

Review

Open Access

Toward a theoretical model of quality-of-life appraisal: Implications of findings from studies of response shift

Bruce D Rapkin*¹ and Carolyn E Schwartz^{2,3,4,5}

Address: ¹Department of Psychiatry and the Behavioral Sciences, Memorial Sloan Kettering Cancer Center, New York, NY, USA, ²QualityMetric Incorporated, Waltham, MA, USA, ³Health Assessment Lab, Waltham, MA, USA, ⁴Division of Preventive and Behavioral Medicine, Department of Medicine, University of Massachusetts Medical School, Worcester, MA, USA and ⁵DeltaQuest Foundation, Concord, MA, USA

Email: Bruce D Rapkin* - rapkinb@mskcc.org; Carolyn E Schwartz - caroln.schwartz@deltaquest.org

* Corresponding author

Published: 15 March 2004

Received: 23 January 2004

Health and Quality of Life Outcomes 2004, **2**:14

Accepted: 15 March 2004

This article is available from: <http://www.hqlo.com/content/2/1/14>

© 2004 Rapkin and Schwartz; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Mounting evidence for response shifts in quality of life (QOL) appraisal indicates the need to include direct measurement of the appraisal process itself as a necessary part of QOL assessment. We propose that directly assessing QOL appraisal processes will not only improve our ability to interpret QOL scores in the traditional sense, but will also yield a deeper understanding of the appraisal process in the attribution of and divergence in meaning. The published evidence for response shift is reviewed, and an assessment paradigm is proposed that includes the explicit measurement of QOL appraisal process parameters: 1) induction of a frame of reference; 2) recall and sampling of salient experiences; 3) standards of comparison used to appraise experiences; and 4) subjective algorithm used to prioritize and combine appraisals to arrive at a QOL rating. A QOL Appraisal Profile, which measures key appraisal processes, is introduced as an adjunct to existing QOL scales. The proposed theoretical model, building on the Sprangers and Schwartz (1999) model and highlighting appraisal processes, provides a fully testable theoretical treatment of QOL and change in QOL, suggesting hypothesized causal relationships and explanatory pathways for both cross-sectional and longitudinal QOL research.

Quality of life (QOL) assessment involves a class of measurement fundamental to many aspects of health care planning and outcomes research. It is relevant for assessing symptoms, side effects of treatment, disease progression, satisfaction with care, quality of support services, unmet needs, and appraisal of health and health care options. Patient self-report is the most desirable, and often the only way to obtain this critical information. Thus, accurate and meaningful measures of the various dimensions of QOL are vitally important. Here, we 1) review evidence from the response shift literature regard-

ing different cognitive processes that influence QOL appraisal; 2) build upon the Sprangers and Schwartz [1] model, to develop a theoretically grounded measurement model that addresses the phenomenology of QOL appraisal and suggest methods of assessing this phenomenology; and 3) discuss how appraisal assessment can be incorporated in statistical and clinical judgment models of QOL, to provide a coherent and empirically-testable definition of response shift.

Reconciling inconsistent findings in QOL research

The importance of QOL makes it critical to improve and refine measures to understand patients' experiences of health, illness, and treatment. Unfortunately, pervasive paradoxical and counterintuitive findings raise questions about what QOL measures actually assess and how scores should be interpreted: people with severe chronic illnesses report QOL equal or superior to less severely ill or healthy people [2-8], and consistent disparities arise between clinical measures of health and patients' own evaluations [9-11]. Indeed, several studies show that health care providers and significant others tend to underestimate patients' QOL compared with patients' evaluations [12-15]. In short, QOL measures do not consistently distinguish known groups, are often only weakly related to objective criteria, and show little convergence across measurement perspectives.

These inconsistent findings support the notion of underlying differences in the phenomenology of QOL appraisal between people and are a function of coping with chronic or life-threatening illness and other sources of stress [16]. Rather than reflecting lack of validity, measurement bias, denial, or willful distortion, these phenomenological factors may reflect individual differences and intra-individual changes in internal standards, values, and meaning of QOL [1,17]. Differences in QOL appraisal are part of human adaptation and inherent in all QOL measurement.

QOL can mean different things to different people at different times [18,19]. Indeed, many QOL measures have been specifically crafted to be as generic as possible to circumvent such differences (as discussed by Stewart and Napoles-Springer [20]). In contrast, methods to detect response shift phenomena have assessed individual differences and intra-individual changes in the meaning of QOL. As we argue below, QOL research has much to gain by using methods that encompass these phenomena. Theoretical and empirical work on response shifts in QOL support the notion that differences in appraisal enter into all self-ratings of QOL and shed light on the nature of appraisal processes, demonstrating ways that appraisal might be directly assessed.

Background on response shift

The concept of response shift is grounded in research on educational training interventions [21-25] and organizational change [26]. The original definition of response shift specified recalibration of internal standards of measurement [21-24] and reconceptualization of the meaning of items [26]. Sprangers and Schwartz [1] added reprioritization of values as a third aspect of response shift phenomena and proposed a theoretical model (Figure 1), revised and updated here (Figures 2, 3), to clarify and pre-

dict changes over time in perceived QOL as a result of the interaction of catalysts, antecedents, mechanisms, and response shifts. **Catalysts** (a) refer to health states or changes in health states, as well as other health-related events, treatment interventions, the vicarious experience of such events, and other events hypothesized to have an impact upon QOL (life events). **Antecedents** (b) include characteristics of the person, culture, and environment hypothesized to influence the likelihood and type of catalysts and mechanisms of appraisal. **Mechanisms** (c) encompass behavioral, cognitive, or affective processes to accommodate changes in catalysts (initiating social comparisons, reordering goals). **Response shift** (d) includes changes in the meaning of one's self-evaluation of QOL resulting from changes in internal standards, values, or conceptualization. This model posits a dynamic feedback loop to explain how quality of life scores can be stabilized despite changes in health status.

Although useful for hypothesizing relationships among key constructs relevant to QOL assessment, the model presents some problems of logical circularity as operationalizations of mechanisms or outcomes may be synonymous with operationalizations of response shift [27]. The model required expansion to distinguish these components from response shift and to differentiate response shift phenomena as initial responses to catalysts from those that reverberate or continue the process (feedback loop). Thus, in this paper, we attempt to resolve these problems by introducing new models incorporating constructs based on Schwartz and Sprangers' work and direct measures of QOL appraisal processes to account for unexplained change in QOL ratings.

