University of Bath

1

EXPLORING RESPONSE SHIFT IN THE QUALITY OF LIFE OF HEALTHY ADOLESCENTS OVER ONE YEAR.

Fiona Gillison, BSc, MSc, PhD[a]

Suzanne Skevington, BSc, PhD[b]

Martyn Standage, BSc, MSc, PhD[a]

Corresponding author: Fiona Gillison

[a] School for Health

University of Bath

Bath, BA2 7AY

United Kingdom

Tel: 0044 1225 386696

Fax: 0044 1225 383275

f.b.gillison@bath.ac.uk

m.standage@bath.ac.uk

[b] WHO Centre for the Study of Quality of Life and Department of Psychology

University of Bath

Bath, BA2 7AY

United Kingdom

s.m.skevington@bath.ac.uk

Abstract

Background: The assessment of child and adolescent quality of life (QoL) is increasingly incorporated into practice and research within health and social care. However, little is known about the stability of the conceptualisation of QoL in the face of normative developmental processes against which to compare changes in response to non-normative events, such as chronic illness. The present study examined the stability of the QoL construct over one year in a cohort of healthy community dwelling adolescents through applying the response shift (RS) model.

Method: A cohort of 356 adolescents (*M* age =14.05) completed the Kidscreen QoL measure on two occasions, one year apart. Both research-design and statistical techniques were employed to assess three types of RS; recalibration, reprioritisation and reconceptualisation.

Results: A significant recalibration of small effect size was found for three QoL domains using a then-test, which was confirmed in the domain of autonomy through statistical modelling. No evidence was found for reprioritisation or reconceptualisation RS using either analytical approach.

Conclusion: The findings support the assumption that the conceptualisation of QoL is stable during adolescence in the absence of significant health threats. However, researchers and clinicians should be aware of potential recalibration effects of some life domains.

Introduction

In recent years the assessment of children and adolescents' quality of life (QoL) has been increasingly incorporated into practice and research within both health and social care settings [1,2]. The construct has been included as an outcome measure to evaluate and compare clinical outcomes, such as treatment for diabetes, asthma and cystic fibrosis [3], and within social services QoL has been incorporated into the decision making processes relating to taking children into care [4,5]. Indeed, within the UK, a brief QoL measure is now included as standard in social work child case assessments [5]. If QoL information is to be relied upon as a basis for clinical or practical judgements, it is particularly important that we are confident that our measures are providing accurate and meaningful results. The study of children and adolescents' QoL has lagged behind that of adults, and although numerous well-validated measures have now been developed specifically for these age groups (e.g., the Pediatric Quality of Life Inventory (PedsQL), [6] and Kindl [7]), many of the assumptions behind methods of assessment (e.g., stability and comprehension) have not been directly tested with younger populations. In particular, it is assumed that in the absence of significant life-events, the conceptualisation and internal standards against which a person measures their own QoL should remain stable. The present study argues that given the range of normative but often challenging developmental tasks that young people face during adolescence, the foundations on which QoL is based, i.e., its conceptualisation and measurement, may not remain stable. As such, this paper represents an original research direction to explicitly test whether the assumption of the conceptual stability of QoL holds in an adolescent population.

The World Health Organisation (WHO) defines QoL as "an individual's perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations values and concerns" [8, p1403]. A person's perception of their QoL therefore stems from their evaluation of how their life is going relative to the life domains they consider to be relevant and important at that time. This subjective definition

infers that QoL may change in quantitative terms (e.g., life is not as good as this time last year) but also in qualitative terms as a result of a change in the domains of life considered most important. The most common application of QoL data in children and adolescents, as with adults, is in assessing the impact of illness, disability and of interventions aimed at their management (see [1] for a review). The validity of such work clearly relies on the assumption that without significant health threat, no change in the conceptualisation of QoL would be experienced. However, adolescents experience a series of meaningful and potentially significant changes in their lives which could present considerable challenges to their values and priorities [9]. At different stages of development these may be biological/physical (e.g. reaching puberty), cognitive (e.g., ability to engage in social comparisons), educational (e.g., facing important exams), environmental (e.g., transition to senior school), and social (e.g., developing romantic relationships) [10,11]. In adapting to these normative developmental processes it is possible that the factors which adolescents consider to constitute their QoL, and the relative importance of these domains, may not remain stable.

The theoretical framework of response shift (RS) provides a useful basis by which to investigate qualitative changes in QoL at a time when goals and priorities may be changing [12]. RS has been be defined as "a change in the meaning of one's self-evaluation of a target construct as a result of: (a) a change in the respondent's internal standards of measurement (scale recalibration, in psychometric terms), (b) a change in the respondent's values (i.e., the importance of component domains constituting the target construct), or (c) a redefinition of the target construct (i.e., reconceptualisation)" [12, p.1508]. The study of RS stems from observations within health care settings that despite the continuing functional or social limitations experienced by patients after the onset of chronic disease or disability, many still often reported a level of QoL similar to, or even higher than that prior to an objective deterioration in their health [e.g., 13,14]. It is the processes underlying such

4

apparent adjustments to changes in ones circumstances that are the focus of research into RS.

Three distinct types of RS in QoL have been documented [12]. *Recalibration* relates to a change in an individual's standards, and occurs when an individual adjusts their expectations to enable them to be satisfied with a different objective standard of circumstances. For example, following adjustment to the onset of a disability, a person may reclassify the health status they would consider as "good" to mean "well enough to carry out key activities of daily living unassisted", whereas prior to their disability they same label would only have been be applied to perfect health. *Reprioritisation* reflects a re-ordering of the importance of the different domains within QoL. For example, a person suffering from a chronic illness may come to place greater value on social relationships and de-emphasise their capacity to work. *Reconceptualisation* refers to a more significant change in values, whereby an individual may redefine the factors that constitute their QoL. Thus, rather than changing the relative importance of domains (i.e., reprioritisation), previously important domains are discarded and/or new domains emerge. The RS model assumes that these three mechanisms are underpinned by change along any or all of three axes; behaviour (e.g., coping skills), cognitions (e.g., reframing expectations), and affect (e.g., detachment mechanisms) [12].   Thus, previous RS research has documented how QoL may be protected through processes such as changing social comparison group [15] or life goals [16].

