

# Evaluating quality of life and response shift from a couple-based perspective: a study among patients with colorectal cancer and their partners

Marjan J. Traa · Johan Braeken · Jolanda De Vries ·  
Jan A. Roukema · Ricardo G. Orsini ·  
Brenda L. Den Oudsten

Accepted: 19 November 2014 / Published online: 28 November 2014  
© Springer International Publishing Switzerland 2014

## Abstract

**Objectives** To examine (1) measurement invariance of quality of life (QoL) domains over time for patients with colorectal cancer and partners (i.e., response shift—recalibration, reprioritization, and reconceptualization), (2) between dyad-member measurement invariance and (3) QoL trajectories.

**Methods** Participants completed the WHOQOL-Bref preoperative (Time-0) and 3 (Time-1) and 6 months (Time-2) postoperative. A stepwise procedure, using nested

factor models, examined the viability of restricting specific model parameters to be equal across measurements and between dyad members.

**Findings** No reconceptualization and reprioritization was detected, but indications for recalibration were present. Therefore, comparisons were restricted to group-level statistics at factor level. For patients, a decrease in the Physical Health domain occurred at Time-1 ( $p < 0.001$ ), with partial recovery to baseline at Time-2 ( $p = 0.055$ ). For partners, factor means in this domain remained constant ( $p's > 0.05$ ) and were at each time point higher than patients' factor means ( $p's < 0.05$ ). Patients' and partners' Psychological Health decreased at Time-1 ( $p's < 0.05$ ), with stabilization at Time-2 ( $p's > 0.05$ ). Patients and partners' Social Relationship factor means decreased at Time-1 ( $p's < 0.05$ ), which decreased further for patients ( $p = 0.011$ ) but stabilized for partners ( $p = 0.214$ ). Partners' factor means were only lower than patients' factor means at Time-1. A similar decrease in the Environmental domain factor means occurred for both patients and partners at Time-1 ( $p's < 0.05$ ), with stabilization at Time-2 ( $p's > 0.05$ ).

**Conclusion** Since both patients and partners are affected by the patients' disease and treatment, we recommend that attention is paid to the couple instead of solely the patient.

**Keywords** Response shift · Partners · Colorectal cancer · Dyadic · Quality of life · Measurement invariance

## Introduction

There is an increasing awareness that not only oncological outcomes, but also patient-reported outcomes, such as

---

M. J. Traa · J. De Vries · J. A. Roukema ·  
B. L. Den Oudsten (✉)  
CoRPS - Center of Research on Psychology in Somatic Diseases,  
Department of Medical and Clinical Psychology, Tilburg  
University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands  
e-mail: B.L.denOudsten@uvt.nl;  
b.l.denoudsten@tilburguniversity.edu

J. Braeken  
Research Methodology Group, Wageningen University and  
Research Center, Wageningen, The Netherlands

J. Braeken  
CEMO - Centre for Educational Measurement,  
University of Oslo, Oslo, Norway

J. De Vries  
Department of Medical Psychology, St. Elisabeth Hospital,  
Tilburg, The Netherlands

J. A. Roukema  
Department of Surgery, St. Elisabeth Hospital, Tilburg,  
The Netherlands

R. G. Orsini  
Department of Surgery, Catharina Hospital, Eindhoven,  
The Netherlands

quality of life (QoL), must be taken into account when evaluating the consequences of disease and treatment. However, QoL standards may change over the course of the disease trajectory [1–3]. This may be inherent to the process of adapting to the disease [1, 2]. Within chronic disease research, Sprangers and Schwartz [1, 2] were the first to theoretically and methodologically address changes in QoL standards. Theoretically, they operationalized the term response shift as a ‘meta-construct,’ which includes three interrelated constructs: Response shift is a change in the meaning of one’s self-evaluation of QoL as a result of (1) recalibration: a change in the respondent’s internal standards, (2) reprioritization: a change in the respondent’s values—the relative importance of an item on a QoL domain changed or the relative importance of the component domains constituting QoL changed, or (3) reconceptualization: a redefinition of QoL [1, 2]. Response shift poses methodological challenges for QoL research: In order to adequately compare assessments completed over time, the questionnaire has to be measurement invariant [4]. However, if response shift occurs, then this is not applicable. Therefore, evaluating response shift is necessary to allow adequate interpretation of QoL findings over time.

To our knowledge, only three studies examined response shift among patients with colorectal cancer [5–7], even though colorectal cancer is among the most prevalent cancer types [8]. A qualitative study found recalibration and reconceptualization in patients with a temporary stoma [5]. In addition, indications for recalibration were found in patients with colon cancer [6] and a definitive colostoma [7]. However, the latter two studies used the Then Test to evaluate response shift [6, 7]. Unfortunately, this test is not without criticism due to its proven susceptibility to recall bias [9]. Therefore, more insight in the occurrence of response shift in patients with colorectal cancer is needed.

Furthermore, it is interesting to evaluate whether response shift occurs in partners of patients with colorectal cancer. These partners also have to incorporate ongoing cancer-related experiences into their lives, which may lead to changes in their QoL standards, values, and conceptualization. To our knowledge, this has not yet been evaluated. Incorporating both patients and partners in analyses can not only show if the QoL measurement is invariant over time for patients and partners separately, but also if measurement of QoL is invariant between both members of the couple. If this is true, then QoL findings of both patients and partners can be compared over time and between both dyad members.

Therefore, the aim of the current study was to examine (1) measurement invariance of QoL domains over time for patients with colorectal cancer and their partners (i.e., response shift—recalibration, reprioritization, and reconceptualization),

(2) between dyad-member measurement invariance and (3) QoL trajectories for patients and partners in which the measurement invariance is taken into account.