Recent empirical research documents the presence and importance of response shifts in both treatment outcome research and naturalistic longitudinal observations of QOL. Several studies suggest that patients make significant response shifts during treatment. Sprangers and colleagues [28] found changes in internal standards for fatigue in two subgroups of cancer patients undergoing radiotherapy: 1) patients experiencing diminishing levels of fatigue, and 2) patients facing early stages of adaptation to increased levels of fatigue. Jansen and colleagues [29] confirmed these changes in internal standards of fatigue and documented changes over time in patients' importance weights for one toxicity (skin reactions) associated with treatment. Multiple sclerosis patients receiving beta-interferon-1b demonstrated changes in the importance of various QOL dimensions over the course of treatment [30]. Adang and colleagues reported that pancreas-kidney transplant recipients retrospectively rate their pretransplant QOL lower when transplantation is successful [31]. Similarly, Ahmed and colleagues reported that measures of improvement of health status differ in prospective versus

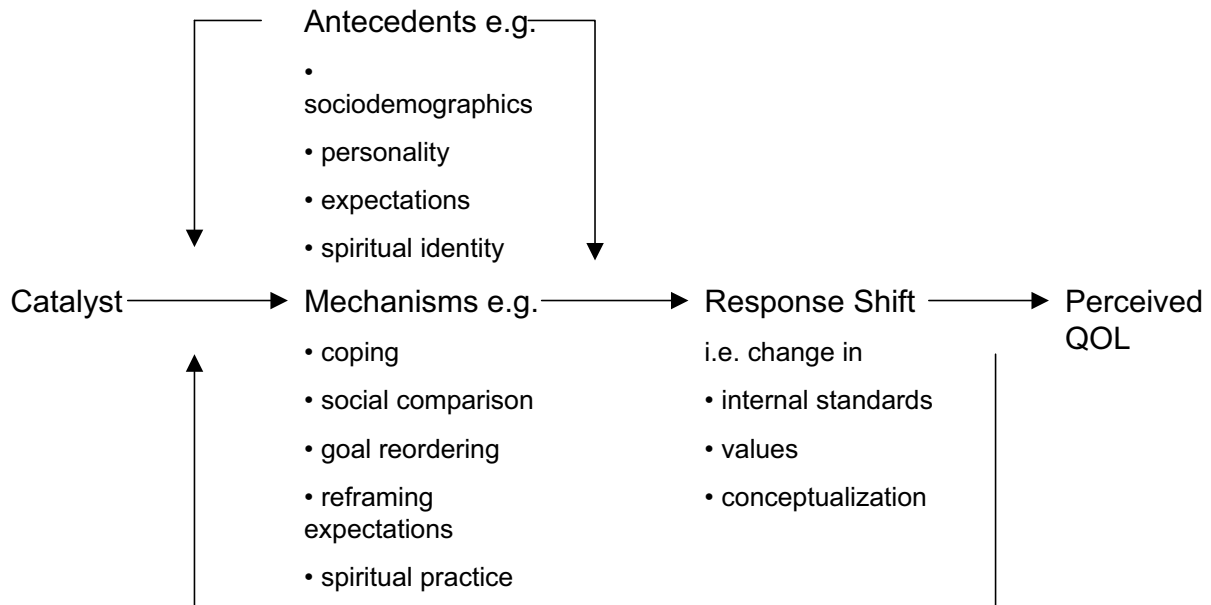


Figure 1
Sprangers and Schwartz (1999) theoretical model of response shift and quality of life

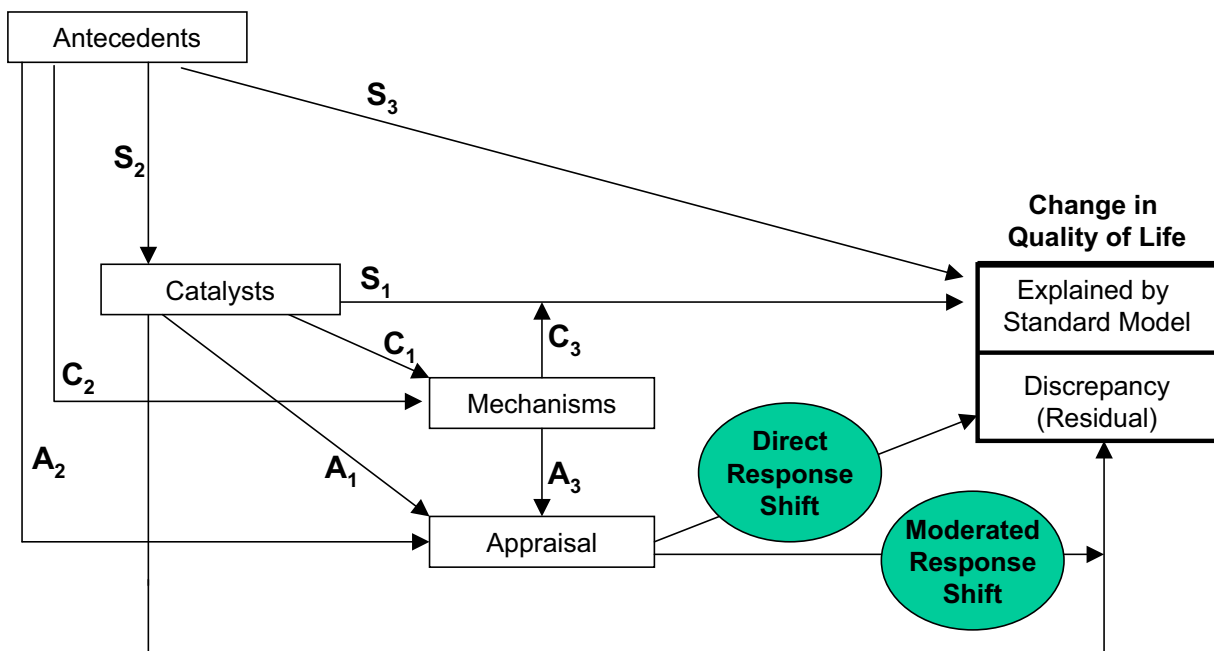


Figure 2
Partitioning response shift effects in the Sprangers and Schwartz (1999) model using a linear regression paradigm: Accounting for changes in Standard influences (S), Coping processes (C), and Appraisal (A) variables

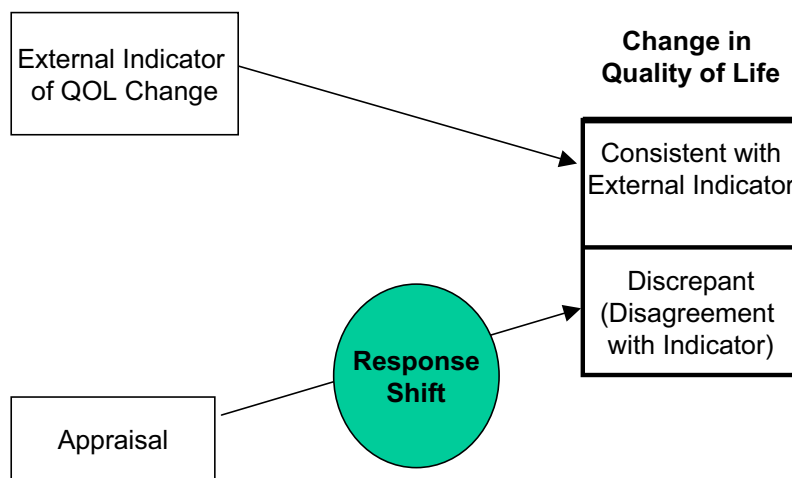


Figure 3
 Response shift in a clinical judgment paradigm based on the Sprangers and Schwartz (1999) model: Accounting for discrepancies between changes in observed clinical status (o) and appraised quality of life

retrospective assessment [32]. Hagedoorn and colleagues found that cancer patients who felt they were better off than others appeared to sustain their QOL under worsening physical condition [33]. Schwartz and colleagues [34] found that an apparently deleterious QOL effect of a psychosocial intervention was largely a function of response shifts in internal standards and conceptualization of QOL. Rees and colleagues found that recalibration response shifts are more likely in the first few months after a threatening event, that patients with more severe symptoms engage in recalibration response shifts longer than patients with milder symptoms [35], and that considering recalibration response shift produced a 10% increase in estimated QOL in prostate cancer patients [35]. Thus, intra-individual comparisons over time may not be comparable or sensitive to change unless they explicitly measure response shifts.

