Understanding changes in quality of life in cancer patients: a cognitive interview approach
Bloem, E.F.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
CHAPTER 5

Using a retrospective pretest instead of a conventional pretest is replacing biases: a qualitative study examining cognitive processes underlying responses to the test items

Taminiau-Bloem EF, van Zuuren FJ, Koeneman MA, Visser MRM, Tishelman C, Koning CCE, Sprangers MAG

Submitted
Abstract

Background
The retrospective pretest-posttest or thentest design is the most commonly used approach to detect and control for recalibration response shift in the prospective measurement of quality of life. The thentest design assumes (1) consistency in the content of the cognitive processes underlying patients’ QoL assessments between posttest and thentest, and (2) accurate recall of pretest functioning. Our objective is to qualitatively examine both assumptions. Building on our prior study [1], we will additionally compare consistency in the content of respondents’ cognitive processes between posttest and thentest with consistency between posttest and conventional pretest.

Methods
We conducted think-aloud interviews with 24 cancer patients prior to and at the end of radiotherapy to elicit cognitive processes underlying their assessment of 7 QoL items at pretest, posttest and thentest. Qualitative analysis of all responses was independently carried out by two researchers. We used an analysis scheme based on the cognitive process models of Tourangeau et al. and Rapkin & Schwartz that yielded five cognitive processes.

Results
The interviews yielded 162 responses per assessment. Regarding the first assumption, comprehension/frame of reference changed in 100 (of 153 applicable) comparisons of posttest and thentest responses; retrieval/sampling strategy in 101 (of 162); standards of comparison in 81 (of 162); judgment/combinatory algorithm in 37 (of 63); and reporting and response selection in 77 (out of 162). Overall, in 157 comparisons (97%) the content of at least one cognitive component changed between posttest and thentest. The content of the five cognitive processes changed in 35 (of 63) to 116 (of 162) comparisons of posttest and pretest responses. With respect to the second assumption, the time frame employed and/or description of pretest functioning provided in 102 thentest responses differed from those provided in the corresponding pretest items.

Conclusions
Both assumptions underlying the thentest design appear not to be supported by the patients’ cognitive processes underlying at least half of their responses. Replacing the conventional pretest with a retrospective pretest may simply be replacing one set of biases for another.
Background

Change in patients’ quality of life (QoL) is most commonly assessed by means of the prospective baseline and follow-up design, also known as the pretest-posttest design [e.g. 2-5]. However, response shift may occur in the interim, which is defined as a change in internal standards (i.e. recalibration), values (i.e. reprioritization) and/or concept of QoL (i.e. reconceptualization) as a result of health changes [6]. Response shift may pose a serious threat to the pretest-posttest design, since it may render QoL assessments over time incomparable [e.g. 7-10].

The retrospective pretest-posttest design is the most commonly used approach to detect and control for recalibration response shift when measuring change in QoL [e.g. 11-14]. The retrospective pretest or thentest extends the pretest-posttest design with a retrospective evaluation of an earlier assessment. Most times, respondents complete the conventional posttest assessment and are subsequently asked to complete the same questions again but with the instruction to report how they perceive themselves to have been prior to the pretest [15]. The first assumption of this design is that by taking posttest and thentest in close proximity, the content of respondents’ underlying cognitive process will be consistent between posttest and thentest. Consequently, comparison of posttest and thentest scores would eliminate treatment induced response shift, and provide an unconfounded assessment of the treatment effect. In addition, the mean change in scores from pretest to thentest would provide an indication of the magnitude and direction of recalibration response-shift effect [16]. The second assumption is that patients are able to accurately recall their pretest functioning when completing the thentest. However, the thentest has been criticized for its susceptibility to memory distortion of pretest functioning [15, 17]. Interestingly, both assumptions have never been tested directly.

To the best of our knowledge, only Westerman et al. [18, 19] have qualitatively examined response strategies underlying QoL assessment using a thentest design. In their study, 23 small-cell lung cancer patients participated in cognitive think-aloud interviews at four time points during a treatment trajectory with chemotherapy. In presenting their results, Westerman et al. focus on patients’ response strategies in the prospective measurement of QoL over time (i.e. comparison of pretest and posttest scores), rather than on the cognitive processes used in answering thentest items. However, they do mention that the interview transcripts indicate that patients have difficulty recalling the previous measurement point and/or their prior functioning when responding to the thentest items [18].

To examine the cognitive processes underlying cancer patients’ responses to QoL (then-test) items, we have developed an analysis scheme [20] based on the frameworks of Tourangeau et al. [21] and Rapkin & Schwartz [22]. Combined, these models distinguish five cognitive processes underlying QoL responses: 1) induction of a frame of reference, 2) recall of relevant information (i.e. samples), 3) use of standards of comparison against which the retrieved information is judged, 4) use of an algorithm to prioritize and combine the retrieved information, and 5) reporting and response selection (see [1]).