The RS model has received intensive investigation with reference to major health events, particularly the onset of disability and the diagnosis and treatment of serious illness. An example of how RS manifests can be seen in work exploring the relationship between QoL and change in life goals in AIDS patients [16]: Rapkin et al. [16] monitored change in patient QoL alongside changes in life goals at six monthly intervals, finding that despite considerable changes in objective health, there was little change in QoL. However, 78% of

5

the sample reported at least one change in their life goals over the period. Moreover regression analysis demonstrated that changes in personal goals moderated the effect of health status on QoL. Patients had maintained an overall positive QoL through prioritising achievable goals such as personal growth and interpersonal relationships, in favour of less achievable goals of work or physical function. We argue that if RS can be brought about by a change in goals, it is plausible that the changing life goals and priorities reported to take place during adolescence could also induce an RS effect in adolescent QoL.

Taking account of RS in the study of longitudinal QoL is important in ensuring that the same construct is compared on any two occasions, both to monitor adjustment and to inform us of how successful adjustment may be achieved. Investigating whether RS is present in a healthy adolescent population is important for reasons beyond the establishment of a sound baseline for comparison with clinical studies. In studies that do not take account of potential RS, significant age related decreases in QoL have been reported during adolescence [17,18,19]. In turn, poor QoL has been associated with the development of health risk behaviours, such as smoking, and drug and alcohol abuse [20], which all have implications for current and future health, and social and educational opportunities. Therefore, establishing whether the reported decrease in QoL is a real effect, or a result of a change in recalibration, reprioritisation or reconceptualisation of the construct is important in building confidence in the interpretation of such findings, and the subsequent implications for the distribution of resources.

To date, no study could be located that assessed RS in child or adolescent populations through subjective self-report, as opposed to proxy data. The methodological framework for assessing RS in adult clinical populations is therefore, as yet, untested with children and adolescents. A range of methods for assessing RS in adults are available, including qualitative interviews, research-design techniques (e.g., the then-test), and statistical analyses (see [21] for discussion). There is debate as to which is the most accurate

6

technique; for example, it is argued that research-design methods are open to recall bias, whereas statistical methods used alone have been criticised for lacking face validity [21]. Therefore, the present study will employ a range of methodological approaches, both research-design and statistical techniques, to assess each type of RS.

In this work we aimed to assess the extent of RS in self-reported QoL in response to normative developmental experiences in a cohort of healthy adolescents over a period of one year. Assessments were undertaken by students during years 9 and 10 at UK senior school (i.e., between the age of approximately 14 and 15). This age group was chosen as one that would encompass a variety of normative but potentially significant changes throughout the cohort (i.e., biological, social and educational). At this age, the majority of both boys and girls in the sample would have experienced at least some of the physical changes associated with puberty, which it is argued are a trigger for the process of adjustment [22]. All of the three types of response shift were evaluated, and hypotheses for each specified on the basis of the adolescent developmental literature.

Hypothesis 1. In the absence of related research indicating the presence of changes in adolescents' internal standards, it was hypothesised that there would be no significant recalibration effect in this sample.

Hypothesis 2. It was hypothesised that a reprioritisation of the domains within QoL would take place as a result of normative changes in adolescents' goals and priorities. In line with existing knowledge of adolescent development, it was predicted that;
- the domains of peer relationships, bullying, autonomy and personal finances would increase in importance, due to the task of developing a self-identity, and increasing reference to peers as a source of information for self-attributions [23,24];

- the domains of family relationships and school would decrease in importance to reflect a decrease in reliance on adult judgements for self-attributions and self-concept [25];
- the importance of the domains of physical health, self-perceptions, psychological well-being, and mood and emotions would remain constant.

Hypothesis 3. It was hypothesised that there would be no reconceptualisation RS within the present sample. Although change in relative importance of domains was expected in line with Hypothesis 2, in the absence of abrupt significant life events as observed in clinical populations, no drastic change in the constituents of QoL was predicted over this period. Furthermore, the initial validation of the QoL instrument used indicated that all domains were found to be of continuing relevance to children and adolescents from age 8 to 18 [26].

Method

Design

Ethical approval was obtained from the authors' local NHS Research Ethics Committee. The study was conducted as a repeated measures cohort study over a one year period. The sampling frame was a complete list of schools from a single Local Educational Authority. Schools were recruited sequentially from the list until the target sample size was achieved (four schools).

Participants

All pupils within the target year group in participating schools were invited to take part. Written consent was provided by Head Teachers, and subsequently letters were sent home to all parents informing them of the study, and inviting them to contact the school or research team if they preferred their child not to take part (passive consent). Verbal consent was

obtained from students on the day of data collection. Students were informed that the research was not compulsory, and that they could withdraw at any time.

Measures

The Kidscreen self-report QoL questionnaire.

QoL was assessed using the Kidscreen self-report health-related QoL questionnaire [26]. The Kidscreen was developed in consultation with adolescents, parents and carers in 13 European countries. The measure consists of 52 subjective items assessing ten dimensions of QOL; physical well-being, psychological well-being, mood and emotions, self-perceptions, autonomy, family relationships, relationships with peers, school environment, bullying, and financial resources. Participants rate each statement on a 5-point Likert scale, anchored by 1 (*never* or *not at all*) to 5 (*always* or *extremely*). Examples of items include *"Have you been in a good mood"* (mood and emotions domain) and *"Have you been able to choose what to do in your free time?"* (autonomy domain). In line with the authors' recommendations, raw scores are transformed using Rasch modelling prior to analysis, such that higher scores represent better QoL [27]. Internal consistency was shown to be acceptable to good in the authors' work with over 20,000 children (α=.77-.89) [26]. Subsequent work with UK school age children has demonstrated the Kidscreen to be sensitive to change; QoL was tracked over a 10 week period following the transition to senior school, and in line with expectations of swift improvements in adjustment over this period [28], reported an improvement of QoL of small to moderate effect size [29]. National data of mean values by domain for UK adolescents are presented alongside ratings for the present sample in Table 1.

Importance Ratings

At both time points participants rated how important they considered each domain to be in contributing to their overall QoL. Importance items were newly constructed for the present study, drawing on techniques used in constructing importance ratings for the World Health Organization QoL instruments [30,31]. The wording of items was kept consistent with that used in the Kidscreen questionnaire, and responses were recorded on a similar 5-point

Likert scale.