## Methods

### Patients

Data were drawn from a larger study examining the (sexual) consequences of colorectal cancer for patients and their partners (NCT01234246). Before surgical treatment, patients diagnosed with colorectal cancer and their partners were asked to participate in this study. Patients were recruited from six Dutch hospitals: St. Elisabeth Hospital (Tilburg), TweeSteden Hospital (Tilburg and Waalwijk), Catharina Hospital (Eindhoven), Jeroen Bosch Hospital (‘s Hertogenbosch), Amphia Hospital (Breda), and Maxima Medical Centre (Eindhoven and Veldhoven). Patients and partners older than 18 years were eligible for participation. The following exclusion criteria were applied: (1) elderly age (older than 75 years), (2) noncuratively treated metastases at baseline, (3) poor expression of the Dutch language, (4) dementia, and/or (5) a history of psychiatric illness. In order to prevent selection bias, partners were still invited to participate if the patients declined participation (and vice versa). During a preoperative visit, patients and partners were asked, by their treating physician, whether they gave permission to be approached by a member of the research team. This member would subsequently explain the study design and purpose. If patients and/or their partners agreed to participate, they were asked to complete a set of standardized questionnaires at home preoperatively (Time-0) and 3 (Time-1) and 6 months (Time-2) postoperative. However, the Dutch guidelines [10] recommend that all rectal cancer patients, except those with a clinical T1 stage without positive lymph nodes, receive neoadjuvant treatments [i.e., radio(chemo)therapy]. Therefore, a subset of patients completed the first set of questionnaires before surgery, but potentially while already receiving or after neoadjuvant therapy. Participants returned the surveys in sealed postage-paid envelopes. Participants who consented but who did not return the surveys within 2 weeks after consenting received reminders [phone call(s) or a reminder letter]. The institutional review board approved this study. All participants gave written informed consent.

### Measures

Participants preoperatively completed questions assessing sociodemographic factors (age, sex, marital status, length of the relationship with the partner, and educational level)

and an adapted Self-administered Comorbidity Questionnaire (SCQ) [11], which evaluated comorbidity at the time of questionnaire completion.

The cross-culturally developed World Health Organization Quality of Life assessment instrument-abbreviated version (WHOQOL-Bref) is a generic multidimensional QoL measure [12]. The participants completed this questionnaire at each time point. The WHOQOL-Bref consists of 26 items of which 24 cover four domains: Physical Health, Psychological Health, Social Relationships, and Environment (see Fig. 1). Furthermore, two items constitute the facet overall QoL and General Health. However, since this facet only consists of two items, it was not included in the response shift analyses. The WHOQOL-Bref provides a valid and reliable alternative for the WHOQOL-100 [12]. Higher scores on a domain or the general facet indicate better QoL. Psychometric properties, such as discriminant validity, content validity, internal

consistency, and test–retest reliability of the WHOQOL-Bref, have been examined on multiple occasions and appeared to be good [13–15].

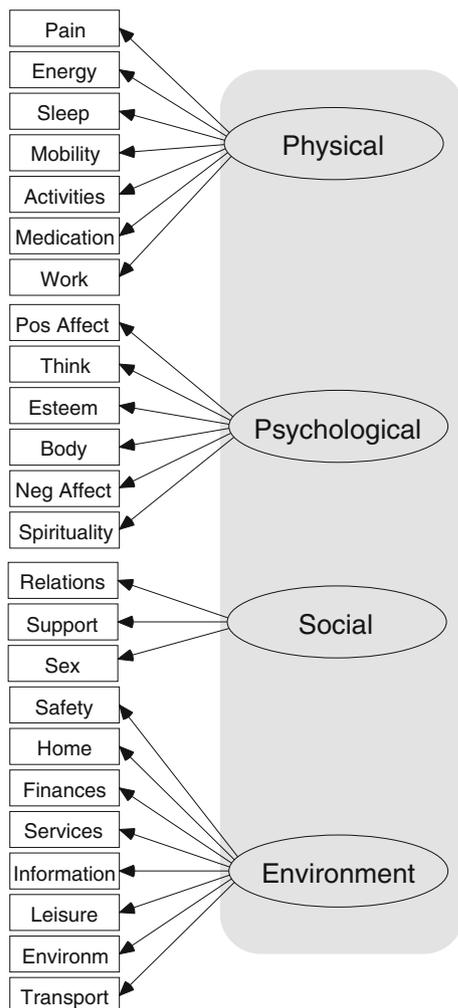
The patient’s clinical information was retrieved from the Eindhoven Cancer Registry (ECR). The ECR routinely collects data on tumor characteristics including date of diagnosis, tumor grade according to the tumor-node-metastasis clinical classification, clinical stage, and treatment. If the ECR had not yet registered all clinical information, then the missing clinical information was retrieved from the patient’s medical records.

### Statistical analysis

Chi-square tests and independent Student’s *t* tests were used to compare patients who declined participation with the participants on age, sex, and tumor type (if applicable). For the participants, descriptive statistics of the sociodemographic and clinical data were calculated. Pairwise difference tests were used to determine whether differences in QoL were noted between colon and rectal cancer for (1) patients and (2) partners.