In a prior study with 50 cancer patients, we used this analysis scheme to examine the
content of the cognitive processes underlying patients’ responses to QoL items in a pretest-posttest design. A specific objective was to study the assumption of consistency in the content of respondents’ cognitive processes over time. Our results showed that the content of all cognitive processes changed over time, ranging from 113 (of 220) to 246 (of 342) pretest-posttest comparisons [1]. The present study builds on these prior results by examining whether the content of each distinct cognitive process underlying patients’ responses remains similar or rather changes between posttest and thentest. Since the thentest design is assumed to control for inconsistencies in respondents’ QoL assessment between pretest and posttest, we expect more similarity in the content of the cognitive processes underlying responses to the posttest and thentest. To examine this hypothesis, we will compare the results of the current study (thentest design) with those of the same subgroup of patients that participated in the previous study (pretest-posttest design). Finally, to comprehensively address both assumptions underlying the thentest design, we will also examine whether patients accurately recall their pretest functioning when completing thentest items.

Methods
This study was part of our prior study [1], in which we examined the consistency assumption in the pretest-posttest design. This study consisted of two consecutively conducted studies: in one study the pretest and posttest assessments were extended with a thentest (current study), and in the other study with transition questions (manuscript submitted for publication). Since the current study includes a subsample of the previous study (N=50), data will only be provided of the subjects participating in the current, thentest design, study. We refer to our prior publication for details on the material and methods used [1]. In this section we will provide a brief summary of the methods used, and will only expand on information that is unique for the present study.

Participants
To include a heterogeneous sample and variation in cognitive processes used in arriving at responses to QoL (thentest) items, we purposefully selected cancer patients undergoing radiotherapy at the Department of Radiation Oncology at the Academic Medical Center (AMC) in Amsterdam. Since participation in this study was not considered intrusive and solely consisted of self-reports, the Medical Ethics Committee (MEC) of the AMC provided exemption from seeking formal approval.

Procedure
In accordance with the design commonly used in treatment evaluation, pretest assessments were administered prior to, and posttest and thentest assessments at the end of a salient health-related intervention. Pretest interviews took place on the same day the patient had a CT-simulator appointment to plan treatment or received first radiation treatment.
posttest and thentest interviews were conducted on patients’ last day of radiotherapy. The interviews were conducted by two researchers (ETB, MAK) not involved in the patient’s clinical care. All interviews were audio-recorded and transcribed verbatim.

To limit patient burden, we selected seven items from the 30-item EORTC QLQ-C30 [23] - a HRQoL instrument widely used in European cancer clinical trials [24] - covering both global and specific content, including physical, psychological and social dimension (see Table 2). All items employ a one week time frame in accordance with the EORTC QLQ-C30 instructions. The thentest questions were adapted versions of these items, e.g. How would you rate your overall QoL in the week prior to the first interview?

We used Hak et al’s Three-Step Test Interview [25] that combines cognitive think aloud interviewing and verbal probing techniques [26] at the pretest, posttest and thentest assessments to enable comparisons of patients’ cognitive processes. In the think-aloud interview, we asked the patients to read each question out loud and subsequently think out loud in assigning a score to the question. Immediately after the think-aloud response to each item, we attempted to elicit more information about participants’ cognitive processes using probes based on the cognitive process models of Tourangeau et al. [21] and Rapkin & Schwartz [22] (see [1]).

After completion of the posttest think-aloud interview, we introduced the thentest by jogging the patient’s memory about the time s/he conducted the pretest think aloud interview. As recommended, we called the day and time of the pretest interview to the patient’s mind, and provided cues to elicit patient’s memory [27]: “Please take a minute to think back to the first interview. The following questions concern the week immediately prior to this interview and the start of radiotherapy. At that time, you might have felt tense or sad, or maybe you didn’t feel tense or sad at all. You might have suffered from physical complaints, or maybe you didn’t suffer from physical complaints at all. Can you remember how you were feeling at that time?” Subsequently, we asked the patients to provide a new judgment about their QoL at the time they conducted the pretest think aloud interview. We emphasized that we were not asking patients to recall their pretest response, but rather to provide a renewed judgment. We used non-leading probes for the thentest items, such as “You just referred to your functioning prior to the start of radiotherapy, could you explain to me how you tried to recall this period?”