Response shift is also important for medical decision-making. Lenert, Treadwell, and Schwartz [36], using preference-assessment methods common in cost-effectiveness analysis to investigate interactions between preferences and health status, found that patients in poor health valued intermediate health states almost as much as near-normal states. Conversely, patients in good health valued intermediate states nearly as little as poor health states.

Patients in poor physical and mental health tended to recalibrate their standards for comparing health states in a manner that downplayed current personal problems, and small gains were more valuable to disabled than to healthy persons. These findings are consistent with those of Cella and colleagues that, among cancer patients, relatively small gains in function and QOL have significant value, whereas comparable declines in status may be less meaningful [37]. Ultimately, cost-effectiveness of medical treatments may depend on the health status of persons rating preferences.

In addition to influencing our approach to outcomes measurement, response shift research also suggests reconsideration of standard QOL designs. Lepore and Eton [38] compared the fit of two theoretical models – suppressor and buffering models – for explaining the lack of association between physical health problems and reported QOL in men with prostate cancer. They used the then-test to operationalize recalibration response shift and a measure of primary life goal changes modeled after one used by Rapkin and Fischer [39] to assess reprioritization response shifts. The then-test, or retrospective-pre-test design, asks respondents to fill out the self-report measure in reference to how they perceive themselves to have been at the pre-test [17]. Thus, the then-test asks for a renewed judgment

about their pre-test level of functioning [17]. Suppressor analyses tested whether response shifts explained a null relation between negative health status changes and QOL. Buffering models examined whether the relation between changes in sexual/urinary problems and QOL was weaker among men who did or did not make response shifts. These linear regression analyses produced some evidence consistent with the buffering model. An interactive effect suggested that response shifts moderated the association between increases in urinary problems and changes in QOL. Specifically, response shifts appeared to buffer men from negative effects of declines in urinary function and QOL. They found no evidence for the suppressor model of response shifts. Thus, the response shift construct may help account for individual differences in QOL among prostate cancer patients who experience post-treatment complications.

Other recent research examines how an explicit consideration of response shifts might elucidate an understanding of QOL in various patient and caregiver populations. Richards and Folkman [40] examined response shifts among bereaved caregivers of men with AIDS from a coping perspective. Using qualitative data, they illustrated the processes through which response shift is achieved and maintained through meaning-based coping: marking loss, evolving new expectations with their own positive meanings, finding meaning in the ordinary, and creating global (deeply held core values) and situational (ordinary events of daily life) meaning for their caregiving experience.

Evidence for response shift has also been demonstrated in studies of the appraisal of health status. In a secondary analysis of cross-sectional data, Daltroy and colleagues [41] compared a measure of function based on observed performance (the Physical Capacity Evaluation, Daltroy et al., [42]) with a self-reported measure of disability (the Health Assessment Questionnaire, [43]) to test several response shift hypotheses, using stepwise linear regression and Fisher's Z transformation to compare correlation coefficients by type of recent loss (illness, fall, pain or stiffness, or perceived decline in function). Their data were consistent with response shift predictions. Specifically, people recalibrate their self-assessments of functional ability based on recent health problems. Additionally, physical performance testing provides salient information for subjects who have not experienced recent decline. Providing objective performance indicators can improve agreement between observed function and self-reported ability, perhaps by counteracting a response shift. They propose performance measures as a universal standard to correct for differential self-report of various subgroups. Further, patients might be reassured by a performance test that counteracts a response shift whereby they overesti-

mate their disability, thereby possibly reducing health care expenditures by anxious patients seeking reassurance.

Finally, Rapkin [44] examined how the impact of life events on QOL were subject to reconceptualization and reprioritization response shifts associated with changes in personal goals in a longitudinal study of people with AIDS. Using idiographic assessment, people were asked to identify changes in personal goals most strongly associated with high life satisfaction. Individuals were free to mention any goals that mattered to them. Response shifts in self-appraised QOL were defined as discrepancies from some expected value that could be explained by direct measures of change in priorities associated with QOL (that is, change in personal goals). Expected QOL values in this study were operationalized using a regression model, taking into account initial QOL and changes in health status, stressful events, and coping resources. Rapkin's analysis attempted to explain discrepancies between observed and expected change in QOL by assessing whether people who changed their personal goals reacted to illness and events differently from those whose goals did not change. In statistical terms, QOL response shifts were operationalized as statistical interaction effects, with changes in goals amplifying or attenuating anticipated (main) effects of disease progression, life events, or treatments on QOL. Rapkin's findings suggested four distinct reprioritization response-shifts associated with changes in personal goals and concerns. People's reaction to life events and disease progression depended upon whether and how their goals changed. Perhaps more fundamentally, these findings provide direct evidence that people's goals and concerns continue to evolve during serious illness, perhaps up to death.

In summary, a variety of assessment methods of response shift confirms that QOL assessment involves a subjective process of appraisal, that individual differences in the appraisal process can affect observed QOL scores, and that individuals can change how they appraise QOL over time. Work on response shift phenomena is still at an early stage, but there is ample evidence to encourage investigators to include explicit methods for evaluating and integrating response shift phenomena in the next generation of QOL studies. Response shift findings point to the need for a broader QOL assessment paradigm that encompasses self-appraisal and meaning.

Definition of QOL based on appraisal

How then should one think about the appraisal of QOL? An individual's answer to any self-evaluative question depends upon this process. Individual differences or longitudinal changes in appraisal will affect how people respond to QOL items. Similarly, factors that correlate with QOL, including differences in personal circum-

stances, stressful events, disease progression, and interventions, also depend upon the criteria individuals use to evaluate QOL. Appraisal is a hidden facet in all measurement of QOL, and all studies involving self-reported QOL are influenced by appraisal.

In sum, any response to a QOL item can be understood as a function of an appraisal process. In other words, we view QOL scores as contingent upon several key variables related to appraisal. In order to describe QOL appraisal adequately, we posit at least four distinct cognitive processes suggested by research on QOL response shift and largely anticipated in Sprangers and Schwartz's original response shift model [1]. These four processes also relate to cognitive processes involved in formulating responses to surveys identified by Tourangeau, Rips and Rasinski [45]. As Jobe [46] points out, there have been a number of variations on the four-process model. We have adopted our operational definitions of appraisal processes to correspond to psychological aspects of coping and adjustment intrinsically related to QOL appraisal. This is an important distinction: Tourangeau and colleagues emphasize psychological processes that arise in the survey situation [45]. Alternatively, we believe that cognitive aspects of quality of life appraisal are not merely a measurement issue, and may themselves become the focus of clinical interventions to help patients understand and think about their QOL in more adaptive ways.