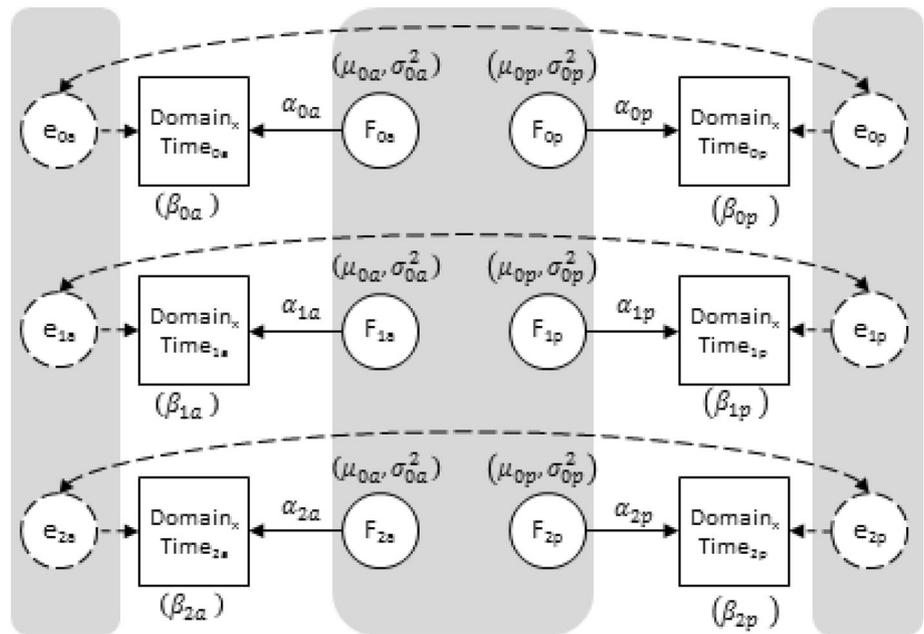
The equivalence of the QoL domains at different time points and between both members of the dyad (patients–partners) was evaluated. This so-called factorial measurement invariance was examined by means of structural equation modeling [16]. An annotated path diagram of the factor model is given in Fig. 2. To account for the dyadic data structure, there are different links (i.e., factor and residual correlations) between the patient- and partner side of the model. For each domain, a stepwise procedure was followed that used nested factor models to examine the viability of restricting specific model parameters to be equal across time points and between dyad members (see Table 1; Fig. 2 and for an overview, see Millsap [17]) [18, 19].

Prior to evaluating the model, it is good practice to verify whether the normality assumption is tenable, for instance, by checking univariate skewness and kurtosis. Skewness is about symmetry around the mean, with negative/positive values indicating that the left/right tail of the distribution is longer than the normal reference distribution. Excess kurtosis is about the size of the tails (i.e., probability of extreme responses), with negative/positive values indicating that the distribution has thinner/fatter tails than the normal reference distribution [20]. In application of statistical models that work under normal-based distributions, it is not necessarily important that the skewness and kurtosis values are ‘significantly’ different from 0, what is more important is the extent to which the model is robust to violations of its assumptions. Based on simulation studies under both structural as well as distributional



**Fig. 1** WHOQOL-Bref structure for patients and partners. Variables within the *gray box* are allowed to intercorrelate

**Fig. 2** Factorial measurement invariance across time and between patients and partners. Circles are latent variables, with the  $F$ 's representing common domain factors and the  $e$ 's unique residuals. Squares are observable variables representing the WHOQOL items. For reasons of clarity only, one indicator per factor is shown, yet it represents the whole set of items for the domain. Numerical subscripts indicate the time point and subscript a/p are for patients (actors) and partners (partners), respectively. Variables within a gray box are allowed to intercorrelate.  $\alpha$  factorloading,  $\beta$  item intercept,  $\mu$  factor mean,  $\sigma^2$  factor variance



**Table 1** Stepwise procedure using nested factor models to examine the viability of restricting specific model parameters to be equal across time points and between patients and partners (see Fig. 2 for the model used)

Invariance	Restrictions
Configural	Means and variances for all factors are fixed, $\mu = 0$ and $\sigma^2 = 1$
Metric	All $\alpha$ s of the same item are equal (i.e., across time points and dyad); variances of all factors are no longer fixed except for the reference factor $F_{0a}$
Scalar	All $\alpha$ s of the same item are equal (i.e., across time points and dyad); all $\beta$ s of the same item are equal (i.e., across time points and dyad); means of all factors are no longer fixed except for the reference factor $F_{0a}$ ; variances of all factors are no longer fixed except for the reference factor $F_{0a}$

$\alpha$  factorloading,  $\beta$  item intercept,  $\mu$  factor mean,  $\sigma^2$  factor variance

assumption violations, Curran, West, and Finch (1996) [21] state that obtained absolute values approaching at least 2 and 7 for skewness and kurtosis are suspect.

In the first step of model evaluation, a baseline model was established to check whether the theoretical domain-specific factor structure holds across time points and between both dyad members (i.e., configural invariance). If this step does not hold, the same unidimensional factor structure, for the specific domain, is not applicable across time points and between patients and partners. This implies a form of reconceptualization of the specific QoL domain. In the second step, metric invariance (i.e., equal factor loadings) was assessed to verify whether the strength of relationship between each item and the underlying domain

construct remains the same across time points and dyad members. If this step does not hold, a form of reprioritization is implied, with for instance one or more items becoming more relevant for the construct domain at a given time point. In the third step, scalar invariance (i.e., equal loadings and intercepts) was assessed to verify whether specific items show a differential score pattern compared to the general trend across time and between dyad members. If this step does not hold, it implies that certain items get recalibrated across time points and/or are being scored higher/lower by patients or partners. In order to evaluate which items are the likely candidates for a specific response shift phenomenon, specific items that do not fit the common general pattern of results across time points and between dyad members (i.e., appear inconsistent with the imposed invariance restrictions) need to be located.