Table 1   Comparison of domain scores (unadjusted for RS) with UK reference values

	Time 1	Time 2	UK comparison scores
Physical health	43.63 (6.42)	42.86 (7.17)	47.04
Family relationships	44.19 (9.93)	43.61 (10.92)	47.74
Self perceptions	43.63 (7.65)	43.34 (7.07)	46.24
Psychological well-being	45.41 (9.24)	44.54 (9.20)	47.15
School	44.06 (8.56)	43.96 (8.55)	46.52
Peers	48.89 (9.10)	48.88 (9.83)	49.96
Financial	47.35 (9.25)	47.26 (9.52)	50.32
Moods & emotions	46.18 (8.31)	46.02 (9.53)	46.84
Bullying [b]	41.93 (11.04)	43.79 (11.13)	46.78
Autonomy [b]	48.52 (9.93)	45.11 (9.08)	48.50

Then-test

At Time 2, participants completed a then-test [32]. This required the completion of the QoL measure twice at the second time point; once with respect to how the student is currently feeling (post-test), and once responding with a retrospective judgement of their QoL at the initial baseline (then-test). By completing the post-test and then-test at the same point in time it is assumed that the items are being judged with the same internal standard of measurement [12]. In line with previous research and in order to reduce the response burden, the then-test consisted of only one item from each Kidscreen domain (that demonstrating the strongest item-scale correlations) [13]. Participants were assisted in thinking back to this time last year through orientation to specific events in the school calendar and reminders of their tutor group. Items were preceded with the instruction "Thinking about this time last year….", and rephrased in the past tense.

Procedure

Participants completed the Kidscreen QoL questionnaire at the end of a battery of measures for a wider study [33]. Due to a shortage of time allowed for data collection in one school, only participants from two schools completed the then-test.

Analysis

Research-design techniques

T-tests were used to calculate the significance of differences between pre-test and then-test scores indicative of recalibration. Each QoL domain was analysed separately. Reprioritisation was assessed through the comparison of the rank order of importance ratings across domains. Reconceptualisation was assessed through testing change in self-reported importance ratings, through examination of 95% confidence intervals (CIs). Specifically, importance ratings were tested for significant difference from zero at both time points (i.e., zero was not within the CI), and these findings compared across measurement occasions.

Statistical techniques

Structural equation modeling (SEM) was used to provide a concurrent statistical test for all three types of RS using the four stage approach recommended by Oort et al. [34]. The first stage establishes a measurement model of the mean and covariance structure of the observed variables, providing good fit to the data at both time points [34,35]. The ten domain scores were used as the observed variables in the model, each loading onto the common single factor of QoL. To achieve model fit, three additional covariance terms were added to the model; allowing the residual factors of the mood and emotion, self-perceptions, and bullying domains to covary[1] (Figure 1). The stage 1 model, for which no constraints of invariance across time points are imposed is termed Model 1.

Stage 2 involves constraining (i) the intercepts of each domain, (ii) the factor loadings of each domain, and (iii) the variance of residual terms to equality across time. This represents

11

Figure 1: Measurement model and factor loadings for response shift analysis

re1 → **Physical health**	(.43; 1.00)
re2 → **Well-being**	(.81; 2.52)
re3 → **Moods & emotion**	(.63; 1.89)
re4 → **Self-perceptions**	(.46; 1.17)
re5 → **Autonomy**	(.62; 2.05)
re6 → **Peers**	(.45; 1.45)
re7 → **Parents**	(.69; 2.42)
re8 → **School**	(.60; 1.73)
re9 → **Bullying**	(.28; 1.07)
re10 → **Finances**	(.55; 1.79)*

Quality of Life
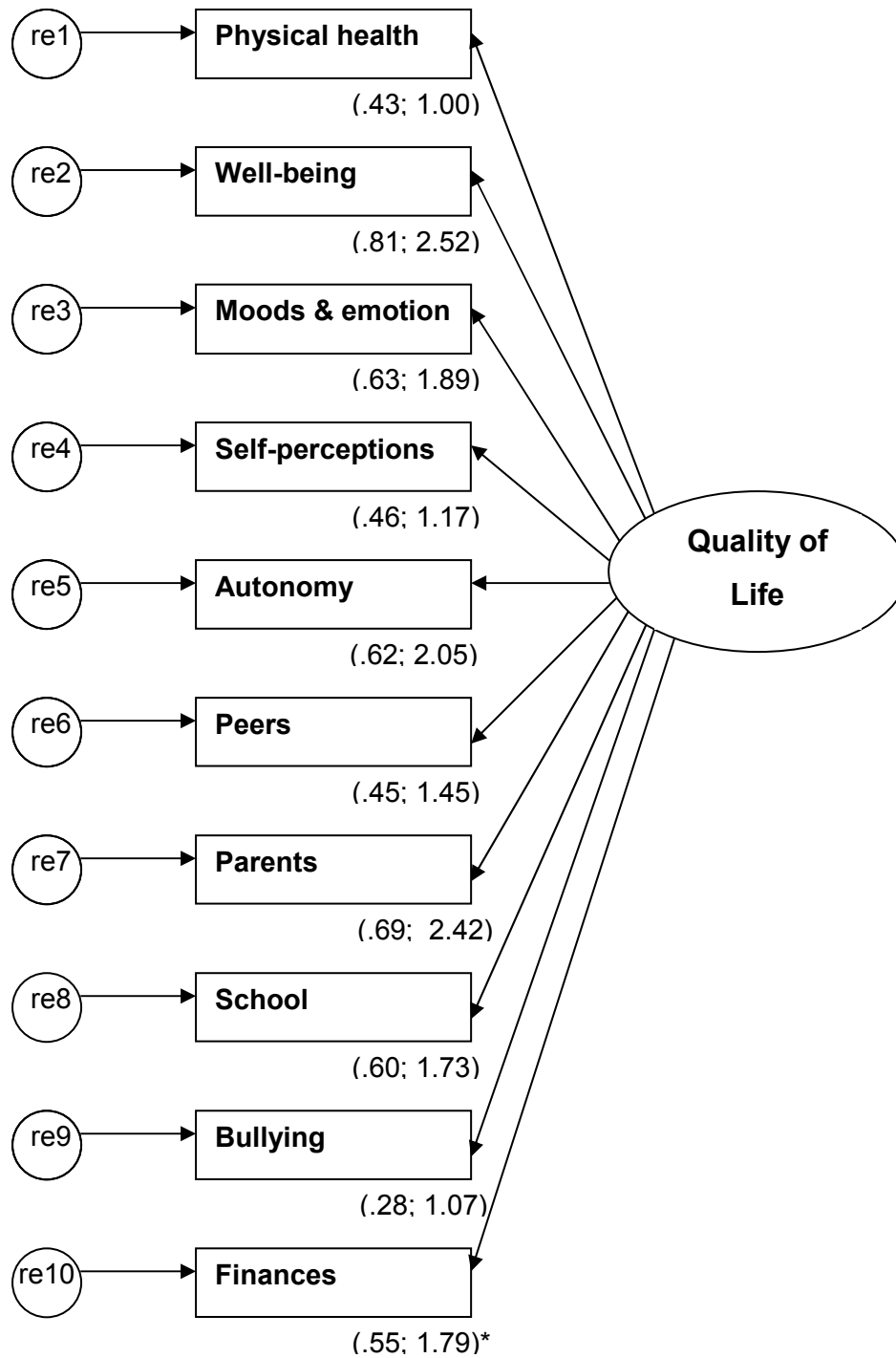
Key: re = residual factor; * figures denote (standardised; unstandardised) factor loadings