Data analysis

Qualitative analysis of all interviews was independently carried out by two researchers (ETB, MK) using MAXqda software [28]. As was the case in our prior study [1], we used our qualitative analysis scheme [20] based on the cognitive process models of Tourangeau et al. [21] and Rapkin & Schwartz [22] for the coding of patients’ cognitive processes. The two researchers independently assigned codes related to the content of the underlying cognitive processes to all items of the pretest, posttest, and thentest interviews of each patient. They discussed their findings and, in case of differences, achieved agreement about the assigned codes through negotiated consensus [29]. Once agreement was established, the assumption of consistency in the content of the cognitive processes underlying QoL
appraisal between posttest and thentest was examined. To that end, the researchers independently compared the content of each cognitive process at posttest and thentest to determine whether it was similar or had changed (see [1] for an illustration of the use of our analysis scheme, and examples of similarity and dissimilarity in the content of all five cognitive processes for all seven items).

The second assumption of accurate recall underlying the thentest was operationalized by examining whether the time frame employed and the description of pretest functioning provided in answering each thentest item were similar to those of the corresponding pretest item. Again, all findings were discussed and consensus was negotiated in case of differences. Throughout the period of data collection and analysis, all codes and subsequent analyses were also discussed with co-authors FvZ and MS.

In our prior studies, we found that the content of the cognitive processes used by a respondent to assess QoL over time was not constant for all questionnaire items, but instead, varied per item [1, 20]. Again, in analyzing patients’ think aloud response to each thentest item, we found that the content of the cognitive processes differed within the same patient across items. In the present study, we therefore studied each response to all items separately.

Results

Of 38 eligible patients approached, 12 refused to participate explaining they considered it too burdensome to be cognitively interviewed prior to and at the end of radiotherapy. Twenty-six patients gave written informed consent. One patient was unable to participate in the posttest and thentest interviews due to severe health deterioration, and one patient could not be interviewed at the end of radiotherapy due to logistical problems. The median number of days between the pretest interview, and the posttest and thentest interviews was 46 days (Mean 44 days, SD 8.7, range 27-57). Table 1 depicts the characteristics of the 24 patients who completed all three interviews (median age 61 years, SD 9.7, range 46-82). Twenty patients completed all seven items for all three assessments, with an additional two patients providing interpretable data for six items, and another two patients for five items. This yielded 162 responses suitable for qualitative analysis per assessment.
Table 1- Patient characteristics

<table>
<thead>
<tr>
<th>Gender:</th>
<th>No. of patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>12</td>
</tr>
<tr>
<td>Women</td>
<td>12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>No. of patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>40-49</td>
<td>4</td>
</tr>
<tr>
<td>50-59</td>
<td>7</td>
</tr>
<tr>
<td>60-69</td>
<td>10</td>
</tr>
<tr>
<td>70-79</td>
<td>2</td>
</tr>
<tr>
<td>≥ 80</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tumor site:</th>
<th>No. of patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bladder</td>
<td>2</td>
</tr>
<tr>
<td>Breast</td>
<td>4</td>
</tr>
<tr>
<td>Colorectal</td>
<td>4</td>
</tr>
<tr>
<td>Esophageal</td>
<td>4</td>
</tr>
<tr>
<td>Gynecological</td>
<td>3</td>
</tr>
<tr>
<td>Lung</td>
<td>2</td>
</tr>
<tr>
<td>Prostate</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Length of interval between baseline and follow-up interview (median)</th>
<th>No. of patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 46 days</td>
<td>12</td>
</tr>
<tr>
<td>≥ 46 days</td>
<td>12</td>
</tr>
</tbody>
</table>

Assumption 1 of the thentest design: similarity in the cognitive processes underlying responses to the posttest and thentest

1) Comprehension / frame of reference

Seven different patients could not provide a frame of reference (i.e. definition of the item’s target construct) for nine individual items during either the posttest or thentest assessments. Therefore, we could determine (dis)similarity in this cognitive process for 153 out of 162 comparisons. We found that the content of this cognitive process changed in 100 out of 153 paired posttest and thentest responses (see Table 2). The following interview excerpts illustrate change in the frame of reference of ‘fatigue’. In answering the posttest item ‘Were you tired?’ the patient defined fatigue as “loss of energy”; “I am a little tired, I do notice that I am out of form. I kept putting on my track-suit to go for a run. But I noticed that when I had run 10 minutes on end, I lost my energy whereas normally that isn’t troublesome for me at all.” (Answer: a little). However, in answering this item in thentest version (i.e. ‘Were you tired in the week prior to the first interview?’), the patient defined fatigue as “being out of balance”; “I was out of balance because I was feeling tense.” (Answer: a little) [Female, 48 years, breast cancer].
2) Retrieval / sampling strategy