To evaluate a configural invariance baseline model, we used the Chi-square test of absolute fit; (2) the root-mean-square error of approximation to assess close fit (RMSEA,  $<0.08 =$  acceptable,  $<0.05 =$  good fit), and (3) the Comparative Fit Index to assess deviations from the null model (CFI  $0.90\text{--}0.95 =$  acceptable,  $\geq 0.95 =$  good fit). To allow for a clear interpretation of the CFI, we also fitted the null model (i.e., a model that restricts all correlations between items to be zero). If this null model already shows a reasonable goodness of fit (as shown in a relatively small RMSEA), this implies that the average correlation in the data is fairly low and that it will be unlikely to see very large CFI values for other models (as it is hard to greatly improve upon the null model). For the nested model comparisons between the different steps, we relied on Chi-square

difference tests and differences in the comparative fit index ( $\Delta$ CFI: a nonreduced CFI or a reduction smaller than or equal to 0.01 indicates that the null hypothesis of invariance should not be rejected) [22], as well as relative model comparisons based on the sample-size-adjusted Bayesian Information Criterion (saBIC; a lower value indicates a better trade-off between fit and complexity) [23]. When invariance restrictions did not hold, modification indices were used to locate the items showing potential response shift and (if possible) partial invariance was established by relaxing the invariance restrictions for these items [24]. All data analyses were done in the open-source statistical software R using the lavaan package for structural equation modeling [25].

## Results

### Descriptive statistics

In total, 672 eligible patients agreed to be contacted by a member of the research group, who informed them about the study. In total, 313 (47 %) patients agreed to participate. Of these patients, 279 (89 %) had a partner of which 205 (74 %) participated. Due to the dyadic nature of the analyses, only these couples were included ( $n = 205$ ). An overview of the number of participants at each time point is presented in Table 2. Sociodemographic and clinical characteristics of patients and partners (if applicable) are presented in Table 3. No differences were seen in the mean QoL domains scores between colon and rectal cancer patients/partners ( $p > 0.05$ ).

To make use of all available data and account for missing data, the CFA approach was carried out using full-information maximum likelihood under the missing at random assumption. The distributions of the WHOQOL-Bref items remained within the robust range: For patients, skewness of the WHOQOL-Bref items ranged between  $-1.35$  and  $0.21$ , while excess Kurtosis ranged from  $-1.15$  to  $3.51$ . For partners, these ranges were  $-1.42$  to  $0.28$  and  $-1.30$  to  $6.01$ , respectively.

### Equivalence of the QoL domains at different time points and between patients and partners

The null model showed a reasonable goodness of fit as can be seen from the relatively small RMSEA (see Table 4). The nested model comparisons showed that for each of the four QoL domains no support for reconceptualization and reprioritization was found. Hence, the QoL domains demonstrated configural and metric invariance as can be seen from the decreasing saBIC across subsequent models, negligible  $\Delta$ CFI differences, and RMSEA's of  $\leq 0.05$  (Table 4).

Even though absolute fit indices for scalar invariance were still acceptable, relative comparisons pointed at some minor noninvariance issues at the item intercept level for the QoL domains (i.e., recalibration). Thus, for some items, a higher or lower score at specific time points was reported than would be expected based upon the general trend at the domain level. For the Physical Health domain, item 10 'Do you have enough energy for everyday life' had a slightly higher intercept at Time-0 ( $\Delta = 0.28$ ) while item 3 'To what extent do you feel that physical pain prevents you from doing what you need to do?' and item 4 'How much do you need any medical treatment to function in your daily life?' had slightly lower intercepts at Time-0 ( $\Delta = -0.15, -0.48$ , respectively). Thus, for item 10, surgery has a higher impact compared to the general trend. Surgery also had a substantial impact on item 3 and 4 resulting in lower preoperative scores on these items compared with the general trend. For partners, item 15 'How well are you able to get around' had a slightly lower intercept at Time-0 compared with the patients ( $\Delta = -0.19$ ), which indicates that the difference between patients and partners on this item was smaller before surgery. For the Psychological Health domain, item 6 'To what extent do you feel your life to be meaningful?' had a slightly higher intercept at Time-1 ( $\Delta = 0.12$ ), compared to Time-0 and Time-2 for the patients. Thus, at the first assessment after surgery, patients experienced a general decrease on the Psychological Health domain. However, patients showed less decrease on the extent to which they felt their life to be meaningful. For

**Table 2** Number of participants at each assessment point

	Time-0		Time-1		Time-2	
	Patients	Partners	Patients	Partners	Patients	Partners
Completed	164	166	148	141	123	114
Participated from Time-1 onwards	42	40				
Missing			10	10	14	14
Dropout at Time-1			23	28	23	28
Dropout at Time-2					35	35
Not yet completed the questionnaire <sup>a</sup>			25	27	35	35

<sup>a</sup> Due to the ongoing nature of the study, not all participants had already completed all time points

**Table 3** Demographic and clinical characteristics of patients and their partners

	Patients ( <i>n</i> = 205)	Partners ( <i>n</i> = 205)
Age at time of survey (mean ± SD)	62.0 ± 8.6	62.6 ± 9.9
Relationship in years (mean ± SD)	35.9 ± 16.3	35.9 ± 16.3
	<i>n</i> (%)	<i>n</i> (%)
Male gender	146 (71.2)	59 (28.4)
Educational level		
Low	13 (6)	10 (5)
Medium	125 (61)	142 (68)
High	56 (27)	47 (23)
Missing	11 (6)	9 (4)
Type of cancer		
Colon	125 (61)	–
Rectum	80 (39)	–
Type of surgery		
Low anterior resection	84 (41)	–
Abdominoperineal resection	35 (17)	–
Hemicolectomy/ileocecal resection	29 (14)	–
Sigmoid resection	27 (13)	–
Other	27 (13)	–
Unknown type of surgery	3 (2)	–
Type of radiotherapy (RT)		
No RT	135 (66)	–
Preoperative RT	42 (21)	–
Intra-operative RT	19 (9)	–
Other RT	5 (3)	–
Unknown	4 (1)	–
Chemoradiation		
No	125 (61)	–
Yes	74 (36)	–
Unknown	5 (3)	–
Type of chemotherapy (CT)		
No CT	145 (71)	–
Postoperative CT	50 (24)	–
Other CT	7 (3)	–
Unknown	3 (2)	–
Stoma		
Yes	122 (60)	–
No	79 (38)	–
Unknown	4 (2)	–
Comorbidity		
Cardiovascular disease	21 (11)	17 (9)
Lung disease	8 (4)	19 (10)
Diabetes mellitus	11 (6)	17 (9)
Arthrose	27 (14)	44 (24)

