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, 75(3):617–638, September 1998.

[27] T.D. Wilson, T. Wheatley, J.M. Meyers, D.T. Gilbert, and D. Axsom. Focalism: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, 78(5):821–836, May 2000.

[28] M.F. Luce. Decision making as coping. *Health Psychology*, 24(4):S23–S28, July 2005.

[29] G. Salkeld, M. Ryan, and L. Short. The veil of experience. do consumers prefer what they know best? *Health Economics*, 9(3):267–270, April 2000.

[30] Tversky A. and Kahneman D. Loss aversion in riskless choice: a reference-dependent model. *The Quarterly Journal of Economics*, 106(4):1039–1061, 1991.

[31] S. Frederick and G. Loewenstein. Hedonic adaptation. In D. Kahneman, E. Diener, and N. Schwarz, editors, *Well-Being The foundations of hedonic psychology*, number 16, pages 302–329. Russell sage Foundation, New York, 1999.

[32] S.E. Taylor. Adjustment to threatening events - a theory of cognitive adaptation. *American Psychologist*, 38(11):1161–1173, 1983.

[33] P. Brickman, D. Coates, and R. Janoffbulman. Lottery winners and accident victims - is happiness relative. *Journal of Personality and Social Psychology*, 36(8):917–927, 1978.

[34] E. Diener, R.E. Lucas, and C.N. Scollon. Beyond the hedonic treadmill - revising the adaptation theory of well-being. *American Psychologist*, 61(4):305–314, May 2006.

[35] P.A. Ubel. *You're stronger than you think*, volume 1. The McGraw-Hill companies, New York, 2006.

[36] P. Dolan and D. Kahneman. Interpretations of utility and their implications for the valuation of health. *Economic Journal*, 118(525):215–234, January 2008.

[37] P.A. Ubel, G. Loewenstein, N. Schwarz, and D.M. Smith. Misimagining the unimaginable: The disability paradox and health care decision making. *Health Psychology*, 24(4):S57–S62, July 2005.

[38] L.J. Damschroder, B.J. Zikmund-Fisher, and P.A. Ubel. The impact of considering adaptation in health state valuation. *Social Science & Medicine*, 61(2):267–277, July 2005.

[39] L.J. Damschroder, B.J. Zikmund-Fisher, and P.A. Ubel. Considering adaptation in preference elicitations. *Health Psychology*, 27(3):394–399, May 2008.

[40] P. Dolan. The effect of experience of illness on health state valuations. *Journal of Clinical Epidemiology*, 49(5):551–564, May 1996.

[41] M.A. Sprangers and C.E. Schwartz. Integrating response shift into health-related quality of life research: a theoretical model. *Social Science & Medicine*, 48(11):1507–1515, June 1999.

[42] M. Ross. Relation of implicit theories to the construction of personal histories. *Psychological Review*, 96(2):341–357, April 1989.

[43] H.P. Lacey, D. Smith, and P.A. Ubel. Hope i die before i get old: mispredicting happiness across the adult lifespan. *Journal of Happiness Studies*, 7:167–182, 2006.

[44] M.G.T. Dolders, M.P.A. Zeegers, W. Groot, and A. Ament. A meta-analysis demonstrates no significant differences between patient and population preferences. *Journal of Clinical Epidemiology*, 59(7):653–664, July 2006.

[45] S.E. Taylor, M.E. Kemeny, G.M. Reed, J.E. Bower, and T.L. Gruenewald. Psychological resources, positive illusions, and health. *American Psychologist*, 55(1):99–109, January 2000.

[46] K.E. Bremner, C.A. Chong, G. Tomlinson, S.M. Alibhai, and M.D. Krahn. A review and meta-analysis of prostate cancer utilities. *Medical Decision Making*, 27(3):288–298, May 2007.

[47] D. Feeny, W. Furlong, G.W. Torrance, C.H. Goldsmith, Z. Zhu, S. DePauw, M. Denton, and M. Boyle. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. *Medical Care*, 40(2):113–128, February 2002.

[48] M.W. Lipsey and D.B. Wilson. *Practival meta-analysis.* Number 49. Sage Puclications, Inc, California, 2001.

[49] D. Smith, R.L. Sherriff, L. Damschroder, G. Loewenstein, and P.A. Ubel. Misremembering colostomies? former patients give lower utility ratings than do current patients. *Health Psychology*, 25(6):688–695, November 2006.

[50] J.E Hunter and F.L. Schmidt. *Methods of meta-analysis correcting error and bias in research findings.* Newbury Park, CA: Sage, 1990.

[51] M. Borenstein, L. Hedges, J. Higgins, and H. Rothstein. Comprehensive meta analysis, 2007.

[52] L.V. Hedges and I. Olkin. *Statiscal methods for meta-analysis.* Academic Press, Orlando, FL, 1985.

[53] S. Duval and R. Tweedie. A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, 95(449):89–98, March 2000.

[54] A.S. Elstein and G.B. Chapman. Patients' values and clinical substituted judgments: The case of localized prostate cancer. *Health Psychology*, 24(4):S85–S92, July 2005.

[55] J.F. O'Leary, D.L. Fairclough, M.K. Jankowski, and J.C. Weeks. Comparison of time-tradeoff utilities and rating scale values of cancer patients and their relatives: evidence for a possible plateau relationship. *Medical Decision Making*, 15(2):132–137, April 1995.

[56] E.A. Calhoun, C.L. Bennett, P.A. Peeples, J.R. Lurain, P.Y. Roland, J.M. Weinstein, and D.A. Fishman. Perceptions of cisplatin-related toxicity among ovarian cancer patients and gynecologic oncologists. *Gynecologic Oncology*, 71(3):369–375, December 1998.