a model of no RS (Model 2). If the fit of Model 2 is not significantly poorer than Model 1 then it is interpreted to indicate that no response shift has taken place (i.e., all changes in means and covariances can be attributed to changes in the common factor mean and variance).

If Model 2 has a significantly poorer fit than Model 1, stage 3 involves specifying a third model, Model 3, following the sequential release of untenable constraints in the model until all apparent response shifts are accounted for (i.e., until the fit of Model 3 is not significantly poorer than Model 1). Improved fit through release of constraints between the intercepts of observed variables (i.e., domain scores) is indicative of uniform recalibration RS (i.e., size and direction of effect does not vary with true QoL). Non-uniform recalibration RS (i.e., size or direction of effect does vary with true QoL) is indicated by untenable constraints between the variance of the observed variables. Evidence for reprioritisation RS is detected through either (i) improved fit following the release of constraints on factor loadings of the observed variables (i.e., domain scores), suggesting that a domain has become more or less indicative of the concept which it is intended to measure, and/or (ii) improved fit following the release of constraints in covariance terms between residual covariances, indicating a shift in the relationships between domains. Reconceptualisation is determined when the release of constraints to factor loadings results in change in the zero and non-zero factor loadings.

The final stage involves the assessment of true change in QoL through including constraints of the common factor (i.e., QoL) mean and variance across time. If Model 4 proves to have a significantly different fit than Model 3, then these values can be interpreted as a measure of true change.

The models were tested using the maximum likelihood estimation (*mle*) of the statistical package Amos (Version 6; [36]). Overall fit was examined using the Chi-Square Test (χ^2), and the evaluation of incremental [i.e., Comparative Fit Index (CFI) and Incremental Fit

Index (IFI)], and absolute [i.e., Standardized Root Mean Square Residual (SRMR) and Root

Mean Square Error of Approximation (RMSEA)] fit indexes. Cut-off values of ≥.95 for CFI

and IFI, and values of ≤.08 and .06 for SRMR and RMSEA respectively, indicate a relatively

good fit between the hypothesized model and observed data [37]. Invariance was

investigated by calculating χ^2 difference (significant when $p<.05$), and CFI difference

(significant when >.01) [38].

Results

Four schools in South West England participated in the study at baseline, comprising 580

participants; only 12 pupils and/or their parents opted out of the study. One school withdrew

at Time 2 due to staff changes, and therefore pupils from this school are not included in the

analysis. Complete data were obtained from 356 students; 180 boys (51%) and 176 girls,

with mean age of 14.05 years (SD = .31) at Time 1, and 14.9 years (SD = .30) at Time 2.

There was no pattern of missing data either within or across schools. Then-test data was

obtained from 279 students in two schools.

Tests of response shift though research-design methods

Assessment of recalibration analysis (then-test)

Significant differences were reported between pre-test and then-test scores for three

domains; peer relationships, mood and emotions, and autonomy (Table 2). This indicated

that a recalibration RS had taken place. Participants retrospectively judged their peer related

QoL to be better than they had rated it at the time, indicating that the standard of QoL

participants labelled 'good', was lower at Time 2 than at Time 1. However, the mood and

emotions and autonomy domains were rated as poorer retrospectively, suggesting that a

higher standard of QoL was necessary to be considered 'good' for these domains at Time 2.

While these significant changes were contrary to Hypothesis 1, the effect size were small in

each case (*d* = -.29, -.17 and .15 for mood and emotions, autonomy and peer relationships respectively). There was no significant change in the remaining seven domains.

Table 2   Comparison of mean scores for Kidscreen items at pre-test and then-test

Domain	pre-test Mean (SD)	then-test Mean (SD)	Then-test (pre-test vs then-test)	
			t statistic (*df*)	*d* [a]
Physical health	3.39 (.91)	3.40 (1.02)	.23 (235)	.02
Family relationships	3.60 (1.20)	3.60 (1.15)	.00 (234)	0
Self perception	3.06 (1.17)	3.06 (1.21)	.00 (230)	0
Psychological well-being	3.61 (1.05)	3.58 (1.00)	-.31 (230)	-.02
School	3.51 (1.01)	3.39 (.97)	-1.89 (232)	-.12
Peers	3.70 (.98)	3.86 (.95)	2.24 (232)*	.15
Financial	3.63 (1.20)	3.61 (1.05)	-.31 (230)	-.02
Mood & emotions	4.24 (.99)	3.91 (1.12)	-4.43 (234)***	-.29
Bullying	4.17 (1.03)	4.05 (1.16)	-1.79 (229)	-.12
Autonomy	3.91 (1.02)	3.72 (.93)	-2.59 (233)*	-.17

*p<.05; ** p<.01; *** p<.001; *df* = degrees of freedom, [a] *d* =*t*/square root(*df*+1): small *d* >.2, moderate *d* >.5, large *d* >.8

Assessment of reprioritisation and reconceptualisation response shift

There was little change in the relative importance of domains when rank order was inspected (see Table 3). Contrary to Hypothesis 2, this indicated that no reprioritisation RS had taken place.

Table 3   Ranks of self-reported importance of each of the Kidscreen QoL domains at
              pre-test and post-test.