**Table 3** continued

	<i>n</i> (%)	<i>n</i> (%)
Back pain	28 (15)	37 (20)

Low = up to 10 years of education, middle = 10–14 years of education, high = more than 14 years of education

patients, item 7 ‘How well are you able to concentrate?’ had a slightly higher intercept at baseline ( $\Delta = 0.21$ ). Thus, surgery and adjuvant treatment (if applicable) have substantial impact on the patients’ concentration. For the Social Relationships domain, both patients and partners had a slightly higher intercept at Time-0 for item 20 ‘How satisfied are you with your personal relationships?’ compared with Time-1 and Time-2 ( $\Delta = 0.20, 0.29$ , respectively). This implies that, after surgery, the decrease in satisfaction with personal relations is even stronger than the general decreasing trend in the Social Relationships domain. Finally, the Environmental domain item 9 ‘How healthy is your physical environment?’ had slightly lower intercept for partners than for patients ( $\Delta = -0.43$ ), which might be linked to the intrinsically worse physical condition of patients. The occurrence of recalibration indicates that the comparison of individual QoL domain scores may be biased. In total, 8 of the 24 items were affected to some extent by recalibration, yet the intercept deviations were relatively small and the majority of items for each separate domain could still be considered scalar invariant, hereby providing a broad enough base to warrant comparison of the domains at the factor level [23]. Therefore, the results of the QoL domains over time and between dyad members, based on group-level statistics of the partial scalar invariant factor model, are described below.

#### QoL domain trajectories

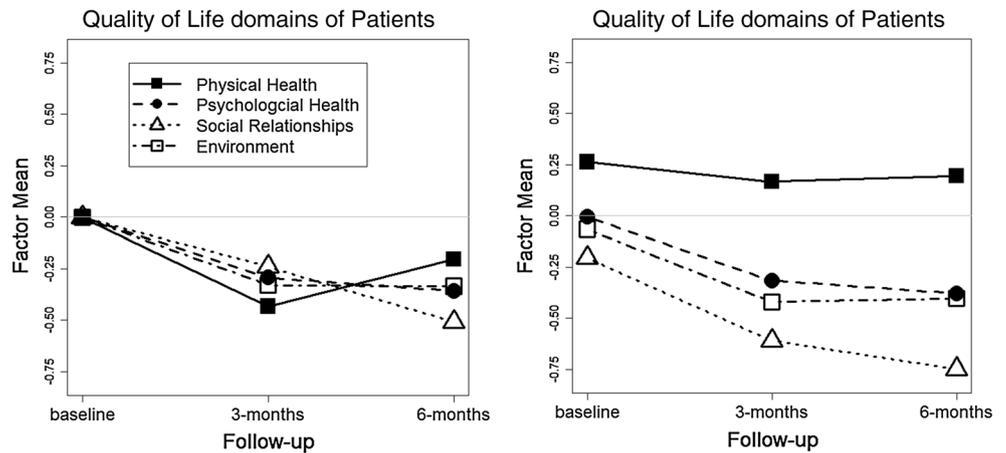
The established partial scalar measurement invariance for the factor structure of the QoL domains allowed comparison of patients and partner over time and between both members of the dyad on these domains at the factor level. For patients and partners, the results of these analyses are presented in Fig. 3a, b. For patients, an initial decrease in the Physical Health domain occurred at Time-1 compared with Time-0 ( $p < 0.001$ ), after which partial recovery to baseline was observed at Time-2 ( $p = 0.055$ ). For partners, factor means in the Physical Health domain remained constant over time ( $p = 0.189, p = 0.679$ ) and were at each time point higher than the patients’ factor means ( $p = 0.007, p < 0.001$ , and  $p = 0.001$ , respectively). Patients and partners’ factor means on the Psychological Health domain significantly decreased at Time-1 ( $p = 0.002$  and  $p < 0.001$ ) after which

**Table 4** Factorial invariance for patient–partner dyads and across time points and between dyad members

Invariance model	$\chi^2$	<i>df</i>	<i>p</i>	CFI	RMSEA	SaBIC	$\Delta\chi^2$	$\Delta df$	<i>p</i>	$\Delta CFI$
<b>Physical Health</b>										
Null	4,186	931	<0.001	0.000	0.130	15,690				
Configural	1,034	729	<0.001	0.900	0.045	12,974				
Metric	1,072	759	<0.001	0.897	0.045	12,947	37.67	30	0.1583	0.003
Scalar	1,160	789	<0.001	0.878	0.048	12,971	88.45	30	<0.001	0.019
Scalar <sup>a</sup>	1,115	885	<0.001	0.892	0.045	12,934	42.908	26	0.0198	0.005
<b>Psychological Health</b>										
Null	3,144	690	<0.001	0.000	0.131	11,716				
Configural	793	510	<0.001	0.880	0.052	9,754				
Metric	831	535	<0.001	0.874	0.052	9,737	37.35	25	0.0534	0.006
Scalar	882	560	<0.001	0.863	0.053	9,735	51.74	25	0.0014	0.011
Scalar <sup>a</sup>	865	558	<0.001	0.870	0.052	9,722	34.23	23	0.0618	0.004
<b>Social Relationships</b>										
Null	968	183	<0.001	0.000	0.144	5,971				
Configural	140	93	0.001	0.933	0.050	5,337				
Metric	154	103	0.001	0.927	0.049	5,330	13.96	10	0.1748	0.006
Scalar	177	113	<0.001	0.909	0.053	5,331	23.00	10	0.0108	0.018
Scalar <sup>a</sup>	164	111	0.001	0.924	0.048	5,323	1.4	8	0.238	0.003
<b>Environment</b>										
Null	3,712	931	<0.001	0.000	0.120	13,965				
Configural	1,123	705	<0.001	0.843	0.054	11,863				
Metric	1,160	735	<0.001	0.840	0.053	11,836	37.46	30	0.1642	0.003
Scalar	1,235	765	<0.001	0.824	0.055	11,845	73.97	30	<0.001	0.016
Scalar <sup>a</sup>	1,185	760	<0.001	0.841	0.052	11,806	24.169	25	0.50960	0.001