First, QOL assessment induces a **frame of reference**, experiences individuals deem relevant to their response. This frame of reference depends upon the meanings the individual attaches to questions [47], as well as demand characteristics of the testing situation. This aspect of appraisal relates to the process of comprehension described by Tourangeau and colleagues [45]. Individuals implicitly understand that QOL items refer to certain aspects of their life, although what aspects are only partly determined by the overt item content. Questions about global well-being, general health, social functioning, or mood can each invoke a range of different issues and concerns idiosyncratic to the individual.

Second, in order to respond to any item, individuals necessarily sample specific experiences within their frame of reference. We posit a subjective **sampling strategy** that is at least in part determined by the item *per se* and the broader context of the QOL measure and the assessment situation [48-51]. Tourangeau and colleagues [45] discuss this in terms of retrieval of autobiographical information.

Third, each sampled experience is judged against relevant, subjective **standards of comparison**. This represents a special case of the answer-estimation heuristics discussed by Tourangeau and colleagues [45]. There has been con-

siderable interest in how medical patients make comparisons to judge their health. Such comparisons may be based upon personal reference points [52], including prior functioning, lost capacities, and extreme experiences [53]. Observations of other patients, past encounters with illness, and communication from providers may also enter into appraisal [54,55]. Of course, individuals may select standards in a biased fashion, leading to criteria that are more or less demanding or strident, as Gruder [56] pointed out in the distinction between self-enhancing and self-evaluative comparisons.

Fourth, to arrive at a QOL score, individuals must apply some **combinatory algorithm** to summarize their evaluation of relevant experiences and formulate a response [45]. Individuals may combine their experiences in an additive and linear fashion, using subjective salience weights to increase or decrease the relative importance of different experiences [48,57,58]. However, the combinatory strategy may be more complicated. The significance of a particular experience may be determined only in contrast to other relevant experiences. Thus, individuals may place added emphasis on recent patterns or unusual events. Frequently repeated or continuous difficulties may be treated as a single experience and so receive less emphasis than if they were appraised as separate events.

It will be useful to depict these appraisal processes in more formal terms. As noted above, the prevailing assessment paradigm presumes that an observed QOL score at a specific time represents the sum of the latent true score plus random error, or $Q_t = q_t + e_t$. Although investigators rarely state this equation, this measurement model is the foundation for all research using QOL scales, including global well-being or life satisfaction and specific domains. It follows from our discussion of appraisal and response shift phenomena that there is much more to observed QOL scores than meets the eye. To understand better the nature of the latent variable q_t , the QOL true score at time t , we must "unpack" q_t to specify how different appraisal processes enter into QOL assessment. Given this formulation, a measurement model for q_t can be formally represented in terms of information about the appraisal processes that constrain and qualify QOL, as follows:

Equation 1 (Induction): $\{FR_t\} \subset \{K_{kt}\}_{k=1 \rightarrow K}$

The induction of a frame of reference for QOL: A person's frame of reference for responding to a particular QOL scale or item is represented by the symbol $\{FR_t\}$, depicted as a set comprising one or more subsets $\{K_i\}$. These subsets can be understood as categories of experiences or events that the individual considers relevant to that QOL scale at that time. $\{K_{kt}\}$ stands for the k^{th} -specific category at time t . In other words, Equation 1 states that the frame of reference may be understood as a set of categories of

experience. Any individual's QOL rating is necessarily shaped and constrained by this frame of reference. A person thinking about QOL may consider a single category or multiple categories of concerns, including such areas as activities of daily living, emotional well-being, personal growth, social roles, and interpersonal relationships. If the frame of reference changes, we might expect the QOL score – or at least QOL correlates – to change. For example, if someone decides that work is no longer a priority, correlation between work-related events and QOL change should decrease correspondingly. Thus, the true QOL score at any given time is contingent upon $\{FR_t\}$, the person's frame of reference.

Equation 2 (identification): $x_{ikt} \in \{K_{kt}\} | S_{kt}$

Identification and sampling of specific experiences or events within that frame of reference in making QOL appraisals: The second equation indicates that QOL appraisal calls upon the individual to sample experiences from the various categories $\{K_{kt}\}$ that make up his/her frame of reference (from Equation 1). x_{ikt} represents the i^{th} experience from the k^{th} -category at time t . These experiences are sampled according to a strategy S_{kt} , represented as a constraint in this model. In other words, individuals consider specific experiences sampled from categories within their frames of reference and determined or constrained by some way of thinking that leads them to pay attention to some things and not others. Again, QOL true scores are contingent upon this identification process. Even if the frame of reference remains constant, sampling different experiences may lead to different ratings of QOL. Although the specific experiences individuals consider may change over assessment occasions, the strategy they use to recall or "sample" experiences must be articulated. For example, paying attention to "recent instances of pain" could lead to different ratings than considering "times when pain interfered with my activities".

Equation 3 (evaluation): $[A_t] = [X_t] - [O_t] | R_t$

Evaluation of sampled events or experiences against some standard of comparison: The third equation includes all the experiences an individual considers at time t (all of the x_{ikt} from Equation 2), arrayed in a one-dimensional vector $[X_t]$. Each experience is compared with some optimal situation or desired outcome according to the individual's standards of comparison at time t . These standards are represented as a vector $[O_t]$, which is the same dimension as $[X_t]$. Just as experiences are sampled according to a specific strategy (indicated by S_{kt} in Equation 2), standards for desired outcomes are derived relative to specific reference groups or external criteria, indicated as R_t . The difference between $[X_t]$ and $[O_t]$ yields a vector of appraisals of each experience under consideration at that time, $[A_t]$. Equation 3 makes it explicit that appraisal of the experiences related to QOL depends on standards that may be

subject to change. For example, recent pain experiences might be compared to "the worst pain I ever had" or to "what my doctor told me to expect" or "how I wish things were". Clearly, QOL true scores are necessarily contingent on the standards an individual invokes.

Equation 4 (combination): $q_t = [W_t]' [A_t]$

Combination of evaluations into a summary appraisal of QOL: The fourth and final equation indicates that the QOL true score is the point product of the vector of appraisals $[A_t]$ premultiplied by a (transposed) vector of the same dimension $[W_t]'$. $[W_t]$ consists of the weights needed to combine appraisals across experiences. In other words, these weights represent a combinatory algorithm that dictates the relative impact or importance of specific experiences on QOL at time t . Equation 4 shows that any QOL rating is based on an amalgam of appraisals of different experiences that depends upon weights by their importance at a given time. For instance, a patient may give greater importance to recent instances when pain medications failed to provide relief, or on new sensations or locations.

In sum, this measurement model addresses the fact that QOL ratings are not intrinsically meaningful and can only be accurately understood through an underlying appraisal process. These four equations represent an attempt to identify and organize psychological processes involved in QOL appraisal to yield a definition of q_t as a contingent construct. The QOL true score is depicted as a direct function of weighted judgments of experiences in equation 4. This vector of experiences is identified according to equation 2 and evaluated according to equation 3. Available experiences are determined by the frame of reference in equation 1. This formulation allows us to specify this process in more precise terms. Thus, rather than speak of a true score in an absolute sense, QOL measures yield an estimation of $q_t | \{FR_t\}, S_{kt}, R_t, [W_t]$, or the true QOL score at a given point in time, contingent upon the individual's particular frame of reference, the ways that s/he identified and sampled relevant experiences, the reference groups and standards considered in evaluating those experiences, and the relative importance of each experience.