	Pre-test[a]			Post-test		
	Rank	Rating (SD)	95% CI	Rank	Rating (SD)	95% CI
Peers	1	4.77 (.54)	4.73-4.85	1	4.48 (.81)	4.38-4.56
Family relationships	2	4.59 (.75)	4.51-4.68	3	4.13 (1.03)	4.03-4.26
Bullying	3	4.47 (.90)	4.42-4.62	2	4.15 (1.10)	4.03-4.28
Autonomy	4	4.46 (.63)	4.38-4.53	4	4.03 (.86)	3.93-4.12
Self perceptions	5	4.24 (.83)	4.13-4.33	6	3.86 (.91)	3.77-3.97
Psychological well-being[a]	6	4.20 (.86)	4.01-4.30	5	3.99 (.85)	3.92-4.11
Physical Health	7	4.17 (.75)	4.08-4.25	7	3.86 (.88)	3.71-3.91
Financial	8	4.01 (.92)	3.87-4.09	9	3.63 (.96)	3.52-3.74
School	9	3.91 (.97)	3.83-4.05	8	3.66 (.93)	3.59-3.80

[a] - mood and emotion and psychological well-being were combined for importance ratings

Means for the importance of domains also remained relatively stable in absolute terms, i.e.,

mean scores ranged from 3.82 and 4.76 on a 5-point scale at pre-test, and from 3.60 to 4.46

at post-test (a score of 3 represents *moderately important*, and 4 *very important)*. These

related to a small effect size even in the domains where there was greatest change (bullying;

d=.22 and peer; *d*=.26), and minimal change for all other domains (*d*≤.2). Confidence

intervals (95%) indicated that mean importance ratings for all domains were significantly

different from zero (i.e., all were indicative of QoL). Thus, in line with Hypothesis 3, there

was no evidence of reconceptualisation RS.

Statistical modelling of response shift

Following the addition of three covariances between residual error terms of three theoretically linked domains (mood and emotions, bullying and self-perceptions) the CFA for the measurement model underpinning QoL showed an adequate fit to the data (χ^2(64)=221.875, p<.001; IFI =.92, CFI =.92; SRMR =.05; RMSEA =.065 (90% CI: .055 - .074). Model 2 showed a significant decline in goodness of fit, indicating that some form or degree of response shift was present [34] (Table 4). Constraints on the model parameters were released sequentially following scrutiny of differences between Time 1 and Time 2 estimates of factor loadings, variances and intercepts (no additional paths were suggested according to modification indices). Comparisons of fit against sequential models were conducted following the release of each constraint until the model fit represented no significant change. The final model was specified with the release of the invariance constraints of two intercepts; autonomy and bullying. This was indicative of recalibration RS for these two domains. The intercept for autonomy was lower, and bullying higher at Time 2 than at Time 1, consistent with the outcomes obtained by the research-design technique. All constraints on factor loadings and residual covariances between domains remained tenable, indicating an absence of reprioritization or reconceptualisation RS.

Assessment of true change in QoL

A secondary outcome of the present study is to comment on the *true change* in QoL within the present sample, once RS has been either accounted for or ruled out. In view of the recalibration RS detected in three domains, true change in QoL was assessed by comparing post-test ratings with then-test scores (see Table 5) [12]. A significant improvement in QoL was reported for three domains, all of small effect size; self-perceptions (*d*=.24), mood and emotions (*d*=.16), and bullying (i.e., participants perceived less bullying, *d*=.27). A significant deterioration was reported for physical health (*d*=-.14). It is of note that failing to adjust for recalibration would have resulted in a falsely significant result for change in the autonomy and peer domains, and significant results overlooked for the bullying domain. Furthermore,

adjustment for RS in the mood and emotion domain resulted in the finding of a significant improvement in functioning, which without adjustment would have resulted in a significant finding of a deterioration in QoL (i.e., pre-test M=4.24, post-test M=4.09, but then-test M=3.91). No significant true change was detected through SEM; Model 4 was not significantly different from Model 3 (see Table 4).

Table 4   Fit indices following constraints of estimated pathway coefficients for the invariance of factor loadings over time.

Model	IFI	RMSEA	SRMR	CFI	χ^2	Df	p
Model 1: unconstrained	.917	.065 (CI*: .055-.074)	.053	.916	221.88	64	
Model 2: no response shift	.897	.058 (CI: .051-.066)	.060	.897	288.59	96	
Difference from Model 1				.019	66.71	32	<.001
Model 3: response shift	.910	.055 (CI: .047-.063)	.060	.910	261.76	94	
Difference from Model 1[†]				.013	39.88	30	.107
Model 4: no real change	.909	.055 (CI: .047-.063)	.060	.909	264.31	95	
Difference from model 3				.001	2.55	1	.110

*CI = 90% confidence interval; [†] = Model 3 is computed by releasing constraints of Model 2 until fit is not significantly poorer than Model 1

df= degrees of freedom; p= significance value for χ^2 difference test; IFI= incremental fit index; CFI = comparative fit index, RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual

Table 5   Comparison of mean scores for Kidscreen items at then-test and post-test
                (i.e., *true change*)

Domain	then-test Mean (SD)	post-test Mean (SD)	True effect (then-test vs post-test)	
			t statistic (*df*)	*d* [a]
Physical health	3.40 (1.01)	3.26 (1.00)	-2.15 (232)*	-.14
Family relationships	3.60 (1.15)	3.50 (1.25)	-1.79 (233)	-.12
Self perception	3.08 (1.21)	3.31 (1.13)	3.56 (228)***	.24
Psychological well-being	3.58 (1.00)	3.65 (1.02)	1.25 (231)	.08
School	3.39 (.97)	3.30 (.99)	-1.44 (229)	-.09
Peers	3.87 (.95)	3.94 (.89)	1.04 (231)	.07
Financial	3.60 (1.05)	3.53 (1.21)	-1.11 (234)	-.07
Mood & emotions	3.92 (1.11)	4.09 (1.01)	2.49 (233)*	.16
Bullying	4.05 (1.16)	4.29 (1.05)	4.05 (232)***	.27
Autonomy	3.72 (.92)	3.66 (1.05)	-1.01 (229)	-.07

 *p<.05; ** p<.01; *** p<.001 *df* = degrees of freedom,
 [a] *d* =*t*/square root(*df*+1): small *d* >.2, moderate *d* >.5, large *d* >.8

Discussion

The present study aimed to examine the extent of RS in QoL as a result of normative yet
significant developmental processes in a healthy mid-adolescent sample. The purpose of
this was to (i) explore whether the process of adolescent development may itself be a trigger
to significant change in adolescents' conceptualisation and expectations with relation to QoL,
and (ii) by so doing relate this evidence to the assumption made by clinical studies that QoL
is stable in healthy adolescent control group. Change in self-reported QoL was assessed
between two time points for evidence of recalibration, reprioritisation and reconceptualisation
RS.