<sup>a</sup> Scalar indicates partial invariance: freeing the intercepts of recalibrated items

**Fig. 3** Factor means across time for the QoL domains for patients and partners. The factor mean at baseline for patients is set as reference (i.e., *dashed line*) and fixed at zero to identify the longitudinal Structural Equation Model



stabilization occurred at Time-2 ( $p = 0.471$ , and  $p = 0.362$ ). Patients and partners’ factor means were at each time comparable ( $p = 0.844$ ,  $p = 0.835$ , and  $p = 0.838$ , respectively). Patients and partners’ factor means on the Social Relationships domain significantly decreased at Time-1 ( $p = 0.018$  and  $p < 0.001$ ), which decreased further on for patients ( $p = 0.011$ ), but stabilized for partners ( $p = 0.214$ ). Partners factor means were significantly lower than patients

factor means at Time-1, but not at the other time points ( $p = 0.088$ ,  $p < 0.001$ ,  $p = 0.098$ , respectively). For the Environmental domain, a decrease in factor means was noted for both patients and partners at Time-1 ( $p < 0.001$  and  $p < 0.001$ ), after which stabilization occurred at Time-2 ( $p = 0.952$  and  $p = 0.830$ ). For this domain, patients and partners’ factor means were at each time comparable ( $p = 0.459$ ,  $p = 0.346$ , and  $p = 0.484$ , respectively).

For each domain, significant correlations in the QoL construct domains were found across time points, for both groups (see Table 5). These relatively high overtime correlations indicate that the relative position of patients and partners stayed stable over time; however, their actual QoL scores might still have changed. Between patients and partners, the QoL domain constructs correlated significantly at each time point (see Table 5). For the Social Relationship and Environmental domains between dyad-

member correlations were relatively homogenous. However, for the Physical and Psychological Health domains, an auto-regressive correlation structure was noted: Correlations reduced over time. In addition, at each time point, the correlations between patients and partners for the Physical Health domain were low. Finally, the partners' lower variances for the Physical Health and Psychological Health domains might indicate that, for these domains, the partners are a more homogeneous group than the patients.

**Table 5** Correlations at different time points and between patients and partners and variances for patients and partners at each time point for the four QoL domain factors

	Patient			Partner		
	Time-0	Time-1	Time-2	Time-0	Time-1	Time-2
<i>Physical Health</i>						
Patients						
Time-0	1.000					
Time-2	0.496*	0.940				
Time-3	0.420*	0.692*	1.052			
Partners						
Time-0	0.281**	0.141**	0.152***	0.882		
Time-2	0.366*	0.123**	0.321**	0.681*	0.702	
Time-3	0.111	-0.037	0.100	0.398*	0.706*	0.692
<i>Psychological Health</i>						
Patients						
Time-0	1.000					
Time-2	0.729*	1.252				
Time-3	0.668*	0.721*	0.939			
Partners						
Time-0	0.416*	0.282**	0.269***	0.848		
Time-2	0.329**	0.340**	0.309**	0.269***	0.888	
Time-3	0.093	0.181	0.215	0.309**	0.760*	0.715
<i>Social Relationships</i>						
Patients						
Time-0	1.000					
Time-2	0.863*	0.847				
Time-3	0.890*	0.736*	1.059			
Partners						
Time-0	0.462*	0.435**	0.241	1.310		
Time-2	0.493**	0.410***	0.426**	0.681*	0.816	
Time-3	0.507**	0.353***	0.300***	0.398*	0.808*	1.077
<i>Environment</i>						
Patients						
Time-0	1.000					
Time-2	0.765*	1.074				
Time-3	0.856*	0.738*	0.861			
Partners						
Time-0	0.496*	0.487*	0.501*	0.973		
Time-2	0.524*	0.596*	0.592*	0.789*	1.041	
Time-3	0.336**	0.376**	0.548*	0.661*	0.850*	0.998

Variances on the diagonal and correlations on the off-diagonal  
 \*  $p < 0.0001$ ; \*\*  $p < 0.01$ ;  
 \*\*\*  $p < 0.05$

## Discussion

For both patients with colorectal cancer and their partners, no indication for reconceptualization (i.e., no dissimilarity in the conceptual frame of the construct domains across time points) and reprioritization (i.e., similar relative importance of one or more of the items relevant for the construct domain at a given time) was found. Furthermore, conceptualization and prioritization was similar between patients and partners.