[57] M. Happich and B. Mazurek. Priorities and prospect theory. *European Journal of Health Economics*, 3(1):40–46, 2002.

[58] X. Badia, M. Herdman, and P. Kind. The influence of ill-health experience on the valuation of health. *Pharmacoeconomics*, 13(6):687–696, June 1998.

[59] X. Badia, A. Diaz-Prieto, M. Rue, and D.L. Patrick. Measuring health and health state preferences among critically ill patients. *Intensive Care Medicine*, 22(12):1379–1384, December 1996.

[60] J.H. Goldberg. Being there is important, but getting there matters too: the role of path in the valuation process. *Medical Decision Making*, 26(4):323–337, July 2006.

[61] S. Hallan, A. Asberg, B. Indredavik, and T.E. Wideroe. Quality of life after cerebrovascular stroke: a systematic study of patients' preferences for different functional outcomes. *Journal of Internal Medicine*, 246(3):309–316, September 1999.

[62] V.S. Naraine, N.A. Risebrough, P. Oh, V.S. Blanchette, S. Lee, A.M. Stain, D. Hedden, J.M. Teitel, and B.M. Feldman. Health-related quality-of-life treatments for severe haemophilia: utility measurements using the standard gamble technique. *Haemophilia*, 8(2):112–120, March 2002.

[63] S.T. Stewart, L. Lenert, V. Bhatnagar, and R.M. Kaplan. Utilities for prostate cancer health states in men aged 60 and older. *Medical Care*, 43(4):347–355, April 2005.

[64] A.N.A. Tosteson, S.E. Gabriel, T.S. Kneeland, M.M. Moncur, P.D. Manganiello, I. Schiff, B. Ettinger, and L.J. Melton. Has the impact of hormone replacement therapy on health-related quality of life been undervalued? *Journal of Womens Health & Gender-Based Medicine*, 9(2):119–130, March 2000.

[65] J. Hall, K. Gerard, G. Salkeld, and J. Richardson. A cost utility analysis of mammography screening in australia. *Social Science & Medicine*, 34(9):993–1004, May 1992.

[66] M.E. Suarez-Almazor and B. Conner-Spady. Rating of arthritis health states by patients, physicians, and the general public. implications for cost-utility analyses. *Journal of Rheumatology*, 28(3):648–656, March 2001.

[67] M.E. Suarez-Almazor, B. Conner-Spady, C.J. Kendall, A.S. Russell, and K. Skeith. Lack of congruence in the ratings of patients' health status by patients and their physicians. *Medical Decision Making*, 21(2):113–121, March 2001.

[68] J.A. Hayman, D.L. Fairclough, J.R. Harris, and J.C. Weeks. Patient preferences concerning the trade-off between the risks and benefits of routine radiation therapy after conservative surgery for early-stage breast cancer. *Journal of Clinical Oncology*, 15(3):1252–1260, March 1997.

[69] J.A. Hayman, M.U. Kabeto, M.J. Schipper, J.E. Bennett, F.A. Vicini, and L.J. Pierce. Assessing the benefit of radiation therapy after breast-conserving surgery for ductal carcinoma-in-situ. *Journal of Clinical Oncology*, 23(22):5171–5177, August 2005.

[70] A.E. Molzahn, H.C. Northcott, and J.B. Dossetor. Quality of life of individuals with end stage renal disease: perceptions of patients, nurses, and physicians. *Annals Journal*, 24(3):325–333, June 1997.

[71] V.R. Grann, J.S. Jacobson, V. Sundararajan, S.M. Albert, A.B. Troxel, and A.I. Neugut. The quality of life associated with prophylactic treatments for women with brca1/2 mutations. *Cancer Journal From Scientific American*, 5(5):283–292, September 1999.

[72] L.A. Prosser, K.M. Kuntz, A. Bar-Or, and M.C. Weinstein. Patient and community preferences for treatments and health states in multiple sclerosis. *Multiple Sclerosis*, 9(3):311–319, June 2003.

[73] S.J. Cunningham and N.P. Hunt. A comparison of health state utilities for dentofacial deformity as derived from patients and members of the general public. *European Journal of Orthodontics*, 22(3):335–342, June 2000.

[74] J. Vanderdonk, P.C. Levendag, A.J. Kuijpers, F.H.J. Roest, J.D.F. Habbema, C.A. Meeuwis, and P.I.M. Schmitz. Patient participation in clinical decision-making for treatment of t3 laryngeal-cancer - a comparison of state and process utilities. *Journal of Clinical Oncology*, 13(9):2369–2378, September 1995.

[75] M. Happich and T. von Lengerke. Valuing the health state ’tinnitus’: Differences between patients and the general public. *Hearing Research*, 207(1-2):50–58, September 2005.

[76] S.E. Gabriel, T.S. Kneeland, L.J. Melton, M.M. Moncur, B. Ettinger, and A.N.A. Tosteson. Health-related quality of life in economic evaluations for osteoporosis: Whose values should we use? *Medical Decision Making*, 19(2):141–148, April 1999.

[77] M. Cappelli, L. Surh, L. Humphreys, S. Verma, D. Logan, A. Hunter, and J. Allanson. Measuring women's preferences for breast cancer treatments and brca1/brca2 testing. *Quality of Life Research*, 10(7):595–607, 2001.

[78] S. Chen, A. Shaheen, and A. Garber. Cost-effectiveness and cost-benefit analysis of using methotrexate vs goeckerman therapy for psoriasis. a pilot study. *Archives of Dermatology*, 134(12):1602–1608, December 1998.