All three types of RS were analysed using both a research-design technique and statistical modelling. There is no standard rule to apply in reconciling results obtained using different methodologies when results from both are available [39]. We therefore propose an approach which combines information from both sources, but considers results from the then-test to provide the primary evidence of RS. The results of statistical modelling are interpreted to provide secondary evidence to either support findings when the results are consistent, or to provide a check of the degree to which then-test results constitute change of a meaningful magnitude or are generalisable, where differences are found. This approach is based on recommendations that detection of RS should not be purely statistically driven without recourse to theory [34]. The then-test represents a theory-driven within-subject test of change [20], consistent with this recommendation. SEM provides an assessment of group level effects over time, which effectively tests the fit of observed data to a hypothesised model of statistical changes, i.e., represents a confirmatory, rather than *a priori* test of effect.

Hypothesis 1 predicted that there would be no recalibration RS. The hypothesis was only partially supported, as a small but significant recalibration effect was found using the then-test for three of the ten domains assessed (autonomy, peers, mood and emotions). This was reinforced by the convergence of research-design and statistical analyses for just one domain; autonomy. Participants rated the autonomy domain more highly at pre-test (Time 1) than it would be judged by current standards (Time 2), indicating standards for achieving "good" QoL had become more demanding. In suggesting possible reasons for this effect, the recalibration of autonomy may be as a result of adolescents responding to objective changes in their permitted autonomy experienced as part of the process of growing up. As they get older, children and adolescents are provided with more choice and greater responsibility both at home and at school, reflected by noticeable changes in objective markers such as choosing school subject options, and being allowed greater freedom to spend time away from home and parents. Research suggests that children can be satisfied

with either low or high levels or autonomy, as long as it is consistent with their experience and expectations [40]. Thus, when retrospectively judging autonomy against current standards, it is possible that participants would assume prior dissatisfaction with the lesser freedom they were afforded in the previous year, and hence adjust (recalibrate) their judgement to rate their QoL as poorer.

Although two other domains (peer relationships and mood and emotions) showed significant recalibration when assessed by then-test, and bullying was found to vary across time in the statistical model, these effects were not confirmed by the alternative method of RS assessment. Given that the effect size of the difference between pre- and then-tests was small in each case, the lack of convergence between analyses was interpreted to indicate that the shift may not reflect a meaningful degree of change. Investigation with further healthy samples would be informative. The statistical and research design methods concurred that no recalibration effect was reported for the remaining six domains.

Contrary to Hypothesis 2, there was no evidence that the sample had undergone a reprioritisation RS when assessed either through self-report, or through the statistical comparison of factor structures. The most important domains contributing towards QoL on both occasions were peers, family relationships and autonomy. These domains are consistent with the developmental model [9] that frames the task of adolescence as developing a personal identity separately from parents, and relying to a greater degree on peer judgements and support as a basis for global self-perceptions [24]. The domains that were least important, although still not unimportant, were personal finances and school. Students were attending state run comprehensive schools in mixed areas, therefore this lower level of importance does not result from participants experiencing abnormally high financial security or privileged schooling. Instead, it seems to suggest that relationships and self-development remain of a higher priority to young people than either their financial situation, or doing well at school during this period. In line with Hypothesis 3, both the

statistical and research design analyses reported no reconceptualisation RS effect. All ten domains included within the Kidscreen QoL measure remained of importance in adolescents' conceptualisation of QoL over the study period.

The lack of a reprioritisation was unexpected given the large body of adolescent development literature that reports considerable changes in many areas related to QoL, such family and peer relationships [41], physical self-perceptions [22,42] and life goals [9]. However, the findings are consistent with prior research, not directly assessing RS, but which has reported similar stability in the relative importance of domains. For example, in adolescents aged 11 to 17, Petito and Cummins found no change in the relative importance or satisfaction of subjective domains of the ComQoL instrument [43]. One reason for why no reprioritisation was observed may be due to the subjective nature of QoL, which implies that good QoL can potentially be achieved irrespective of whether expectations/standards etc. are high or low, as long as these can be realistically met [8]. Thus, rather than bringing about a reprioritisation of the constituents of QoL, the changes in priorities and values that are documented to take place during adolescence may instead lead to a change in the way each domain is satisfied. For example, children aged 10 or 18 may expect very different levels of support and involvement from their close friendships in order to perceive their peer related QoL to be good, yet as long as both are achievable at each respective age then QoL may remain stable. Qualitative work with adolescents of different age groups would help to explore this hypothesis further.

Having accounted for recalibration effects there was found to be a small but significant, positive true change in QoL for four domains using the then-test; physical health, self-perceptions, mood and emotions, and bullying. These findings were not supported by the statistical analysis. In line with our proposed method of reconciling the two research-design and statistical approaches, we concluded that while statistically significant, the change in physical health and mood and emotions were associated with effect sizes at or below the

point conventionally accepted to indicate meaningful change. However, small but meaningful improvements (i.e., effect size $d \geq .2$) in self-perceptions and bullying were accepted. These limited effects are largely in line with the outcome of no overall effect obtained through statistical modelling. However, they suggest that the then-test approach is more sensitive to detecting change of multidimensional QoL inventories that statistical techniques.

Limitations

There are several limitations to the longitudinal study of QoL and the application of the theory of RS. First, there is a risk of biased recall effects influencing the then-test results. The one year period between the initial measurement and retrospective then-test in the present study posed a considerable challenge to memory. However, by age 15, adolescents are considered to have cognitive capacities similar to adults [44], with whom acceptable, if imperfect, reliability of the then-test has been established [e.g. 45]. The risk of recall bias would thus not be expected to be greater for this age group than for adults, with whom such measures were developed. In fact, more cues to memory may be available within a structured school environment than are available to adults, but future work would be useful in assessing the reliability of the then-test in a younger age group.

A further limitation resulted from the SEM with which the statistical tests of RS were conducted reporting a model fit slightly below the levels usually considered acceptable (specifically; CFI and IFI =.92, vs cutoff $\geq .95$, and RMSEA = $\leq .065$, vs cutoff .06) [cf. 37]. Furthermore, to obtain this fit, the addition of three covariance terms were necessary, indicating a common source of variance between three of the ten domains (bullying, self-perceptions and mood and emotions). A two factor solution was specified, but resulted in poorer fit than the single factor model used, therefore the single factor solution was retained for the RS analysis.