The lack of indications for reconceptualization and reprioritization might be surprising, since earlier studies did find these indications in other populations [19, 26–31]. The questionnaire used to assess response shift may in part explain the outcomes [32]. Questionnaires routinely used to assess response shift are in fact questionnaires that assess health status [e.g., The Short Form (12/36) Health Survey; SF-12/36] [3]. The terms QoL, health status, and health-related QoL are often used interchangeably; however, the distinction between these definitions is important [33]. Health status evaluates the impact of disease on functioning [33]. Questionnaires evaluating health status, such as the SF-12/36, therefore assess physical possibilities, social activities, and psychological function, but not the participants' feelings concerning this functioning [33]. QoL is defined by the World Health Organization Quality of Life Group as 'an individual's perception of his/her position in life in the context of the culture and value systems in which he/she lives and in relation to his/her goals, expectations, standards, and concerns' [34]. This definition implies that QoL is always subjective in nature. Therefore, QoL entails participants' subjective interpretation of their well-being. Thus, health status questionnaires measure the impact of disease on functioning, whereas QoL questionnaires measure the extent to which a participant is bothered by these limitations [32]. Therefore, questionnaires using the SF12/36 might assess response shift in health status as opposed to response shift in QoL. In addition, our study only examined within-domain reconceptualization. However, in order to detect reconceptualization between the several domains, it is also necessary to evaluate the cross-loadings. Unfortunately, we were unable to conduct these multidimensional analyses due to a limited sample size. In addition, in both research and clinical practice, the WHOQOL-Bref is only used to evaluate each domain individually (i.e., no total score is computed); therefore, comparing the domains does also not fit within the established fixed framework. Thus, in this study, the current within-domain analyses were justified, but future studies in this area are still needed. In line with earlier studies [6, 7], indications for recalibration were found. That is, some items showed a differential score pattern compared to the general trend over time and between dyad members. It is not surprising

that patients preoperatively reported to have more energy (item 10), a better ability to concentrate (item 7), less prevented in their activities by physical pain (item 3), and in less need of medical treatment (item 4) compared with the general domain trend, especially since surgery and adjuvant treatments (if applicable) can have a substantial impact on these aspects. It is also not surprising that the preoperative difference between patients and partners was smaller on their ability to get around (item 15). However, it is noteworthy that the decrease in satisfaction with personal relations is stronger than the general decreasing trend in the Social Relationships domain. Potentially, the social support received after treatment did not live up to patients and partners' expectations. In addition, even though patients experienced a general decrease on the Psychological Health domain at Time-1, patients showed less decrease on the extent to which they felt their life to be meaningful. In our opinion, this reflects what is in clinical practice often meant with response shift. Finally, for the Environmental domain, partners reported their physical environment to be less healthy, which might be linked to the intrinsically worse physical condition of the patients.

While the indications for recalibration prohibited straightforward comparison of individual QoL domain scores, the established (partial) scalar measurement invariance did allow comparison of patients and partners' QoL domain factor means over time and between both members of the dyad. For both patients and partners, these factor level analyses showed, for the domain Psychological Health and the Environmental domain (consisting of questions such as 'To what extent do you have the opportunity for leisure activities?'), an initial significant decrease at Time-1 compared with Time-0, after which stabilization was observed at Time-2. This makes sense, since patients and partners both need to recover psychologically from the diagnosis and impact of cancer treatment. Patients and partners' factor means on the Social Relationships domain also decreased at Time-1. However, at Time-2, this dissatisfaction increased for patients, while a partial recovery occurred for partners. As can be expected, partners had higher scores on the Physical Health domain compared with patients and correlations between patients and partners for this domain were low. All in all, these results show that both patients and partners are affected by the patients' illness. Therefore, it is recommended that healthcare professionals and researchers pay attention to the QoL of both patients and partners.

To our knowledge, this is the first study that examined measurement invariance and QoL over the first 6 months after colorectal cancer surgery from a dyadic perspective. The results of the factor level QoL domain analyses acknowledge the notion that a coupe-based perspective to cope with colorectal cancer is warranted. Using the

WHOQOL-Bref was an important advantage, as it assesses the subjective evaluation of functioning instead of functioning alone. However, there are also some limitations that need to be noted. For instance, although the first questionnaire was completed prior to surgery (Time-0), it probably does not reflect the patients and partners' actual baseline QoL. The time after diagnosis but prior to surgery can be very distressing for patients, as the consequences and results of cancer treatment are still unknown. Furthermore, the patient population was quite heterogeneous. Within and between patients with colon and rectal cancer, a wide variety of treatment schedules and possible consequences of treatment (e.g., a stoma) can occur. This might have influenced our results, even though no differences in QoL were found between patients with colon and rectal cancer. In addition, future studies are needed to evaluate multidimensional reconceptualization and how the QoL trajectories develop in the future. Finally, even though patients' self-report is the most common and straightforward way to obtain QoL information, questionnaires may not completely capture the response shift experienced. In this regard, Rapkin and Schwartz [32] rightly advise to incorporate direct measures of QoL appraisal and interviews in order to account for inter-individual and temporal differences in the meaning attributed to QOL scales.

## Conclusion

For patients with colorectal cancer and their partners, no reconceptualization and reprioritization were detected within each of the QoL domains, but indications for recalibration of specific items were found. The reported partial scalar factorial invariance implied that straightforward comparison of individual QoL domain scores was not warranted, but that comparisons over time and between patients and partners were feasible if restricted to group-level statistics at the factor level. The QoL domain trajectories showed that both patients and partners are affected by the patient's disease and treatment. Therefore, it is recommended that healthcare professionals and researchers give attention to the couple as opposed to the patient alone.