[79] J. Ashby, M. Ohanlon, and M.J. Buxton. The time trade-off technique - how do the valuations of breast-cancer patients compare to those of other groups. *Quality of Life Research*, 3(4):257–265, August 1994.

[80] J. Souchek, M.M. Byrne, P.A. Kelly, K. O' Malley, M. Richardson, C. Pak, H. Nelson, and M.E. Suarez-Almazor. Valuation of arthritis health states across ethnic groups and between patients and community members. *MedicalCare*, 43(9):921–928, September 2005.

[81] I.J. Korfage, H.J. de Koning, J.D.F. Habbema, F.H. Schroder, and M.L. Essink-Bot. Side-effects of treatment for localized prostate cancer: are they valued differently by patients and healthy controls? *British Journal of Urological International*, 99(4):801–806, April 2007.

[82] E.A. Calhoun, D.A. Fishman, J.R. Lurain, E.E. Welshman, and C.L. Bennett. A comparison of ovarian cancer treatments: analysis of utility assessments of ovarian cancer patients, at-risk population, general population, and physicians. *Gynecologic Oncology*, 93(1):164–169, April 2004.

[83] C.D. Wells, W.B. Murrill, and M.R. Arguedas. Comparison of health-related quality of life preferences between physicians and cirrhotic patients: Implications for cost-utility analyses in chronic liver disease. *Digestive Diseases and Sciences*, 49(3):453–458, March 2004.

[84] A.E. Clarke, M.K. Goldstein, D. Michelson, A.M. Garber, and L.A. Lenert. The effect of assessment method and respondent population on utilities elicited for gaucher disease. *Quality of Life Research*, 6(2):169–184, March 1997.

[85] D.L. Sackett and G.W. Torrance. The utility of different health states as perceived by the general public. *Journal of Chronic Diseases*, 31(11):697–704, 1978.

[86] S.M. van Osch. *Construction of TTO utilities*. PhD thesis, Leiden University, 2007.

[87] R. Bender, C. Bunce, M. Clarke, S. Gates, S. Lange, N.L. Pace, and K. Thorlund. Attention should be given to multiplicity issues in systematic reviews. *Journal of Clinical Epidemiology*, 61(9):857–865, September 2008.

[88] D. Polsky, R.J. Willke, K. Scott, K.A. Schulman, and H.A. Glick. A comparison of scoring weights for the euroqol (c) derived from patients and the general public. *Health Economics*, 10(1):27–37, January 2001.

[89] R. Brooks. Euroqol: The current state of play. *Health Policy*, 37(1):53–72, July 1996.

[90] L.S. Matza, K.S. Boye, N. Yurgin, J. Brewster-Jordan, S. Mannix, J.M. Shorr, and B.L. Barber. Utilities and disutilities for type 2 diabetes treatment-related attributes. *Quality of Life Research*, 16(7):1251–1265, September 2007.

[91] H.A. Llewellyn-Thomas, E.C. Thiel, and M.J. McGreal. Cancer patients' evaluations of their current health states: the influences of expectations, comparisons, actual health status, and mood. *Medical Decision Making*, 12(2):115–122, April 1992.

[92] M. Hagedoorn, S.G. Uijl, E. Van Sonderen, A.V. Ranchor, B.M.F. Grol, R. Otter, B. Krol, W. Van den Heuvel, and R. Sanderman. Structure and reliability of ware's patient satisfaction questionnaire iii - patients' satisfaction with oncological care in the netherlands. *Medical Care*, 41(2):254–263, February 2003.

[93] M. Rosenberg. *Society and the Adolescent Self-Image.*, volume Revised edition. CT: Wesleyan University Press, Middletown, 1989.

[94] Ci3. Ci3 skimsoftware division rotterdam. Technical report, Sawtooth-Software, USA, 2000.

[95] J.F. Fries, P. Spitz, R.G. Kraines, and H.R. Holman. Measurement of patient outcome in arthritis. *Arthritis and Rheumatism*, 23(2):137–145, 1980.

[96] H.P. Lacey, A. Fagerlin, G. Loewenstein, D. Smith, J. Riis, and P.A. Ubel. It must be awful for them: Perspective and task context affects ratings for health conditions. *Judgement and Decision Making*, 1(2):146–152, 2007.

[97] S.J. Walters and J.E. Brazier. Comparison of the minimally important difference for two health state utility measures: Eq-5d and sf-6d. *Quality of Life Research*, 14(6):1523–1532, August 2005.

[98] D.C. Hadorn. The role of public values in setting health-care priorities. *Social Science & Medicine*, 32(7):773–781, 1991.

[99] D.G. Froberg and R.L. Kane. Methodology for measuring health-state preferences .3. population and context effects. *Journal of Clinical Epidemiology*, 42(6):585–592, 1989.

[100] Y. Peeters and A.M. Stiggelbout. Health state valuations of patients and the general public analytically compared: A meta-analytical comparison of patient and population health state utilities. *Value in Health*, 13(2):306–309, March 2010.

[101] M.F. Janssen, E. Birnie, J.A. Haagsma, and G.J. Bonsel. Comparing the standard eq-5d three-level system with a five-level version. *Value in Health*, 11(2):275–284, March 2008.

[102] P.F. Krabbe, M.E. Stouthard, M.L. Essink-Bot, and G.J. Bonsel. The effect of adding a cognitive dimension to the euroqol multiattribute health-status classification system. *Journal of Clinical Epidemiology*, 52(4):293–301, April 1999.