Measurement of appraisal constructs

As noted in our review of the response shift literature, many different approaches have been used to measure these appraisal constructs. This new measurement model conforms best with direct idiographic approaches that ask people to identify areas of concern, and then to sample, evaluate and prioritize salient experiences in those areas [59-61]. Such approaches literally ask individuals to "write" their own QOL items at each point of assessment. Although these techniques provide very rich data, they can be quite unwieldy to use and difficult to score. As an alter-

native, we have tried to develop more conventional self-report methods that directly assess parameters in the appraisal model. Such assessment is necessary to account for the effects of appraisal in QOL research, as we discuss below.

Appendix 1 (see Additional file: 1) provides the longitudinal version of the QOL Appraisal Profile (QOLAP) that we have developed to assess each of the appraisal parameters in this model. The QOLAP is designed to be used as an adjunct to standard QOL measures. Instructions presume that patients have just completed one or several such measures. The first item, based on questions used by Wernicke and colleagues [62], asks respondents to provide their perspective on quality of life. Items 2 through 7 are based on Rapkin and colleagues [44,61] assessment of personal goals. Item 8 asks people to highlight which of their personal goals (if any) were on their minds when they responded to QOL measures during the preceding portion of the interview. Together, these two sections provide a broad assay of an individual's frame of reference. Responses to these items are coded to identify the specific life domains and developmental themes of current concern to respondents. Occurrence of different codes can be tallied across different responses, to determine the range and variety of concerns mentioned. For example, an individual's goals may all pertain to health or family or mood, or they may pertain to multiple concerns.

Items 9a-9n are face valid questions, written to capture different implicit strategies people use to sample experiences. Items focus on the window of time patients may have considered, positive or negative aspects of experiences that may have made them stand out, clinically relevant features, as well as perceived demand characteristics of the interview situation. Similarly, Items 10a-10i provide a face valid assessment of possible comparison groups respondents may consider in rating their QOL. Following Suls and Miller's [63] classic work on social comparison theory, we tried to identify both historical and social reference groups that would provide for both self-critical and self-enhancing evaluations.

Item 11 provides a modified semantic differential to assess factors that respondents may use in weighing or prioritizing experience. Following Schwartz et al., [64] discussion of weights for individualized Quality-Adjusted Life Years (QALYs), we viewed the notion of "importance" as a multidimensional construct that may be related to several features of experience. For one individual, an unexpected event or experience may seem most important while others may be more concerned with typical or enduring problems. Importance may or may not be associated with other people's priorities. Positive events may receive greater, lesser or the same weight as negative

events. This set of items tap these different possibilities by asking individuals to reflect on the factors that mattered most in their responses to the preceding QOL items.

Items 12a-12b represent a standard use of the retrospective pre-test [22,65,66]. This follow-up version of the QOLAP presumes that item 12a alone was asked at baseline. As we shall discuss below, studies of response shift have focused on the discrepancy between the original answer to 12a (QOL rating obtained at the time of prior assessment) and 12b (the retrospective rating of QOL at the time of prior assessment, obtained at the current assessment). Item 12c is included to help gauge, recall effects. Finally, item 12d gives respondent an opportunity to reflect on discrepancies in their answers, to provide insight into how their criteria for appraising QOL have changed over time. This is similar to approaches used in cognitive interviewing [46,67]. Although we expect that responses to item 12d will be particularly related to changes in standards of comparison, it is possible that the reasons given for discrepant ratings may reflect changes in the meaning of the term "overall health" (frame of reference) or the salience of different experiences.

Finally, items 13 and 14 of this follow-up measure ask people to consider their original verbatim responses to items 1-7 from the prior time of measurement. As item 13 demonstrates, it is entirely possible that people may use slightly different language to express similar concerns from time to time. People may also inadvertently omit concerns that were mentioned previously. We want to rule out these possibilities before assessing change in goal content. Finally, Item 14 asks people to reconcile earlier and later statements concerning QOL, as a way of assessing how their definitions may have changed over time.

At the present time, we are gathering QOLAP follow-up data from a cohort of Medicaid HIV/AIDS patients. This is an ideal sample to evaluate this measure, including an ethnically diverse mix of asymptomatic and symptomatic patients, all of whom are lower socioeconomic status, and approximately 50% with a significant history of substance use. Interviews are being administered in both English and Spanish. To date, over 200 patients have completed the baseline portion of the interview (Items 1-11 and 12a), in an average of 15-20 minutes. We anticipate re-interviewing 70-80% of this sample at six-month follow-up. Follow-up interviews are necessary to address our primary concern, the impact of response shift on the measurement of QOL outcomes.

Appraisal and response shift in the regression paradigm

In linear regression and all related approaches (e.g., SEM, HLM, GEE), the goal is to account for variance in change

of QOL using a variety of different predictors. The model in Figure 2 includes several different families of hypothetical relationships that are frequently considered in QOL research, and shows how they are related to appraisal and response shift. It will be useful to consider each set of relationships in turn.

We refer to the first family of relationships as the "standard" QOL research model, whose primary hypothesis is that catalysts (e.g., changes in health, treatment, life events) are directly related to QOL (S_1). Negative catalysts are related to lower QOL and positive catalysts to higher QOL. The effects of antecedents (e.g., demographic factors, personality, cultural, and historical influences) on QOL are mediated through catalysts (S_2). For example, poverty may cause more negative life events leading to worse QOL. Antecedents may also be controlled as exogenous covariates (S_3).

The second family of hypotheses involves coping mechanisms. First, catalysts are hypothesized to encourage or disrupt coping mechanisms (C_1). There may also be hypothesized differences in coping associated with background variables (C_2). Mechanisms of problem-focused coping that reduce the impact of catalysts on QOL are included as moderators or buffering effects (C_3). Note that taking into account the direct, indirect, and moderator effects of catalysts, antecedents, and coping mechanisms on QOL effectively controls all of these variables, making it possible to isolate effects associated with appraisal in later steps of the model. For this reason, we have partitioned Change in QOL to distinguish variance associated with standard predictors such as overt health status and treatment from residual variance that remains after these familiar variables are controlled.

Our third family of hypothesized relationships concerns appraisal processes. The path from mechanisms to appraisal (A_3) indicates that coping mechanisms can lead to changes in the appraisal of QOL. However, catalysts (A_1) or antecedent variables (A_2) can also influence appraisal. Regardless of their cause, changes in appraisal may affect QOL ratings directly ("direct response shift" path) or by attenuating the impact of catalysts ("moderated response shift" path).

These different paths serve to demonstrate important distinctions and relationships among three broad constructs: coping, appraisal, and response shift. Emotion-focused coping represents cognitive behavior that individuals engage in intentionally, directed at changing the way that they understand QOL (or threats to QOL). Appraisal constructs represent the content of what an individual considers relevant to their QOL. For example, people may attempt to cope by reordering their goals. Appraisal

assessment would explicitly describe how those goals have changed. However, changes in appraisal need not depend on intentional efforts to cope. Rather, such changes may be due to other mechanisms including habituation, trauma, or socialization to patient status.