(Dis)similarity in retrieval/sampling strategy could be assessed for all 162 comparisons of posttest and thentest responses. Patients’ sampling strategy changed in 101 out of 162 of these comparisons (see Table 2). In the majority of responses at posttest (N=105), patients retrieved experiences from their radiotherapeutic treatment. In responses to the thentest items, however, most patients retrieved samples from the period prior to (N=47 responses) or following cancer diagnosis (N=53 responses). For example, in responding to the posttest question ‘Has your physical condition or medical treatment interfered with your social activities?’ the following patient retrieved a sample from his radiotherapeutic treatment: “You have to drive to the hospital and back home every day, which really is a burden. And during that time you can’t be in contact with other people.” (Answer: a little). In responding to the corresponding thentest item, this patient retrieved a sample from the period prior to cancer diagnosis: “No, not at all. I felt healthy, I wasn’t sick at that time. I kept company with the neighbours and close friends as always.”

[Male, 78 years, esophageal cancer]

3) Standards of comparison

In all 162 comparisons of posttest and thentest responses, the reference group(s) patients used to rate their own functioning could be discerned. Patients’ standards of comparison were different in 81 out of 162 comparisons of posttest and thentest responses (see Table 2). Patients made a comparison with their own functioning prior to cancer diagnosis and treatment in 108 responses at the posttest and in 73 responses at the thentest. However, in 23 thentest responses they used their current, posttest functioning as reference in providing their retrospective assessment, and in 12 responses varying comparisons were used, e.g. own functioning after surgery. For example, when answering to the posttest, a patient referred to her QoL prior to cancer diagnosis and treatment in rating her overall QoL: “I hope my life will re-establish the way it used to be, and that was pretty good. (…) Currently, I feel put out of action. The only thing you do is lie on the couch, you no longer take part in normal, daily life.” (Score: 3) In answering the thentest item, this patient referred to her current, posttest functioning: “My quality of life was fine at that time. The fact that I already was diagnosed with cancer played a part, but I was able to do everything I wanted back then. And that is very different compared to my current quality of life.”

[Score 7; Female, 59 years, gynaecological cancer]

4) Judgment / Combinatory algorithm

In responding to a QoL item, patients can retrieve positive samples (e.g. “I’m not tired at all, because I still sleep soundly.”) and negative samples (e.g. “The fact that I have to come to the hospital every day really tires me.”) [Male, 52 years, oesophageal cancer (both quotes)]. If patients retrieve both positive and negative samples, they need to combine this information to arrive at an answer. In doing so, patients can either emphasize the positive or negative samples, or find a balance between both. Patients retrieved positive and negative samples in 63 paired posttest and thentest responses. Dissimilarity in the cognitive process
judgment/combinatory algorithm was found in 37 out of these 63 comparisons (see Table 2). At posttest, patients emphasized positive samples in 23 responses, emphasized negative samples in 16 responses, and balanced between both types of samples in 24 responses. However, in the majority of the responses at posttest in which patients emphasized the negative samples, or in which both types of samples were balanced, patients emphasized the positive samples at thentest (N=10 and N=13 responses respectively). For example, in the following interview excerpts the patient emphasized the negative samples at posttest in responding to the item ‘How would you rate your overall health during the past week?’ in rating his health a ‘4’ since [his health] “wasn’t that good”: “I was able to carry out some nice activities, we visited some friends, my e-mail. But I am no longer able to do what I want to do, without considering whether or not the distance is too long or whether it will exhaust me. So for the past week, it wasn’t that good. I was very tired, a ‘4’.” whereas at thentest he emphasized the positive samples in retrospectively assessing his pretest health by choosing a ‘6’ because “I felt fine, I wasn’t in pain”; “I was suffering from the after-effects from my stroke back then, but I experienced lots of positive impulses. (...) I felt fine, I wasn’t in pain. So it was good, a ‘6’.” [Male, 58 years, prostatic cancer]

5) Reporting and response selection
In all 162 comparisons of responses, patients explained how they arrived at the selected response category when responding to the item. Changes in this cognitive process were found in 77 out of 162 comparisons of posttest and thentest responses (see Table 2). Patients arrived at their score in varying ways at both the posttest and thentest interviews. For example, patients employed editing processes aimed at mitigating their initial response at either the posttest or thentest assessment (N=14 responses). To elucidate further, patients did not refer to a previously provided score when answering the posttest items, however, in the majority of responses at thentest, patients arrived at their score by referring to the score previously provided to the same item at posttest (N=24 responses), or at pretest (N=7 responses), or by referring to their previously provided score to another thentest item (N=10 responses). To illustrate, at posttest the patient ‘normalized’ his score on ‘overall health’ (an editing process) by referring to his older age: “In the past couple of weeks, my health hasn’t been how I would like it to have been. But I guess it can’t be excellent anyway considering my age.” (Answer: 5). At the thentest, the patient retrospectively assessed his pretest health by referring to his score previously provided at the posttest: “It [health] wasn’t excellent and it wasn’t very poor either. So I’ll stick with the ‘5’ I’ve just chosen.” [Male, 78 years, esophageal cancer].
Table 2 - Assumption 1: Dissimilarity per cognitive process underlying responses to the posttest and thentest (numbers in bold) and to the posttest and pretest (numbers between brackets) for 7 EORTC QLQ-C30 items