[103] R.A. Carr-Hill. Health related quality of life measurement–euro style. *Health Policy*, 20(3):321–328, 1992.

[104] C. Gudex. Are we lacking a dimensions of energy in the euroqol instrument? In S. Bjork, editor, *EuroQol conference proceedings: Lund October 1991*, number 6, pages 61–81. The Swedish Institute for Health Economics, Lund, 1992.

[105] The EuroQol Group. Euroqol–a new facility for the measurement of health-related quality of life. the euroqol group. *Health Policy*, 16(3):199–208, December 1990.

[106] Rigby AS. Statistical methods in epidemiology. v. towards an understanding of the kappa coefficient. *Disability and Rehabilitation 2000 May 20;22(8):339-344.*, 22(8):339–344, 2000.

[107] J. Riis, G. Loewenstein, J. Baron, and C. Jepson. Ignorance of hedonic adaptation to hemodialysis: A study using ecological momentary assessment. *Journal of Experimental Psychology-General*, 134(1):3–9, February 2005.

[108] N.D. Weinstein. Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, 39(5):806–820, 1980.

[109] M.C. Lennon, B.P. Dohrenwend, A.J. Zautra, and J.J. Marbach. Coping and adaptation to facial pain in contrast to other stressful life events. *Journal of Personality and Social Psychology*, 59(5):1040–1050, November 1990.

[110] Hilgard ER Ruch JC Lange AF Lenox JR Morgan AH and Sachs LB. The psychophysics of cold pressor pain and its modification through hypnotic suggestion. *American Journal of Psychology*, 87:93–105, 1974.

[111] N. Schwarz. Self-reports - how the questions shape the answers. *American Psychologist*, 54(2):93–105, February 1999.

[112] D. M. Smith, N. Schwarz, T.R. Roberts, and P.A. Ubel. Why are you calling me? how study introductions change response patterns. *Quality of Life Research*, 15(4):621–630, May 2006.

[113] Centraal Bureau voor de Statistiek. Zelfgerapporteerde medische consumptie, gezondheid en leefstijl. Technical report, CBS, 2010.

[114] P.F.M. Stalmeier, M.K. Goldstein, A.M. Holmes, L. Lenert, J. Miyamoto, A.M. Stiggelbout, G.W. Torrance, and J. Tsevat. What should be reported in a methods section on utility assessment? *Medical Decision Making*, 21(3):200–207, May 2001.

[115] Y. Peeters and A.M. Stiggelbout. Valuing health: does enriching a scenario lead to higher utilities? *Medical Decision Making*, 29(3):334–342, May 2009.

[116] N.F. Boyd, H.J. Sutherland, K.Z. Heasman, D.L. Tritchler, and B.J. Cummings. Whose utilities for decision-analysis. *Medical Decision Making*, 10(1):58–67, January 1990.

[117] Centraal Bureau voor de Statistiek CBS. Overlevingstafels naar leeftijd en geslacht 2005. Technical report, Centraal Bureau voor de Statistiek, September 2007.

[118] M.L. Essink-Bot. Beschrijvend kwaliteit van leven-onderzoek. In M.P.M.H. Rutten-van Molken, J.J.V. Busschbach, and F.F.H. Rutten, editors, *Van Kosten tot Effecten*, number Hoofdstuk 5, pages 63–72. Elsevier gezondheidszorg, Maarsen, 2000.

[119] J. Tsevat. What do utilities measure? *Medical Care*, 38(9):160–164, September 2000.

[120] R. Arnold. *Quality of life in chronic obstructive pulmonary disease and chronic heart failure*. PhD thesis, Rijksuniversiteit Groningen, 2004.

[121] R.E. Lucas. Adaptation and the set-point model of subjective well-being - does happiness change after major life events? *Current Directions in Psychological Science*, 16(2):75–79, April 2007.

[122] J.E. Sherman, D.J. DeVinney, and K.B. Sperling. Social support and adjustment after spinal cord injury: Influence of past peer-mentoring experiences and current live-in partner. *Rehabilitation Psychology*, 49(2):140–149, May 2004.

[123] E. Martz, H. Livneh, M. Priebe, L.A. Wuermser, and L. Ottomanelli. Predictors of psychosocial adaptation among people with spinal cord injury or disorder. *Archives of Physical Medicine and Rehabilitation*, 86(6):1182–1192, June 2005.

[124] J.R. Rodrigue, W.F. Kanasky Jr, S.I. Jackson, and M.G. Perri. The psychosocial adjustment to illness scale–self-report: factor structure and item stability. *Psycholigy Assessments*, 12(4):409–413, December 2000.

[125] V.S. Helgeson and H.L. Fritz. Cognitive adaptation as a predictor of new coronary events after percutaneous transluminal coronary angioplasty. *Psychosomatic Medicine*, 61(4):488–495, July 1999.

[126] V.S. Helgeson. Applicability of cognitive adaptation theory to predicting adjustment to heart disease after coronary angioplasty. *Health Psychology*, 18(6):561–569, November 1999.

[127] I.W.S. Chan, J.C.L. Lai, and K.W.N. Wong. Resilience is associated with better recovery in chinese people diagnosed with coronary heart disease. *Psychology & Health*, 21(3):335–349, June 2006.

[128] T. Moore, P. Norman, P.R. Harris, and M. Makris. Cognitive appraisals and psychological distress following venous thromboembolic disease: An application of the theory of cognitive adaptation. *Social Science & Medicine*, 63(9):2395–2406, November 2006.

[129] H.E. Stiegelis, M. Hagedoorn, R. Sanderman, K.I. van der Zee, B.R. Buunk, and M.C.M. van den Bergh. Cognitive adaptation: A comparison of cancer patients and healthy references. *British Journal of Health Psychology*, 8:303–318, September 2003.