In the context of a regression paradigm, response shift may be operationally defined in terms of residual variance in the QOL change score that can be explained by changes in appraisal, after taking into account standard influences. These changes in appraisal may be due to coping or other processes. Different appraisal parameters map to the different types of response shift identified by Schwartz and Sprangers [1]:

||- changes in the frame of reference relate to *reconceptualization*;

||- changes in strategies for sampling experience within one's frame of reference deemed relevant to rating QOL as well as changes in the factors that determine the relative salience of different experiences relate to *reprioritization*;

||- changes in standards of comparison for evaluating one's experience relate to *recalibration*.

In the regression model, significant variance in QOL change associated with a given subset of appraisal measures would be taken as evidence for the corresponding type of response shift. In this sense, these response shifts may be considered "epiphenomena" that involve unexpected (e.g., unpredicted) or discrepant (e.g., residual) changes in QOL that can be explained by specific kinds of changes in the ways that individuals understand and appraise QOL.

Appraisal and response shift in clinical judgments and decision making

As noted above, measuring response shift involves accounting for changes in ratings of QOL that are discrepant from some expected value. In the regression model, the expected level of change in QOL is estimated statistically by adjusting observed QOL for catalysts and antecedents. Alternatively, investigators have also compared self-reported change in QOL to other external measures of QOL change that are independent of patient self-report – clinician judgment, performance tests, or family caregiver ratings. Even the popular retrospective-pre-test approach [22] described above asks the individual to re-rate past QOL from an "independent" perspective. Discrepancies between self-reported and external criteria for QOL have also been used to identify response shifts. Figure 3 describes how measures of appraisal can be used in this context.

In this figure, as in the regression model, we have again partitioned change in QOL outcome variables. However, rather than using a residual score, the partition here may be derived by taking a simple difference score between observed and self-reported QOL (scaled using the same metric). The important feature in this model is that self-reported change in QOL is directly compared with an external measure of change. QOL discrepancy here reflects the difference between two different measurement perspectives.

Both external and self-reported measures of QOL are subject to the effects of catalysts, antecedents and coping constructs in the Sprangers and Schwartz [1] model. For the sake of clarity, we omit these paths from Figure 3. Figure 3 demonstrates that it is possible to determine whether discrepancies between external measures and self-reported QOL are explained by changes in the ways that individuals appraise QOL. If individuals' ratings of change in QOL closely map to the external measure, discrepancies would be small and there would be little variance left to explain. However, if a significant portion of variance in discrepancies can in fact be attributed to changes in appraisal, this effect would be evidence of response shift.

It is important to emphasize that the model presented in Figure 3 can be used to describe response shifts at the level of a single case. For example, a clinical interview may identify a patient who is very frustrated by a slow process of rehabilitation. Her rating of self-reported functioning may be very negative compared to an external performance measure. Over time, this patient's overt performance may change very gradually. However, this patient's subsequent ratings of QOL may improve as she adjusts her expectations of progress and begins to take satisfaction from small gains. This kind of change represents a (recalibration) response shift. As this example demonstrates, response shift cannot be considered merely a statistical artifact.

Conclusions

An adequate description of QOL appraisal is fundamental to our understanding of response shift phenomena. Findings from this proposed line of research should yield an approach to QOL assessment that surpasses the relatively superficial treatments of QOL currently available. Studies using comprehensive methodology to assess appraisal will help us to determine what should be included in briefer, more portable appraisal assessments. We envision studies of QOL outcomes designed according to the models presented in Figures 2 and 3, which use direct measures of appraisal to account for inter-individual and temporal differences in the meaning attributed to QOL scales.

This paper proposes that QOL response shift may best be understood as an epiphenomenon: individuals' ratings of QOL can respond to changes in illness, treatment and other life events in atypical (e.g., statistically different from some expected value) ways or in ways that do not gibe with external observation. Changes in QOL appraisal may be able to account for these discrepancies. This definition of response shift provides a way to unify and harmonize many of the different methods that have been used in the literature. Some studies have attempted to infer changes in appraisal parameters from changes in coping mechanisms. These are related but not identical (e.g., determining that one has coped by getting a better outlook does not, in and of itself, describe what that new outlook is). Other studies have used the retrospective pretest to obtain a discrepancy score, but have not assessed the psychological reasons underlying this discrepancy. Still other studies have inferred changes in appraisal in a sample based on changes in the factor structure of items. Although such methods can be used to point to possible QOL response shifts, actual measurement of response shift per se requires direct assessment of changes in appraisal to account for discrepant changes in QOL ratings.

Elucidating QOL appraisal processes over time should lead to a more interpretable link between patient-reported indicators of QOL and external observers' (e.g., clinicians, caregivers) perspectives. Although we have focused this discussion around self-reported measures, this model of appraisal can readily be extended to clinician judgments, proxy ratings, and the like. The reason for discrepant scores between concurrent ratings of QOL measures from different perspectives ought to be explained by differences in perspective. Indeed, it would be interesting to observe whether discrepant scores between observers (e.g., patient and care giver) fall into line if they are first asked to come to consensus on what QOL appraisal criteria they will use.

Appraisal concepts and methods have bearing on the emerging interest in the use of cognitive assessment of survey methods applied to QOL research [46,67]. Cognitive methods attempt to determine the appraisal processes associated with a given item, scale or instructional set. Consistent with Sprangers and Schwartz' [1] original response shift model and the appraisal model presented here, cognitive techniques emphasize the assessment of psychological processes of comprehension, retrieval, judgment, and response. Cognitive methods have been used in a variety of research areas to arrive at self-report measures that have well-articulated, and widely shared meanings and to facilitate comparisons across individuals and over time. However, there is an important tension between applications of cognitive methods to refine standard QOL measures and the methods presented here.

Refinement of QOL measures to narrow the range of appraisals that they pull for may be quite valuable. Indeed, measures presented herein could be used as adjunct to cognitive interviews to determine how well this goal is achieved. For example, do most people answering a given QOL measure adopt the same frame of reference, rely on similar standards of comparison, or prioritize responses in the same (or at least, similar enough) ways?

There is a danger, however, that writing measures to constrict or constrain differences in QOL appraisals for the sake of comparability of ratings may obscure important aspects of individual experience. Methods discussed here and other cognitive methods provide a viable alternative to one-size-fits-all approaches. By including direct measures of appraisal parameters as an adjunct to standard QOL ratings, individual differences in cognitive processes can be detected and controlled in outcomes research. Indeed, in some studies, change in appraisal parameters themselves may be the main phenomenon of interest. In either case, both evidence and logic compel us to conclude that the study of change in QOL requires methodology to articulate the process of appraisal. By building on the Sprangers and Schwartz [1] model and explicitly integrating a formulation of QOL appraisal in both statistical and judgment models of QOL, we hope to pave the way for research that links response shift phenomena to other critical areas of research where self-evaluation comes into play.