<table>
<thead>
<tr>
<th>Comprehension / Frame of reference</th>
<th>1 Do you have any trouble taking a short walk outside of the house?</th>
<th>2 Have you had pain?</th>
<th>3 Were you tired?</th>
<th>4 Did you worry?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dissimilar</td>
<td>Number of responses: 18</td>
<td>Number of responses: 24</td>
<td>Number of responses: 22</td>
<td>Number of responses: 22</td>
</tr>
<tr>
<td>Dissimilar</td>
<td>Trouble: 10(8)</td>
<td>17(12)</td>
<td>15(12)</td>
<td>9(9)</td>
</tr>
<tr>
<td>Dissimilar</td>
<td>Walk: 0(1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dissimilar</td>
<td>Both: 2(4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dissimilar</td>
<td>Trouble and/or walk: 12(13)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Retrieval / Sampling strategy</th>
<th>15(15)</th>
<th>16(14)</th>
<th>11(18)</th>
<th>12(12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dissimilar</td>
<td>Number of responses: 23</td>
<td>Number of responses: 24</td>
<td>Number of responses: 23</td>
<td>Number of responses: 24</td>
</tr>
<tr>
<td>Standards of comparison</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dissimilar</td>
<td>9(4)</td>
<td>12(13)</td>
<td>12(11)</td>
<td>7(9)</td>
</tr>
<tr>
<td>Judgment / Combinatory algorithm</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dissimilar</td>
<td>Number of responses: 5</td>
<td>Number of responses: 8</td>
<td>Number of responses: 6</td>
<td>Number of responses: 9</td>
</tr>
<tr>
<td>Reporting and response selection</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dissimilar</td>
<td>7(5)</td>
<td>6(7)</td>
<td>8(9)</td>
<td>13(11)</td>
</tr>
</tbody>
</table>

Number of responses: 153
<table>
<thead>
<tr>
<th>5 Has your physical condition or medical treatment interfered with your social activities?</th>
<th>6 How would you rate your overall health during the past week?</th>
<th>7 How would you rate your overall Qol during the past week?</th>
<th>Total dissimilarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of responses: 21</td>
<td>Number of responses: 22</td>
<td>Number of responses: 23</td>
<td>Total number of responses: 153</td>
</tr>
<tr>
<td>Interference: 6(10)</td>
<td>Interference: 16(14)</td>
<td>Interference: 16(13)</td>
<td>100(90)</td>
</tr>
<tr>
<td>Social activities: 1(1)</td>
<td>Social activities: 15(17)</td>
<td>Social activities: 15(16)</td>
<td></td>
</tr>
<tr>
<td>Both: 9(6)</td>
<td>Both: 13(20)</td>
<td>Both: 19(21)</td>
<td></td>
</tr>
<tr>
<td>Interference and/or social activities: 15(17)</td>
<td>15(16)</td>
<td>13(20)</td>
<td></td>
</tr>
<tr>
<td>Number of responses: 23</td>
<td>Number of responses: 22</td>
<td>Number of responses: 23</td>
<td>Total number of responses: 162</td>
</tr>
<tr>
<td>Number of responses: 23</td>
<td>Number of responses: 22</td>
<td>Number of responses: 23</td>
<td>101(116)</td>
</tr>
<tr>
<td>Number of responses: 23</td>
<td>Number of responses: 22</td>
<td>Number of responses: 23</td>
<td></td>
</tr>
<tr>
<td>Number of responses: 12(8)</td>
<td>Number of responses: 14(14)</td>
<td>Number of responses: 15(11)</td>
<td>81(70)</td>
</tr>
<tr>
<td>Number of responses: 8</td>
<td>Number of responses: 14</td>
<td>Number of responses: 13</td>
<td></td>
</tr>
<tr>
<td>Number of responses: 4(6)</td>
<td>Number of responses: 7(12)</td>
<td>Number of responses: 7(4)</td>
<td>37(35)</td>
</tr>
<tr>
<td>Number of responses: 23</td>
<td>Number of responses: 22</td>
<td>Number of responses: 23</td>
<td>Total number of responses: 162</td>
</tr>
<tr>
<td>Number of responses: 7(7)</td>
<td>Number of responses: 18(14)</td>
<td>Number of responses: 18(12)</td>
<td>77(65)</td>
</tr>
</tbody>
</table>
Conclusion

Overall, in 157 out of 162 comparisons (97%) the content of at least one cognitive component changed between posttest and thentest. Dissimilarity in the content of the five cognitive processes varied per patient, i.e. in each patient the content of each cognitive process changed in a different number of items, and was similar for the remaining items.