[130] V.S. Helgeson. Cognitive adaptation, psychological adjustment, and disease progression among angioplasty patients: 4 years later. *Health Psychology*, 22(1):30–38, January 2003.

[131] J.R. Ickovics, S. Milan, R. Boland, E. Schoenbaum, P. Schuman, and D. Vlahov. Psychological resources protect health: 5-year survival and immune function among hiv-infected women from four us cities. *Aids*, 20(14):1851–1860, September 2006.

[132] C.F. Ratelle, R.J. Vallerand, Y. Chantal, and P. Provencher. Cognitive adaptation and mental health: A motivational analysis. *European Journal of Social Psychology*, 34(4):459–476, July 2004.

[133] B. Major, C. Richards, M.L. Cooper, C. Cozzarelli, and J. Zubek. Personal resilience, cognitive appraisals, and coping: An integrative model of adjustment to abortion. *Journal of Personality and Social Psychology*, 74(3):735–752, March 1998.

[134] P.L. Tomich and V.S. Helgeson. Cognitive adaptation theory and breast cancer recurrence: Are there limits? *Journal of Consulting and Clinical Psychology*, 74(5):980–987, October 2006.

[135] J.D. Creswell, W.T. Welch, S.E. Taylor, D.K. Sherman, T.L. Gruenewald, and T. Mann. Affirmation of personal values buffers neuroendocrine and psychological stress responses. *Psychological Science*, 16(11):846–851, November 2005.

[136] B. Spilker and D.A. Revicki. Taxonomy of quality of life. In P. Spinhoven, editor, *Quality of Life and Pharmacoeconomics in Clinical Trails*, volume Second Edition, pages 25–32. Lippincott-Raven Publishers, Philadelphia, 1996.

[137] A.D. Wu and B.D. Zumbo. Understanding and using mediators and moderators. *Social Indicators Research*, 87(3):367–392, July 2008.

[138] E.J. Hurkmans, S. Maes, V. De Gucht, K. Knittle, A.J. Peeters, H.K. Ronday, and T.P.M. Vliet Vlieland. Motivation as a determinant of physical activity in patients with rheumatoid arthritis. *Arthritis Care & Research*, 62(3):371–377, March 2010.

[139] A.M. Stiggelbout, E. Vogel-Voogt, E.M. Noordijk, and T.P.M. Vliet Vlieland. Individual quality of life: adaptive conjoint analysis as an alternative for direct weighting? *Quality of Life Research*, 17(4):641–649, May 2008.

[140] Jr. Ware, J.E., M.K. Snyder, W.R. Wright, and A.R. Davies. Defining and measuring patient satisfaction with medical care. *Evaluation and Program Planning*, 6(3-4):247–263, 1983.

[141] M.F. Scheier and C.S. Carver. Optimism, coping, and health - assessment and implications of generalized outcome expectancies. *Health Psychology*, 4(3):219–247, 1985.

[142] L.I. Pearlin and C. Schooler. The structure of coping. *Journal of Health Social Behavior*, 19(1):2–21, March 1978.

[143] J.E. Ware and C.D. Sherbourne. The mos 36-item short-form health survey (sf-36) .1. conceptual-framework and item selection. *Medical Care*, 30(6):473–483, June 1992.

[144] A.M. Stiggelbout, G.M. Kiebert, J. Kievit, J.W. Leer, J.D. Habbema, and J.C. de Haes. The "utility" of the time trade-off method in cancer patients: feasibility and proportional trade-off. *Journal of Clinical Epidemiology*, 48(10):1207–1214, October 1995.

[145] F.E. van Nooten, X. Koolman, and W.B. Brouwer. The influence of subjective life expectancy on health state valuations using a 10 year tto. *Health Economics*, 18(5):549–558, May 2009.

[146] J.M. Morea, R. Friend, and R.M. Bennett. Conceptualizing and measuring illness self-concept: A comparison with self-esteem and optimism in predicting fibromyalgia adjustment. *Research in Nursing & Health*, 31(6):563–575, December 2008.

[147] M.F. Scheier, C.S. Carver, and M.W. Bridges. Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem) - a reevaluation of the life orientation test. *Journal of Personality and Social Psychology*, 67(6):1063–1078, December 1994.

[148] R.M. Baron and D.A. Kenny. The moderator mediator variable distinction in social psychological-research - conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6):1173–1182, December 1986.

[149] M.E. Sobel. Asymptotic confidence intervals for indirect effects in structural equitions models. In S. Leinhart, editor, *Sociological methodology*, pages 290–312. Jossey-Bass, San Francisco, 1982.

[150] R. Arnold, A.V. Ranchor, R. Sanderman, G.I.J.M. Kempen, J. Ormel, and T.P.B.M. Suurmeijer. The relative contribution of domains of quality

of life to overall quality of life for different chronic diseases. *Quality of Life Research*, 13(5):883–896, June 2004.

[151] D.A. Revicki and R.M. Kaplan. Relationship between psychometric and utility-based approaches to the measurement of health-related quality-of-life. *Quality of Life Research*, 2(6):477–487, December 1993.

[152] A. Campbell. Subjective measures of well-being. *American Psychologist*, 31(2):117–124, 1976.

[153] J.C. de Haes and K. Welvaart. Quality of life after breast cancer surgery. *Journal of Surgical Oncology*, 28(2):123–125, February 1985.

[154] B.G. Tabachnick and L.S. Fidell. *Using multivariate statistics*, volume Fourth Edition. Allyn & Bacon, Needham Heights, MA, 2001.