Additional material

Additional file 1

Click here for file
[<http://www.biomedcentral.com/content/supplementary/1477-7525-2-14-S1.doc>]

Acknowledgements

We gratefully acknowledge Mirjam Sprangers, Ph.D. and Kathleen Wyrwich, Ph.D. for their helpful comments and discussions as the ideas in this manuscript evolved. We also want to acknowledge the participants in the University of Hull symposium on "Assessing Health-Related Quality of Life – What Can the Cognitive Sciences Contribute?" (December, 2000) and the subsequent special issue of *Quality of Life Research* (Volume 12, Number 3, May 2003), organized and edited by Ivan Barofsky, Ph.D., Keith Meadows, Ph.D. and Elaine McColl, Ph.D. These opportunities for scholarly exchange encouraged us to formulate the ideas in this paper in much greater depth.

References

- Sprangers MAG, Schwartz CE: **Integrating response shift into health-related quality-of-life research: A theoretical model.** *Social Science and Medicine* 1999, **48**:1507-1515.
- Andrykowski MA, Brady MJ, Hunt JW: **Positive psychosocial adjustment in potential bone marrow transplant recipients: cancer as a psychosocial transition.** *Psycho-Oncology* 1993, **2**:261-276.
- Breetvelt IS, van Dam FSAM: **Underreporting by cancer patients: the case of response shift.** *Social Science & Medicine* 1991, **32**:981-987.
- Bach JR, Tilton MC: **Life satisfaction and well-being measures in ventilator assisted individuals with traumatic tetraplegia.** *Arch Phys Med Rehabil* 1994, **75**:626-632.
- Cassileth BR, Lusk EJ, Tenaglia AN: **A psychological comparison of patients with malignant melanoma and other dermatologic disorders.** *Am Acad Dermatol* 1982, **7**:742-746.
- Groenvold M, Fayers PM, Sprangers MAG, Bjorner JB, Klee MC, Aaronson NK, Bech P, Mouridsen HT: **Anxiety and depression in breast cancer patients at low risk of recurrence compared with the general population – unexpected findings.** *Journal of Clinical Epidemiology* 1999, **52**:523-530.
- Stensman R: **Severely mobility-disabled people assess the quality of their lives.** *Scandinavian Journal of Rehabilitation Medicine* 1985, **17**:87-99.
- Albrecht GL, Devlieger PJ: **The disability paradox: high quality of life against all odds.** *Social Science and Medicine* 1999, **48**:977-988.
- Kagawa-Singer M: **Redefining health: living with cancer.** *Social Science and Medicine* 1993, **37**:295-304.
- Padilla GV, Mishel MH, Grant MM: **Uncertainty, appraisal and quality of life.** *Quality of Life Research* 1992, **1**:155-165.
- Wilson IB, Cleary PD: **Linking clinical variables with health-related quality of life: a conceptual model of patient outcomes.** *JAMA* 1995, **273**:59-65.
- Friedland J, Renwick R, McColl M: **Coping and social support as determinants of quality of life in HIV/AIDS.** *AIDS Care* 1996, **8**:15-31.
- Sneeuw KC, Aaronson NK, Sprangers MA, Detmar SB, Wever LD, Schornagel JH: **Value of caregiver ratings in evaluating the quality of life of patients with cancer.** *J Clin Oncol* 1997, **15**:1206-1217.
- Sprangers MAG, Aaronson NK: **The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: A review.** *Journal of Clinical Epidemiology* 1992, **45**:743-760.
- Slevin ML, Stubbs L, Plant HJ, Wilson P, Gregory WM, Armes PJ, Downer SM: **Attitudes to chemotherapy: comparing views of patients with cancer with those of doctors, nurses, and general public.** *British Medical Journal* 1990, **300**:1458-1460.
- Leventhal H, Colman S: **Quality of life: A process view.** *Psychology and Health* 1997, **12**:753-767.
- Schwartz CE, Sprangers MAG: **Methodological approaches for assessing response shift in longitudinal quality of life research.** *Social Science and Medicine* 1999, **48**:1531-1548.
- Campbell A: **Subjective measures of well-being.** *American Psychologist* 1976, **31**:117-124.
- Cantril H: **The pattern of human concerns.** New Brunswick, NJ: Rutgers University Press; 1966.
- Stewart AL, Napoles-Springer A: **Health-related quality of life assessments in diverse population groups in the United States.** *Med Care* 2000, **38**(Review Suppl):II-102-II-124.
- Howard GS, Ralph KM, Gulanick NA, Maxwell SE, Nance SW, Gerber SK: **Internal invalidity in pre-test-posttest self-report evaluations and a re-evaluation of retrospective pre-tests.** *Applied Psychological Measurement* 1979, **3**:1-23.
- Howard GS, Dailey PR, Gulanick NA: **The feasibility of informed pre-tests in attenuating response-shift bias.** *Applied Psychological Measurement* 1979, **3**:481-494.
- Howard GS, Millham J, Slaten S, O'Donnell L: **Influence of subject response style effects on retrospective measures.** *Applied Psychological Measurement* 1981, **5**:89-100.
- Howard GS, Schmeck RR, Bray JH: **Internal invalidity in studies employing self-report instruments. A suggested remedy.** *Journal of Educational Measurement* 1979, **16**:129-135.
- Howard GS, Dailey PR: **Response shift bias: a source of contamination of self-report measures.** *Journal of Applied Psychology* 1979, **64**:144-150.
- Golembiewski RT, Billingsley K, Yeager S: **Measuring change and persistence in human affairs: Types of change generated by OLD designs.** *J Applied Behavioral Science* 1976, **12**:133-157.