23

Beyond the limitations inherent to all studies of RS, a further limitation of the present study was the restricted age group. Although it was expected that some adjustments in response to puberty would be ongoing at ages 14 to 15 in both sexes [46], it is possible that reprioritisation of QoL domains may have begun earlier and stabilised by this age, or may be yet to take place. Thus, while the present study provides an original example of the assessment of RS in adolescents, investigation across a much wider age group is called for.

The present study diverges from the majority of RS research by its aim to test effects in a healthy sample in the absence of an objective indicator of significant change. Our aim was to assess the potential for RS during the transitional phase of adolescence, which represents a gradual and ongoing period of adjustment rather than response to an abrupt change. However, to be in line with RS theory relating change to specific events, proxy measures such as the onset of puberty, starting work or changes in education post sixteen, could be used to permit a tighter analysis. This would allow control of time since the onset of an objective event, and standardisation of the type of significant life event studied against which to measure and compare RS in QoL.

Conclusions and Practical Implications

Overall, the results of the present study showed that there was very little change in the way QoL is conceptualised and prioritised as a result of the normative developmental changes experienced in mid-adolescence. However, a small but significant change in the internal standards used to measure QoL was found in three domains. Evidence of this type is overdue, and has a number of implications for the interpretation of adolescent QoL assessment in research and clinical practice. First, the findings present preliminary evidence that the conceptualisation of QoL is stable within a healthy adolescent sample. This is consistent with the theoretical basis of Sprangers and Schwartz's model, which suggests that meaningful RS occurs as a psychological response only to highly significant, distinctive and threatening life events. Our findings therefore suggest that normative

24

adolescent development is not sufficiently intense an experience to be considered a trigger to reprioritisation and reconceptualisation RS. However, our findings do suggest that recalibration effects may occur in healthy populations at this time. As such, to verify the accuracy of longitudinal comparisons, researchers working with adolescents may be advised to include a simple measure of recalibration RS (e.g., then-test items) in their assessments.

A secondary aim of the study was to test the application of RS methodology within an adolescent sample. Through combining research-design and statistical techniques, we were able to compare the findings that would have resulted from each technique used alone. The outcomes were largely similar, in that where the two methods did vary in terms of statistical significance (e.g., which domains showed a recalibration effect), the differences could be accounted for by inspecting effect size (i.e., all were small or negligible) in place of significance value (i.e., p value). The only other outcome that may have led to different conclusions was that of true change, which were reported as meaningful effects in the self-perceptions and bullying domains according to research-design methods, but not by statistical modelling. Given that only two results from 30 statistical comparisons showed a different outcome (three types of response shift for ten domains), overall the two techniques were found to show good consistency, and thus provide some preliminary evidence of validity within a mid-adolescent age group.

A further application of the study findings is in the interpretation of clinical studies. The majority of studies comparing the QoL of healthy versus clinical populations assume, but do not explicitly test, the stability of the conceptualisation of QoL within healthy controls [47]. The present findings support the assumption for the conceptualisation and prioritisation of QoL, but suggest some recalibration may take place during adolescence within a healthy control sample. This could be controlled for by a range of available quantitative (e.g., inclusion of then-test items), and qualitative (e.g., preference based methods) techniques [21]. Accounting for RS in healthy adolescents may also have implications for the

25

interpretation of cohort studies reporting age-related changes in QoL. The majority of such reports suggest an age-related decrease in QoL [17,18,19]. After controlling for recalibration effects the results of the present study do not support this finding, although this interpretation is limited to the age range of the present sample. Further work to assess RS across the wider age span is called for to investigate whether a true effect of decreasing QoL persists after any potential recalibration of domains is accounted for.

Finally, adolescent QoL data is not only used to assess the impact of illness and treatment, but also the impact of social interventions for healthy children and adolescents considered to be "at-risk" [4]. Better understanding of potential changes in QoL in healthy, community dwelling adolescents is crucial in determining whether or not reported change in QoL may be attributable to normative developmental processes, or indicate non-normative distress. Given the association between adolescent QoL, health risk behaviours and mental health problems [20], identifying at-risk individuals and groups for whom intervention may be particularly crucial is an important direction for research.

Footnote:

An anonymous reviewer suggested the covariance of domains may suggest an alternative two factor solution loading these three domains onto a different higher-order domain. This was tested but did not provide adequate or improved fit to the data, thus the modified single factor solution was retained

Running head: Response shift in healthy adolescents

References

1. Eiser, C. & Morse, R. (2001). The measurement of quality of life in children: Past and future perspectives. *Journal of Developmental and Behavioral Pediatrics,* 22,(4): 248-256.

2. De Civita, M., Regier, D., Alamgir, A. H., Anis, A. H., FitzGerald, M. J. & Marra, C. A. (2005). Evaluating health-related quality-of-life studies in paediatric populations - some conceptual, methodological and developmental considerations and recent applications. *Pharmacoeconomics,* 23,(7): 659-685.

3. Sawyer, M. G., Reynolds, K. E., Couper, J. J., French, D. J., Kennedy, D., Martin, J., et al. (2004). Health-related quality of life of children and adolescents with chronic illness--a two year prospective study. *Quality of Life Research,* 13,(7): 1309-19.

4. Davidson-Arad, B. (2005). Fifteen-month follow-up of children at risk: Comparison of the quality of life of children removed from home and children remaining at home. *Children and Youth Services Review,* 27,(1): 1-20.

5. Department for Education and Skills (2003). *Every Child Matters*. London, DfES Publications.

6. Varni, J. W. & Burwinkle, T. M. (2006). The pedsql as a patient-reported outcome in children and adolescents with attention-deficit/hyperactivity disorder: A population-based study. *Health and Quality of Life Outcomes,* 4: 26.

7. Ravens-Sieberer, U. (2001). The revised KINDL-R: Final results on reliability, validity and responsiveness of a modular HRQOL instrument for children and adolescents. *Quality of Life Research* 10,(3): 199.

8. The WHOQOL Group. (1995). The world health organisation quality of life assessment (WHOQOL): Position paper from the world health organisation. *Social Science & Medicine,* 41: 1403-1409.

9. Goldbeck, L., Schmitz, T. G., Besier, T., Herschbach, P. & Henrich, G. (2007). Life satisfaction decreases during adolescence. *Quality of Life Research,* 16,(6): 969-79.

10. DiClemente, R. J. (1996). *Handbook of adolescent health risk behavior.* New York, Plenum.

11. Williams, P. G., Holmbeck, G. N. & Greenley, R. N. (2002). Adolescent health psychology. *Journal of Consulting and Clinical Psychology,* 70,(3): 828-42.