[155] L.M. McCracken. Learning to live with the pain: acceptance of pain predicts adjustment in persons with chronic pain. *Pain*, 74(1):21–27, January 1998.

[156] P.A. Ubel, R.A. Hirth, M.E. Chernew, and A.M. Fendrick. What is the price of life and why doesn't it increase at the rate of inflation? *Archives of Internal Medicine*, 163(14):1637–1641, July 2003.

[157] M.W.M. Post, A.J. van Dijk, F.W.A. van Asbeck, and A.J.P. Schrijvers. Life satisfaction of persons with spinal cord injury compared to a population group. *Scandinavian Journal of Rehabilitation Medicine*, 30(1):23–30, March 1998.

[158] M.W.M. Post, L.P. de Witte, F.W. van Asbeck, A.J. van Dijk, and A.J. Schrijvers. Predictors of health status and life satisfaction in spinal cord injury. *Archives of Physical Medicine and Rehabilitation*, 79(4):395–401, April 1998.

[159] S. Folkman and J.T. Moskowitz. Positive affect and the other side of coping. *American Psychologist*, 55(6):647–654, June 2000.

[160] R.G. Tedeschi and L.G. Calhoun. The posttraumatic growth inventory: Measuring the positive legacy of trauma. *Journal of Traumatic Stress*, 9(3):455–471, July 1996.

[161] S.J.T. Jansen, J. Kievit, M.A. Nooij, and A.M. Stiggelbout. Stability of patients' preferences for chemotherapy: The impact of experience. *Medical Decision Making*, 21(4):295–306, July 2001.

[162] T.D. Wilson and D.T. Gilbert. Affective forecasting - knowing what to want. *Current Directions in Psychological Science*, 14(3):131–134, June 2005.

[163] C.S. Saigal, J. Gornbein, K. Reid, and M.S. Litwin. Stability of time trade-off utilities for health states associated with the treatment of prostate cancer. *Quality of Life Research*, 11(5):405–414, August 2002.

[164] C.F. van Koppenhagen, M.W. Post, L.H. van der Woude, S. de Groot, L.P. de Witte, F.W. van Asbeck, Heuvel W. van den, and E. Lindeman. Recovery of life satisfaction in persons with spinal cord injury during inpatient rehabilitation. *American Journal of Physical Medicine & Rehabilitation*, 88(11):887–895, November 2009.

[165] C. Collin, S. Davis, V. Horne, and D.T. Wade. Reliability of the barthel adl index. *International Journal of Rehabilitation Research*, 10(3):356–357, 1987.

[166] N.M. Crewe and J.S. Krause. An eleven-year follow-up of adjustment to spinal cord injury. *Rehabilitation Psychology*, 35(4):205–210, 2006.

[167] M.W.M. Post, F.W. van Asbeck, A.J. van Dijk, and A.J. Schrijvers. Dutch interview version of the barthel index evaluated in patients with spinal cord injuries. *Nederlands Tijdschrift Voor Geneeskunde*, 139(27):1376–1380, July 1995.

[168] K.T. Lucke, H. Coccia, J.S. Goode, and J.F. Lucke. Quality of life in spinal cord injured individuals and their caregivers during the initial 6 months following rehabilitation. *Quality of Life Research*, 13(1):97–110, February 2004.

[169] W.B. Mortenson, L. Noreau, and W.C. Miller. The relationship between and predictors of quality of life after spinal cord injury at 3 and 15 months after discharge. *Spinal Cord.*, 48(1):73–79, January 2010.

[170] J.S. Krause. Adjustment after spinal cord injury: A 9-year longitudinal study. *Archives of Physical Medicine and Rehabilitation*, 78(6):651–657, June 1997.

[171] A. Schwartz, G. Hazen, A. Leifer, and P. Heckerling. Life goals and health decisions: What will people live (or die) for? *Medical Decision Making*, 28(2):209–219, March 2008.

[172] J.C. de Haes, J.H. de Ruiter, R. Tempelaar, and B.J. Pennink. The distinction between affect and cognition in the quality of life of cancer

patients–sensitivity and stability. *Quality of Life Research*, 1(5):315–322, October 1992.

[173] A. Robinson, P. Dolan, and A. Williams. Valuing health status using vas and tto: What lies behind the numbers? *Social Science & Medicine*, 45(8):1289–1297, October 1997.

[174] P.F.M. Stalmeier, L.M. Lamers, J.J.V. Busschbach, and P.F.M. Krabbe. On the assessment of preferences for health and duration - maximal endurable time and better than dead preferences. *Medical Care*, 45(9):835–841, September 2007.

[175] D.T. Gilbert, M.D. Lieberman, C.K. Morewedge, and T.D. Wilson. The peculiar longevity of things not so bad. *Psychological Science*, 15(1):14–19, January 2004.

[176] T.D. Wilson and D.T. Gilbert. How happy was i, anyway? a tetrospective impact bias. *Social Cognition*, 21(6):421–446, 2003.

[177] Ubel P. A. Jankovic A. Smith D.M. Langa K. M. and Fagerlin A. What is perfect health to an 85-year-old?: Evidence for scale recalibration in subjective health ratings. *Medical Care*, 43(10):1054–1057, 2005.

[178] G. Loewenstein and D Schkade. Wouldn't it be nice? predicting future feelings. In D. Kahneman, E. Diener, and N. Schwarz, editors, *Well-Being The foundations of Hedonic Psychology*, number 5, pages 85–105. Russell Sage Foundation, New York, 1999.