Assumption 1:
Comparing the thentest design with the pretest-posttest design

The results on (dis)similarity of the cognitive processes underlying the QoL responses for the 26 patients who completed the pretest, posttest, and thentest are presented in Table 2. Contrary to expectation, the content of the cognitive processes was more often dissimilar between posttest and thentest than between posttest and pretest, with the exception of retrieval/sampling strategy.

Assumption 2 of the thentest design:
Accurate recall of pretest functioning – Comparing thentest and pretest

In comparing the thentest responses with the corresponding pretest responses, we found that in 10 out of 162 comparisons of responses patients had employed the one week time frame instructed by the EORTC QLQ-C30 at both time points. In the remaining thentest and pretest responses, a variety of different time periods were used. For the thentest these include: prior to cancer diagnosis and start of treatment, between cancer diagnosis and start of treatment, following other cancer treatment, but prior to radiotherapy, or other, e.g. since diagnosis of another illness. In further comparing the responses to the thentest items with the responses to the corresponding pretest items, we found four patterns (see Figure 1):

1) Similarity in time frame and description of pretest functioning:
In 60 responses to thentest items, the time frames employed and the descriptions of pretest functioning were consistent with those employed in the corresponding pretest items (see Figure 1, box 1). Thus, apart from not employing the time frames as instructed at both the pretest and thentest, these responses were in line with the assumption of accurate recall. In the following example, the patient referred to the time prior to cancer diagnosis and treatment at both pretest and thentest. Additionally, in answering the thentest item the patient provided a description of pretest functioning which is consistent with the description provided at pretest. At pretest, the patient responded to the question ‘Have you had pain?’ as follows: ‘Naturally, I suffer from some small complaints. But that doesn’t bother me that much. (...) I have had these complaints for years now, and you get accustomed to it. If you want to grow old, that’s part of the deal.’ (Answer: not at all). In answering the thentest item, she indicated: ‘I haven’t had pain. Only those small complaints, which are part of your body getting older.’ (Answer: not at all)
[Female, 67, oesophageal cancer].

154
2) Similarity in time frame and dissimilarity in description of pretest functioning:

In 35 responses to thentest items, the time frames employed were consistent with those employed in the corresponding pretest items. However, patients’ descriptions of pretest functioning differed between the responses to the pretest and thentest items (see Figure 1, box 2). Patients defined the target construct of the item differently at the thentest assessment and/or retrieved different information about their pretest functioning than was the case at the pretest assessment. The patient cited below referred to the period between cancer diagnosis and start of treatment at both assessments. However, he provided a different description of pretest functioning in both interviews. In response to the pretest question ‘Have you had pain?’, the patient refers to suffering from psychological pain: “I have experienced pain when I received the diagnosis of having cancer. (...) When they tell you that, your entire world collapses completely. Some time after that message you recuperate, but I really suffered from pain psychologically.” (Answer: quite a bit). Whereas in response to the thentest item, the patient retrospectively described to have suffered pain “both mentally and physically”: “At that time, I suffered ‘a little’ pain because I had trouble swallowing. You are rebellious after the diagnosis, and you keep eating certain things that you can’t eat anymore. That makes you angry, resulting in suffering pain both mentally and physically.” (Answer: a little) [Male, 52 years, esophageal cancer].
3) Dissimilarity in time frame and similarity in description of pretest functioning:
In 10 responses to thentest items, the time frames employed differed from those used in
the corresponding pretest items but the descriptions of pretest functioning were similar
to those provided in the corresponding pretest items (see Figure 1, box 3). To illustrate,
the patient cited below referred to the time prior to cancer diagnosis and treatment in
answering the pretest item ‘Do you have any trouble taking a short walk outside of the
house?’, whereas he referred to the period between cancer diagnosis and start of treat-
ment in answering the thentest item. However, the descriptions of pretest functioning are
similar. At pretest the patient responded as follows: “I haven’t got any trouble taking a walk
outside. (...) I am capable of doing so, and I generally like it.” (Answer: not at all). In answ-
ing the thentest item, the patient explained: “After diagnosis, I walked as I have always done
normally. Without limitations, and without shortcomings.” (Answer: not at all)
[Male, 64 years, colorectal cancer].