12. Schwartz, C. E., Sprangers, M. A., Carey, A. & Reed, G. (2004). Exploring response shift in longitudinal data. *Psychology & Health,* 19,(1): 51-69.

13. 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): 1507-15.

14. Breetvelt, I. S. & Van Dam, F. S. (1991). Underreporting by cancer patients: The case of response-shift. *Social Science and Medicine,* 32,(9): 981-7.

15. Gibbons, F. X. (1999). Social comparison as a mediator of response shift. *Social Science Medicine,* 48,(11): 1517-30.

16. Rapkin, B. D. (2000). Personal goals and response shifts: Understanding the impact of illness and events on the quality of life with people living with aids. *Adaptation to changing health. Response shift in quality of life research.* Schwartz C. E. & Sprangers, M. A. American Psychological Association, Washington DC.

17. Bisegger, C., Cloetta, B., von Rueden, U., Abel, T. & Ravens-Sieberer, U. (2005). Health-related quality of life: Gender differences in childhood and adolescence. *Sozial-Und Praventivmedizin,* 50,(5): 281-91.

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

19. Drukker, M., Kaplan, C., Schneiders, J., Feron, F. J. & van Os, J. (2006). The wider social environment and changes in self-reported quality of life in the transition from late childhood to early adolescence: A cohort study. *BMC Public Health,* 6: 133.

20. Zullig, K. J., Valois, R. F., Huebner, E. S., Oeltmann, J. E. & Drane, J. W. (2001). Relationship between perceived life satisfaction and adolescents' substance abuse. *Journal of Adolescent Health,* 29,(4): 279-88.

21. 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): 1531-48.

22. Graber, J. A., Petersen, A. C., & Brooks Gunn, J. (1996b). Pubertal Processes: Methods; Measures and Models. In Graber, J. A., Brooks Gunn, J. & Petersen, A. C. (Eds.) *Transitions through adolescence: Interpersonal Domains and Context.*

23. Craft, L. L., Pfeiffer, K. A. & Pivarnik, J. M. (2003). Predictors of physical competence in adolescent girls. *Journal of Youth and Adolescence,* 32,(6): 431-438.

24. Harter, S. (1990). Self-representations in childhood and adolescence. *Handbook of child psychology: Social, emotional and personality development*. Ed: Damon, W.

25. Moshman, D. (1999). *Adolescent Psychological Development.* New Jersey, Erlbaum.

26. Ravens-Sieberer, U., Gosch, A., Rajmil, L., Erhart, M., Bruil, J., Duer, W., et al. (2005). Kidscreen-52 quality-of-life measure for children and adolescents. *Expert Review of Pharmacoeconomics & Outcomes Research,* 5,(3): 353-364.

27. The Kidscreen Group (2006). *The Kidscreen Questionnaires. Quality of life questionnaires for children and adolescents.* Pabst Science Publishers.

28. Walls, T. A., & Little, T. D. (2005). Relations Among Personal Agency, Motivation, and School Adjustment in Early Adolescence. *Journal of Educational Psychology, 97*(1), 23-31.

29. Gillison, F., Standage, M. & Skevington, S. (2008) Changes in quality of life and psychological need satisfaction over the transition to secondary school. *British Journal of Educational Psychology, 78*, (1): 149-162.

30. Saxena, S., Carlson, D. & Billington, R. (2001). The who quality of life assessment instrument (whoqol-bref): The importance of its items for cross-cultural research. *Quality of Life Research,* 10,(8): 711-21.

31. Skevington, S. M. & O'Connell, K. A. (2004). Can we identify the poorest quality of life? Assessing the importance of quality of life using the whoqol-100. *Quality of Life Research,* 13,(1): 23-34.

32. Howard, G. S., Schmeck, R. R. & Bray, J. H. (1979). Internal invalidity in studies employing self-report instruments: A suggested remedy. *Journal of Educational Measurement,* 16: 129-135.

33. Gillison, F. B., Standage, M. & Skevington, S. M. (2006). Relationships among adolescents' weight perceptions, exercise goals, exercise motivation, quality of life and leisure-time exercise behaviour: A self-determination theory approach. *Health Education Research,* 21,(6): 836-47.

34. Oort, F. J. (2005). Using structural equation modeling to detect response shifts and true change. *Quality of Life Research,* 14,(3): 587-598.

35. Armenakis, A. A. (1988). A review of research on the change typology. *Research in organizational change and development an annual series featuring advances in theory, methodology and research*. Woodman R. W. & Pasmore W. A., JAI Press.

36. Arbuckle, J. L. (2003). *AMOS (Version 6.0).* Chicago, IL, SmallWaters Corporation.

37. Hu, L. T. & Bentler, P. M. (1999). Cutoff criterion for fit indexes in covariance structural analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling,* 6: 1-55.

38. Cheung, G. W. & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling,* 9,(2): 233-255.

39. Schwartz, C. E., Bode, R., Repucci, N., Becker, J., Sprangers, M. A. G. & 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.

40. Stoll, L., Stobart, G., Martin, S., Freeman, S., Freedman, E., Sammons, P., et al. (2003). *Preparing for Change: Evaluation of the Implementation of the Key Stage 3 Strategy Pilot.*: Department of Education and Skills.

41. Bukowski, W. M. & Sandberg, D. (1999). Peer relationships and quality of life. *Acta Paediatrica,* 88: 108-109.

42. Robbins, L. B., Pender, N. J., Ronis, D. L., Kazanis, A. S. & Pis, M. B. (2004). Physical activity, self-efficacy, and perceived exertion among adolescents. *Research in Nursing & Health,* 27,(6): 435-446.

43. Petito, F. & Cummins, R. A. (2000). Quality of life in adolescence: The role of perceived control, parenting style, and social support. *Behaviour Change,* 17,(3): 196-207.

44. Commons, M. L., Richards, F. A. & Armon, C. (1984). *Beyond formal operations : Late adolescent and adult cognitive development.* New York ; Eastbourne, Praeger.

45. Visser, M. R., Oort, F. J. & Sprangers, M. A. (2005). Methods to detect response shift in quality of life data: A convergent validity study. *Quality of Life Research,* 14,(3): 629-39.

46. Coleman, L. & Coleman, J. (2002). The measurement of puberty: A review. *Journal of Adolescence,* 25,(5): 535-50.

47. Schwartz, C. E. & Sprangers, M. A. (2000). *Adaptation to changing health: Response shift in quality-of-life research.* Washington, D.C, American Psychological Association.