[179] D.T. Gilbert, M.J. Gill, and T.D. Wilson. The future is now: Temporal correction in affective forecasting. *Organizational Behavior and Human Decision Processes*, 88(1):430–444, May 2002.

[180] K.J. Smith, J. Tsevat, R.B. Ness, H.C. Wiesenfeld, and M.S. Roberts. Quality of life utilities for pelvic inflammatory disease health states. *Sexually Transmitted Diseases*, 35(3):307–311, March 2008.

[181] D.A. Schkade and D. Kahneman. Does living in california make people happy? a focusing illusion in judgments of life satisfaction. *Psychological Science*, 9(5):340–346, September 1998.

[182] G. Loewenstein. Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes*, 65(3):272–292, March 1996.

[183] E. Diener, R.A. Emmons, R.J. Larsen, and S. Griffin. The satisfaction with life scale. *Journal of Personality Assessment*, 49(1):71–75, 1985.

[184] Smith D.M. Loewenstein G. Jankovich A. and Ubel P.A. The dark side of hope: Lack of adaptation to a temporary versus permanent disability. Health Psychology (In Press), 2010.

[185] Tversky A. and Kahneman D. Advances in prospect-theory - cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323, October 1992.

[186] R.E. Lucas. Long-term disability is associated with lasting changes in subjective well-being: Evidence from two nationally representative longitudinal studies. *Journal of Personality and Social Psychology*, 92(4):717–730, April 2007.

[187] D.M. Smith, G. Loewenstein, P. Rozin, R.L. Sherriff, and P.A. Ubel. Sensitivity to disgust, stigma, and adjustment to life with a colostomy. *Journal of Research in Personality*, 41(4):787–803, August 2007.

[188] C.E. Schwartz and M.A. Sprangers. Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. *Social Science & Medicine*, 48(11):1531–1548, June 1999.

[189] F. X. Gibbons. Social comparison as a mediator of response shift. social science and medicine. *Social Science and Medicine*, 48(11):1517–1530, 1999.

[190] Daltroy L. Larson M. Eaton H. Phillips C. and Liang M. Discrepancies between self-reported and observed physical function in the elderly: The influence of response shift and other factors. *Social Science and Medicine*, 48:1549–1561, 1999.

[191] Schwartz C. and Sendor R. Helping others helps oneself: Response shift effects in peer support. *Social Science and Medicine*, 48:1563–1575, 1999.

[192] I. B. Wilson. Clinical understanding and clinical implications of response shift. *Social Science and Medicine*, 48(11):1577–1588, 1999.

[193] C.E. Schwartz, R. Bode, N. Repucci, J. Becker, M.A.G. Sprangers, and P.M. Fayers. The clinical significance of adaptation to changing health: A meta-analysis of response shift. *Quality of Life Research*, 15(9):1533–1550, November 2006.

[194] G.L. Albrecht and P.J. Devlieger. The disability paradox: high quality of life against all odds. *Social Science & Medicine*, 48(8):977–988, April 1999.

[195] P. Ekman. *Emotions Revealed.* New York: Times Books., 2003.

[196] J. C. Brunstein. Personal goals and subjective well-being: A longitudinal study. *Journal of Personality and Social Psychology*, 65(5):1061–1070, 1993.

[197] Lepore S. J. and Eton D. T. *Response shifts in prostate cancer patients: An evaluation of suppressor and buffer modes.* Washington, D.C.: American Psychological Association., 2000.

[198] Kahneman D. Wakker P. P. and Sarin R. Back to bentham? explorations of experienced utility. *The Quarterly Journal of Economics*, 112(2):375–405, 1997.

[199] Loewenstein G. and Ubel P.A. Hedonic adaptation and the role of decision and experience utility in public policy. *Journal of Public Economics*, 92(8-9):1795–1810, 2008.

[200] Brossart D. Clay D. and Willson V. Methodological and statistical considerations for threats to internal validity in pediatric outcome data: Response shift in self-report outcomes. *Journal of Pediatric Psychology,*, 27(1):97–107, 2002.

[201] Schwartz C. Sprangers M. Carey A. and Reed G. Exploring response shift in longitudinal data. *Psychology & Health*, 19(1):51–59, 2004.

[202] D.M. Smith, G. Loewenstein, C. Jepson, A. Jankovich, H.I. Feldman, and P.A. Ubel. Mispredicting and misremembering: Patients with renal failure overestimate improvements in quality of life following a kidney transplant. *Health Psychology*, 27(5):653–658, 2008.

[203] D. Kahneman. Determinants of health economic decisions in actual practice: The role of behavioral economics. *Value in Health*, 9(2):65–67, 2006.

[204] F.J. Oort. Using structural equation modeling to detect response shifts an true change. *Quality of Life Research*, 14(3):587–98, 2005.

[205] P. Menzel, P. Dolan, J. Richardson, and J.A. Olsen. The role of adaptation to disability and disease in health state valuation: a preliminary normative analysis. *Social Science & Medicine*, 55(12):2149–2158, December 2002.

[206] J. Halpern and R.M. Arnold. Affective forecasting: An unrecognized challenge in making serious health decisions. *Journal of General Internal Medicine*, 23(10):1708–1712, October 2008.

Publications

Peeters, Y., Vliet Vlieland T.P.M, and Stiggelbout, A.M.; *Focusing illusion, adaptation and EQ-5D health state descriptions: The difference between patients and public.* Health Expectations 2011 Mar 3, Epub ahead of print.