## References

- Sprangers, M. A., & Schwartz, C. E. (1999). Integrating response shift into health-related quality of life research: A theoretical model. *Social Science and Medicine*, *48*(11), 1507–1515.
- Schwartz, C. E., & Sprangers, M. A. (1999). Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. *Social Science and Medicine*, *48*(11), 1531–1548.
- Schwartz, C. E., Bode, R., Repucci, N., Becker, J., Sprangers, M. A., & Fayers, P. M. (2006). The clinical significance of adaptation to changing health: A meta-analysis of response shift. *Quality of Life Research*, *15*(9), 1533–1550.
- Sprangers, M. A. (2010). Disregarding clinical trial-based patient-reported outcomes is unwarranted: Five advances to substantiate the scientific stringency of quality-of-life measurement. *Acta Oncologica*, *49*(2), 155–163.
- Neuman, H. B., Park, J., Fuzesi, S., & Temple, L. K. (2012). Rectal cancer patients' quality of life with a temporary stoma: Shifting perspectives. *Diseases of the Colon and Rectum*, *55*(11), 1117–1124.
- Bernhard, J., Hurny, C., Maibach, R., Herrmann, R., & Laffer, U. (1999). Quality of life as subjective experience: Reframing of perception in patients with colon cancer undergoing radical resection with or without adjuvant chemotherapy. Swiss group for clinical cancer research (sakk). *Annals of Oncology*, *10*(7), 775–782.
- Ito, N., Ishiguro, M., Tanaka, M., Tokunaga, K., Sugihara, K., & Kazuma, K. (2000). Response shift in quality-of-life assessment in patients undergoing curative surgery with permanent colostoma. *Gastroenterology Nursing*, *33*(6), 408–412.
- [www.globocan.iarc.fr](http://www.globocan.iarc.fr).
- Schwartz, C. E., Sprangers, M. A. G., Carey, A., & Reed, G. (2004). Exploring response shift in longitudinal data. *Psychology and Health*, *19*(1), 51–69.
- <http://www.oncoline.nl/rectalcancer>.
- Sangha, O., Stucki, G., Liang, M. H., Fossel, A. H., & Katz, J. N. (2003). The self-administered comorbidity questionnaire: A new method to assess comorbidity for clinical and health services research. *Arthritis Rheum*, *49*(2), 156–163.
- The WHOQOL group. (1998). Development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychological medicine*, *28*(3), 551–558.
- Skevington, S. M., Lotfy, M., & O'Connell, K. A. (2004). The World Health Organization's WHOQOL-BREF quality of life assessment: Psychometric properties and results of the international field trial. A report from the WHOQOL group. *Quality of Life Research*, *13*(2), 299–310.
- Trompenaars, F. J., Masthoff, E. D., Van Heck, G. L., Hodiadmont, P. P., & De Vries, J. (2005). Content validity, construct validity, and reliability of the WHOQOL-BREF in a population of Dutch adult psychiatric outpatients. *Quality of Life Research*, *14*(1), 151–160.
- van Esch, L., Den Oudsten, B. L., & De Vries, J. (2011). The World Health Organization quality of life instrument-short form (WHOQOL-BREF) in women with breast problems. *International Journal of Clinical and Health Psychology*, *11*(1), 5–22.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. New York: Routledge.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*(1), 4–70.
- Schwartz, C. E., Merriman, M. P., Reed, G. W., & Hammes, B. J. (2004). Measuring patient treatment preferences in end-of-life care research: Applications for advance care planning interventions and response shift research. *Journal of Palliative Medicine*, *7*(2), 233–245.
- Westfall, P. H. (2014). Kurtosis as Peakedness, 1905–2014. R.I.P. *The American Statistician*, *68*(3), 191–195.
- Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, *1*(1), 16–29.

22. Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233–255.
23. Sclove, L. S. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika*, 52, 333–343.
24. Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, 105(3), 456–466.
25. Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36.
26. Oort, F. J., Visser, M. R., & Sprangers, M. A. (2005). An application of structural equation modeling to detect response shifts and true change in quality of life data from cancer patients undergoing invasive surgery. *Quality of Life Research*, 14(3), 599–609.
27. Chan, C. S., Rhodes, J. E., Howard, W. J., Lowe, S. R., Schwartz, S. E., & Herrera, C. (2013). Pathways of influence in school-based mentoring: The mediating role of parent and teacher relationships. *Journal of School Psychology*, 51(1), 129–142.
28. Schwartz, C. E., Feinberg, R. G., Jilinskaia, E., & Applegate, J. C. (1999). An evaluation of a psychosocial intervention for survivors of childhood cancer: Paradoxical effects of response shift over time. *Psychooncology*, 8(4), 344–354.
29. Schwartz, C. E., Wheeler, H. B., Hammes, B., Basque, N., Edmunds, J., Reed, G., et al. (2002). Group UME-o-LW. Early intervention in planning end-of-life care with ambulatory geriatric patients: Results of a pilot trial. *Archives of Internal Medicine*, 162(14), 1611–1618.
30. Schwartz, C. E., Sajobi, T. T., Lix, L. M., Quaranto, B. R., & Finkelstein, J. A. (2013). Changing values, changing outcomes: The influence of reprioritization response shift on outcome assessment after spine surgery. *Quality of Life Research*, 22(9), 2255–2264.
31. Barclay-Goddard, R., Lix, L. M., Tate, R., Weinberg, L., & Mayo, N. E. (2011). Health-related quality of life after stroke: Does response shift occur in self-perceived physical function? *Archives of Physical Medicine and Rehabilitation*, 92(11), 1762–1769.
32. Rapkin, B. D., & Schwartz, C. E. (2004). Toward a theoretical model of quality-of-life appraisal: Implications of findings from studies of response shift. *Health and Quality of Life Outcomes*, 2, 14. doi:10.1186/1477-7525-2-14
33. Van der Steeg, A. F., De Vries, J., & Roukema, J. A. (2004). Quality of life and health status in breast carcinoma. *European Journal of Surgical Oncology*, 30(10), 1051–1057.
34. (1998). The World Health Organization quality of life assessment (WHOQOL): Development and general psychometric properties. *Social Science Medicine*, 46(12), 1569–1585.