27. Schwartz CE, Sprangers MAG, Eds: *Adaptation to changing health: Response shift in quality of life research* American Psychological Association: Washington DC; 2000.
28. Sprangers MAG, van Dam FSAM, Broersen J, Lodder L, Wever L, Visser MRM, Oosterveld P, Smets EMA: **Revealing response shift in longitudinal research on fatigue: The use of the then-test approach.** *Acta Oncologica* 1999, **38**:709-718.
29. Jansen SJ, Stiggelbout AM, Nooij MA, Noordijk EM, Kievit J: **Response shift in quality of life measurement in early-stage breast cancer patients undergoing radiotherapy.** *Qual Life Res* 2000, **9**:603-615.
30. Schwartz CE, Coulthard-Morris L, Cole B, Vollmer T: **The quality-of-life effects of Interferon-Beta-1b in multiple sclerosis: An Extended Q-TWiST analysis.** *Archives of Neurology* 1997, **54**:1475-1480.
31. Adang EMM, Kootstra G, Engel GL, van Hooff JP, Merckelback HLGJ: **Do retrospective and prospective quality of life assessments differ for pancreas-kidney transplant recipients?** *Transpl Int* 1998, **11**:11-15.
32. Ahmed S, Mayo NE, Wood-Dauphinee S, Hanley J: **Response shift in the assessment of health-related quality of life (HRQL) post-stroke [Abstract].** *Quality of Life Research* 2001, **10**:204.
33. Hagedoorn M, Sneeuw KC, Aaronson NK: **Changes in physical functioning and quality of life in patients with cancer: Response shift and relative evaluation of one's condition.** *J Clin Epidemiol* 2002, **55**:176-183.
34. Schwartz CE, Feinberg RG, Jilinskaia E, Applegate JC: **An evaluation of a psychosocial intervention for survivors of childhood cancer: Paradoxical effects of response shift over time.** *Psycho-Oncology* 1999, **8**:344-354.
35. Rees JE, Waldron D, O'Boyle CA, MacDonagh RP: **Response shift in individualized quality of life in patients with advanced prostate cancer [abstract].** *Clinical Therapeutics* 2002, **24**(Supplement B):33-34.
36. Lenert LA, Treadwell JR, Schwartz CE: **Associations between health status and utilities: Implications for policy.** *Medical Care* 1999, **37**:479-489.
37. Cella D, Hahn EA, Dineen K: **Meaningful change in cancer-specific quality of life scores: The direction of change matters.** *Quality of Life Research* 2002, **11**:207-221.
38. Lepore S, Eton D: **Response Shifts in Prostate Cancer Patients: An Evaluation of Suppressor and Buffer Models.** In: *Adaptation to changing health: Response shift in quality of life research* Edited by: Schwartz CE, Sprangers MAG. American Psychological Association: Washington DC; 2000.
39. Rapkin BD, Fischer K: **Personal Goals of Older Adults: Issues in Assessment and Prediction.** *Psychology and Aging* 1992, **7**:127-137.
40. Richards A, Folkman S: **Response shift: A coping perspective.** In: *Adaptation to changing health: Response shift in quality of life research* Edited by: Schwartz CE, Sprangers MAG. American Psychological Association, Washington D.C.; 2000.
41. Daltroy LH, Larson MG, Eaton HM, Phillips CB, Liang MH: **Discrepancies between self-reported and observed patient function in the elderly: the influence of response shift and other factors.** *Social Science and Medicine* 1999, **48**:1549-1561.
42. Daltroy LH, Phillips CB, Eaton HM, Larson MG, Partridge AJ, Logigian M, Liang MH: **Objectively measuring physical ability in elderly persons: The Physical Capacity Evaluation.** *American Journal of Public Health* 1995, **85**:558-560.
43. Fries JF, Spitz P, Kraines RG, Holman HR: **Measurement of patient outcome in arthritis.** *Arthritis and Rheumatism* 1980, **23**:137-145.
44. Rapkin B: **Personal goals and response shifts: Understanding the impact of illness and events on the quality of life of people living with AIDS.** In: *Adaptation to changing health: Response shift in quality of life research* Edited by: Schwartz CE, Sprangers MAG. American Psychological Association, Washington D.C.; 2000.
45. Tourangeau R, Rips R, Rasinski K: *The Psychology of Survey Response* Cambridge: Cambridge University Press; 2000.
46. Jobe JB: **Cognitive psychology and self-reports: models and methods.** *Qual Life Res* 2003, **12**(3):219-227.
47. Krause NM, Jay GM: **What do global self-rated health items measure?** *Medical Care* 1994, **32**:930-942.
48. Wilson ME: **A procedure for the determination of thresholds in impaired sensory fields.** *J. Neurol. Neurosurg. Psychiat* 1969, **32**:419-422.
49. McGee H, O'Boyle CA, Hickey A, O'Malley K, Joyce CRB: **Assessing the quality of life of the individual: the SEIQoL with a healthy and a gastroenterology unit population.** *Psychological Medicine* 1991, **21**:749-759.
50. Gonzalez-Calvo J, Gonzalez VM, Lorig K: **Cultural diversity issues in the development of valid and reliable measures of health status.** *Arthritis Care Res* 1997, **10**:448-456.
51. McGraw SA, McKinlay JB, Crawford SA, Costa LA, Cohen DL: **Health survey methods with minority populations: some lessons from recent experience.** *Ethnicity & Disease* 1992, **2**(3):273-287.
52. Allison PJ, Locker D, Feine JS: **Quality of life: A dynamic construct.** *Social Science and Medicine* 1997, **45**:221-230.
53. Bernhard J, Hürny C, Maibach R, Herrmann R, Laffer U: **Quality of life as subjective experience: Reframing of perception in patients with colon cancer undergoing radical resection with or without adjuvant chemotherapy.** *Annals of Oncology* 1999, **10**:775-782.
54. Hoeymans N, Feskens EJM, Kromhout D, VanDenBos GAM: **Ageing and the relationship between functional status and self-rated health in elderly men.** *Social Science and Medicine* 1997, **45**:1527-1536.
55. Stanton AL, Danoff-Burg S, Cameron CL, Snider PR, Kirk SB: **Social comparison and adjustment to breast cancer: An experimental examination of upward affiliation and downward evaluation.** *Health Psychology* 1999, **18**:151-158.
56. Gruder CL: **Choice of comparison persons in evaluating oneself.** In: *Social comparison processes: theoretical and empirical perspectives* Edited by: Suls JM, Miller RL. Hemisphere Publishing Corporation, Washington, D.C.; 1977.
57. Blalock SJ, DeVellis BM, DeVellis RF, Giorgino KB, Sauter SVH, Jordan JM, Keefe FJ, Mutran EJ: **Psychological well-being among people with recently diagnosed rheumatoid arthritis.** *Arthritis and Rheumatism* 1992, **35**:1267-1272.
58. Ramund B, Stensman R: **Quality of life and evaluation of functions among people with severely impaired mobility and non-disabled controls.** *Scandinavian Journal of Psychology* 1988, **29**(3-4):137-44.
59. Klinger E, Barta S, Maxeiner M: **Current concerns: Assessing therapeutically relevant motivation.** In *Assessment strategies for cognitive-behavioral intervention* Edited by: Kendall P, Holton S. New York: Academic Press; 1981:161-196.
60. Pals TS, Little BR: **Perceived Life Satisfaction and the Organization of Personal Project Systems.** *Journal of Personality and Social Psychology* 1983, **44**:1221-1230.
61. Rapkin BD, Smith MY, DuMont K, Correa A, Palmer S, Cohen S: **Development of the ideographic functional status assessment: A measure of the personal goals and goal attainment activities of people with AIDS.** *Psychology and Health* 1993, **9**:111-129.
62. Warnecke RB, Ferrans CE, Johnson TP, Chapa-Resendez G, O'Rourke DP, Chavez N, Dudas S, Smith ED, Martinez Schallmoser L, Hand RP, Lad T: **Measuring quality of life in culturally diverse populations.** *J Nat Cancer Inst Monogr* 1996, **20**:29-38.
63. Suls JM, Miller RL: *Social Comparison Processes: Theoretical and Empirical Perspectives* New York: John Wiley & Sons; 1977.
64. Schwartz CE, Mathias SD, Pasta DJ, Colwell HH, Rapkin BD, Genderson MW, Henning JM: **A comparison of two approaches for assessing patient importance weights to conduct an Extended Q-TwiST analysis.** *Quality of Life Research* 1999, **8**(3):197-207.
65. Sprangers M, Hoogstraten J: **Pretesting effects in retrospective pretest-posttest designs.** *Journal of Applied Psychology* 1989, **74**:265-272.
66. Norman G: **Hi! How are you? Response shift, implicit theories, and differing epistemologies.** *Qual Life Res* 2003, **12**(3):239-249.
67. Collins D: **Pretesting survey instruments: an overview of cognitive methods.** *Qual Life Res* 2003, **12**(3):229-238.