Peeters, Y., Ranchor A.V., Vliet Vlieland T.P.M, and Stiggelbout, A.M.;*Effect of adaptive abilities on utilities, direct or mediated by mental health?* Health Quality of Life Outcomes 2010, 12 (8):130

Peeters, Y., and Stiggelbout, A.M.; *Health state valuations of patients and general public analytically compared; A meta-analytical comparison of patient and population health state utilities.* Value in Health 2010, 13(2): 306-309

Peeters, Y., and Stiggelbout, A.M.; *Valuing health: Does enriching a scenario lead to higher utilities?* Medical Decision Making 2009, 29 (3): 334-342

Peeters, Y., Boersma, S.N., and Koopman, H.M.; *Predictors of quality of life: A quantitative investigation of the stress-coping model in children with asthma.* Health and Quality of Life Outcomes 2008, 6: 24

Ubel P.A., Peeters Y., Smith, D.M.; *Abandoning the Language of "Response Shift": A Plea for Conceptual Clarity in Distinguishing Scale Recalibration from True Changes in Quality of Life.* Quality of Life Research 2010, 19 (4):465-471.

Vermunt, R., Peeters, Y., Berggren, K.; *How fair treatment affects saliva cortisol release in stressed low and high type-A behavior individuals.* Scandinavian Journal of Psychology 2007; 48: 547-555.

Peeters, Y., Putter, H., Snoek, G.J., Sluis, T.A.R., Smit, C.A.J., Post, M.W.M., Stiggelbout, A.M.; *The influence of time and adaptation on health state valuations in patients with Spinal Cord Injury.* Under revision (Medical Decision Making)

Peeters, Y., Smith, D.M., Loewenstein, G., and Ubel, P.A.; *Newly disabled patients expect to adapt to their condition.* Conditionally Accepted (Journal of Happiness Studies)

Stiggelbout A.M., Peeters Y., de Jong Z.; *Health state utilities from patients, partners and the public: in search of an explanation for the gap between the imagined and the real.* Submitted.

Curriculum Vitae

Yvette Peeters werd op 12 augustus 1981 geboren in Venlo, als kind van Hans Peeters en Riki Jacobs en zusje van Vivian. In 1999 behaalde ze haar VWO diploma aan het Bouwens van der Boijecollege in Panningen en begon zij aan haar studie psychologie aan de Universiteit Leiden. Als onderdeel van haar studie volgde ze een onderzoeksstage aan de Högskolan i Skövde (Zweden) en een klinische stage aan het Willem-Alexander Kinder- en Jeugdcentrum van het Leids Universitair Medisch Centrum. Eind 2005 behaalde ze haar doctoraal examen in zowel Methoden en Technieken als Klinische- en Gezondheidspsychologie.

In 2006 begon zij haar promotieonderzoek in dienst van de Universiteit Leiden, bij de afdeling Medische Besliskunde van het Leids Universitair Medisch Centrum waarvan de resultaten staan beschreven in dit proefschrift. In het kader van dit onderzoek heeft ze in 2009 een bezoek gebracht aan het Center for Bioethics and Social Sciences in Medicine (CBSSM) van de University of Michigan (U.S.), waar ze heeft samengewerkt met Prof. dr. Peter A. Ubel en dr. Dylan M. Smith. Naast haar onderzoek heeft ze in 2007 in het bestuur gezeten van het Leids Promovendi Overleg en momenteel is ze associate editor van het Journal of European Psychology Students (JEPS) een tijdschrift gericht op studies uitgevoerd door studenten. Samen met haar partner Jan Edelaar kreeg ze in 2010 een zoon, Kyan.

Na haar promotie onderzoek zal ze in juni 2011 starten op het project “Web-based time trade-off incorporating interviewer help: efficiency with validity” op de afdeling Medische Besliskunde van het Leids Universitair Medisch Centrum. Hiervoor heeft ze samen met haar collega’s een subsidie gekregen van ZonMW.

Je kunt niet anders zijn dan wie je bent

- Stef Bos, Mijn hart gevolgd -

Dankwoord

Prof. dr. A. M. Stiggelbout, Anne; dat je als promotor mij met je kennis en wijsheid hebt bijgestaan hoort erbij, maar je snelle reacties, betrokkenheid, en de vaak praktische tips horen minder vanzelfsprekend bij de functie, ik wil je hier oprecht voor bedanken. Omdat je naast promotor ook mijn dagelijkse begeleider bent geworden zou ik je daar ook nog graag voor willen bedanken, maar dat zou tegen de LUMC regels voor het dankwoord zijn.

Nanny ook jou wil ik graag bedanken voor je betrokkenheid maar ook voor je ondersteuning bij het afnemen en verwerken van de interviews, het voelde als teamwork.

Mijn (ex) kamergenoten en mijn collega's van de afdeling Medische Besliskunde bedankt voor jullie steun, kritische vragen en waardevolle opmerkingen.

Next I want to thank Prof. dr. Peter A. Ubel and Dr. Dylan M. Smith you taught me a lot in a short period of time and your enthusiasm was inspiring. Yet, my stay in Michigan could not have been so nice without the support of all other members of the Center for Bioethics and Social Sciences in Medicine (CBSSM).

Tot slot wil ik me richten op alle respondenten die vrijwillig hebben meegewerkt en natuurlijk ook de medewerkers van de verschillende afdelingen (artsen, psychologen, verpleegkundigen en de medewerkers van de planning) die betrokken zijn geweest bij het project; de polikliniek en Sole Mio van de afdeling Reumatologie van het LUMC, Het Roessingh centrum voor revalidatie, Rijndam revalidatiecentrum, Reade revalidatie, Revalidatiecentrum De Hoogstraat, Sophia Revalidatie en Heliomare.