4) Dissimilarity in time frame and dissimilarity in description of pretest functioning:
In the remaining 57 responses to thentest items, both the time frames employed and the
descriptions of pretest functioning provided differed from those employed in the cor-
responding pretest items (see Figure 1, box 4). These differences were due to changes
in definition of the target construct and/or the retrieval of different information at the
thentest assessment. In the following example, the patient referred to the period prior to
cancer diagnosis and start of treatment at pretest, whereas in the thentest assessment she
referred to the period between cancer diagnosis and start of treatment. Additionally, in
response to the pretest question “Did you worry?”, she indicated: “At first, I thought I had a
haemorrhoid. But it got larger and larger until it grew out of my anus. At that stage I immediately
realised I had to visit a physician. I worried a lot at that point, something needed to be done
quickly.” (Answer: very much). Whereas in response to the thentest item, she explained:
“It’s too extreme to say I didn’t worry at all after receiving the diagnosis of having cancer. At
that time I worried a little about possible metastases. You are in need of so much information.”
(Answer: a little)
[Female, 51 years, colorectal cancer].

Discussion
In this study, we examined the assumption inherent to the thentest design that the content
of respondents’ cognitive processes underlying responses to the posttest and thentest are
consistent. In comparing paired posttest and thentest responses, we found changes in the
content of all cognitive processes. Per cognitive process, changes varied from 37 (of 63)
to 101 out of 162 comparisons of posttest and thentest responses. Contrary to our
expectation, we did not find more agreement between the cognitive processes underlying
responses to the posttest and thentest than between those underlying the posttest and
pretest. To the contrary, the results regarding both designs were highly comparable. Thus,
these findings do not support the thentest design’s assumption that the administration of
the posttest and thentest in temporal proximity would induce a comparable content in
cognitive processes. We also examined thentest’s second assumption of accurate recall of
pretest functioning. Our results show that in 102 out of 162 thentest responses, the time
frames employed and/or descriptions of pretest functioning provided differed from those
employed in the corresponding pretest items. Additionally, patients were found to employ
a variety of time frames besides the instructed one-week time frame in (retrospectively)
assessing pretest functioning. In conclusion, both assumptions underlying the thentest
design appear not to be supported by the patients’ cognitive processes underlying at least
half of their responses.

According to Rapkin & Schwartz, change in the content of each of the cognitive processes
constituting their QoL appraisal model is linked to one of the specific types of response
shift, i.e. change in frame of reference is related to reconceptualization, change in sam-
pling strategy and combinatorial algorithm to reprioritization, and change in standards of
comparison to recalibration [22]. The thentest design is devised to detect and control
for recalibration response shift [6], i.e. change in the respondent’s internal standards of
measurement. Hence, in using the thentest, one would particularly expect similarity in
the content of the cognitive process standards of comparison from posttest to thentest.

However, our data reveal changes in the content of this cognitive process in 81 out of 162
of such comparisons. Conversely, the content of standards of comparison changed in 70
out of 162 posttest and pretest comparisons. Thus, surprisingly, the content of standards of
comparison was more often dissimilar in the thentest design than in the pretest-posttest
design. Moreover, this finding suggests the occurrence of recalibration response shift in half
of these posttest and thentest comparisons.

Several studies have examined the accuracy of retrospective assessment of QoL as op-
posed to its prospective assessment. Some of these studies advocate the use of the then-
test for treatment evaluation, and consider it a more valid approach to measure change
than standard, prospective measurement [e.g. 9, 30-32]. Conversely, others raised validity
concerns, suggesting that recall bias, social desirability, and implicit theory of change may
play a role in retrospective assessments [e.g. 13, 33-35]. Our data indicate that patients
may have used implicit theories of change. According to such a theory, patients infer what
their pretest state must have been by extrapolating backwards from their current state
[34]. In this study, we found patients reconstructing their pretest functioning by using their
current, posttest functioning as standard of comparison in 23 responses to thentest items
(see results section - dissimilarity in standards of comparison).

To increase our insight into patients’ response strategies, including adaptive mechanisms
and response shift effects, future studies may be designed in which patients are confronted
with their pretest scores, after they have completed the thentest. Patients could subse-
quently be invited to comment to their scores. Such interviews may provide valuable
insight into patients’ own explanations of inconsistencies in the time frames employed
and the descriptions of pretest functioning provided. In addition, it would be interesting
to compare the content of all five cognitive processes between pretest and thentest to
further elucidate actual response shift effects.
As is common in employing a thentest, we first have revived patients' memory about their pretest functioning. The thentest is most commonly administered in a written format. However, in the present study, we have orally instructed patients to think back to their pretest functioning. It is plausible that an oral instruction and subsequent think-aloud interview might increase patients' efforts to recall their pretest functioning. Consequently, in the context of conventional treatment evaluation, the results may even show more deviation between pretest and thentest than was the case in the current study. The further limitations of this study are similar to those of the study examining the pretest-posttest design. As these have been addressed previously [1], we will briefly summarize them here. First, patients who refused participation explained they considered it too burdensome, which might indicate that the most severely ill patients were not included in this study's sample. The interpretation of our findings might thus be limited to the less severely ill. Second, we cannot be sure that cognitive think-aloud interviews adequately capture patients' cognitive processes. Third, the heterogeneity of the seven EORTC QLQ-C30 items that we used for our interviews might have induced differences in the content of the cognitive processes used in answering the consecutive items. That is, questionnaire items addressing one specific QoL domain might have resulted in more similarity in the content of the cognitive processes between posttest and thentest.

In conclusion, the cognitive processes underlying a little over half of patients' responses to thentest items appear not to be in line with the assumptions of (1) consistency in the content of the cognitive processes underlying responses to the posttest and thentest, and (2) accurate recall of pretest functioning. Rather, our data suggest that patients select personally meaningful time frames and content when (retrospectively) assessing their QoL, which might deviate from the time frames considered relevant by researchers. Since dissimilarity in the content of the cognitive process was found to be comparable for the pretest-posttest and the thentest design, it is questionable whether the thentest is a suitable alternative in controlling for inconsistencies in respondents' cognitive processes over time. However, as argued previously [36, 37], retrospective assessments, such as the thentest, are a useful method when the measurement goal is to describe change as experienced subjectively by the respondents. In interpreting thentest responses in the context of treatment evaluation, it is important to realize that patients provide assessments that are not necessarily based on the cognitive processes intended by researchers. Replacing the conventional pretest with a retrospective pretest may thus not resolve the bias underlying the prospective measurement of change in QoL, but rather may replace it for other biases.

Conclusions
The thentest design is the most commonly used approach to detect and control for recalibration response shift in the prospective measurement of QoL. In the present study we examined the assumption of consistency in the content of the cognitive processes underlying patients' QoL assessment between posttest to thentest. We found changes in the content of each of the five cognitive processes underlying QoL appraisal.
Cognitive processes underlying responses to thentest items

(i.e. comprehension/frame of reference, retrieval/sampling strategy, standards of comparison, judgment/combinatory algorithm, and reporting and response selection), varying from 37 (of 63) to 101 (of 162) comparisons of posttest and thentest responses. This study builds on our prior study, in which we examined the assumption of consistency in respondents’ QoL responses over time inherent to the pretest and posttest design [1]. In this design, the content of all cognitive processes changed in 35 (of 63) to 116 (of 162) comparisons of posttest and pretest responses. Contrary to our expectation, the content of the cognitive processes was more often dissimilar between posttest and thentest than between posttest and pretest, with the exception of retrieval/sampling strategy. To comprehensively address the assumptions underlying the thentest design, we have also examined whether patients accurately recall their pretest functioning when completing thentest items. Our results show that the time frame employed and/or description of pretest functioning provided in 102 thentest responses differed from those provided in the corresponding pretest items. In conclusion, both assumptions underlying the thentest design appear not to be supported by the patients' cognitive processes underlying at least half of their responses. Consequently, replacing the conventional pretest with a retrospective pretest may simply be replacing one set of biases for another.
References

1. Taminiau-Bloem EF, van Zuuren FJ, Koeneman MA, Rapkin BD, Visser MRM, Koning CCE, Sprangers MAG. A ‘short walk’ is longer before radiotherapy than afterwards: a qualitative study questioning the baseline and follow-up design. Health and Quality of Life Outcomes 2010; 8: 69


19. Westerman MJ, Hak T, Sprangers MAG, Groen HJM, van der Wal G, The AM. Listen to their answers! Response behaviour in the measurement of physical and role functioning. Quality of Life Research 2008; 17: 549-558
27. Schwartz CE, Sprangers MAG. Guidelines for improving the stringency of response shift research using the thentest. Quality of Life Research 2010; 19: 455-464
28. MAXqda, 2004 [www.maxqda.com]
30. Nieuwkerk PT, Tollenaar MS, Oort FJ, Sprangers MAG. Are retrospective measures of change in quality of life more valid than prospective measures? Medical Care 2007; 45: 199-205


36. Hermann D. Reporting current, past and change health status. What we know about distortion. Medical Care 1995; 33: AS77-88