

Investigating the Substantive Aspect of Construct Validity for the Satisfaction with Life Scale Adapted for Children: A Focus on Cognitive Processes

Anne M. Gadermann · Martin Guhn · Bruno D. Zumbo

Accepted: 18 March 2010 / Published online: 23 April 2010
© Springer Science+Business Media B.V. 2010

Abstract The objective of this study was to investigate the substantive aspect of construct validity of the Satisfaction with Life Scale adapted for Children (SWLS-C; Gadermann et al. in Soc Indic Res 96:229–247, 2010). Specifically, the study examined the cognitive processes of children when responding to the items of the SWLS-C to find out how they interpret and respond to the items. Think-aloud protocol interviews were conducted with 55 students in grades 4–7 (58% girls, mean age of 11 years, ranging from 8.8 to 12.8 years) and content analysis was used to analyze the data. The findings indicate that most children had no difficulty in understanding the items, and used mainly two strategies for responding to the items: (1) an absolute strategy, in which children used absolute statements to indicate the presence or absence of something that is important for their judgment of their satisfaction with life, and (2) a relative strategy, in which children used relative or comparative statements. In the absolute statements, children primarily referred to social relationships, personal characteristics, time use, and possessions. In the relative statements, the children primarily compared what they have to (a) what they want (b) what they had in the past, (c) what other people have, and (d) what they feel they need. Furthermore, most children considered it important that information on their life satisfaction is obtained. The results are discussed with regard to multiple discrepancies theory (Michalos in Soc Indic Res 16:347–413, 1985) and previous empirical findings. The results provide insights into children’s cognitive processes when responding to items on life satisfaction and provide validity evidence that the SWLS-C is an appropriate measure to assess life satisfaction in children of this age.

Keywords Satisfaction with Life Scale adapted for Children · Life satisfaction · Subjective well-being · Children · Cognitive processes · Construct validity

1 Introduction

Measuring and monitoring children’s well-being has received increasing attention and interest over the last decade (Ben-Arieh 2006; Ben-Arieh and Goerge 2001). One of the

A. M. Gadermann (✉) · M. Guhn · B. D. Zumbo
University of British Columbia, Vancouver, BC, Canada
e-mail: gaderman@interchange.ubc.ca

reasons is the “movement toward accountability-based public policy that requires increasing amounts of data to provide more accurate measures of the conditions children face and the outcomes various programs achieve” (Ben-Arieh 2005, p. 573). Specifically, measuring and monitoring children’s well-being are important to gain a better understanding of and enhanced knowledge about their well-being and to inform and evaluate policies and programs with the aim of improving children’s well-being (e.g., Ben-Arieh and Goerge 2006; Ben-Arieh et al. 2001; Fronas 2007).

With regard to what is measured, there has been a shift from early indicators that focused on measuring survival or negative facets of children’s lives to an approach that is more holistic by also measuring assets and positive aspects of children’s lives. Furthermore, indicators assessing children’s ‘well-becoming’ (predicting transition to and well-being in adulthood) have been supplemented by indicators assessing current well-being during childhood (Ben-Arieh 2006; Ben-Arieh and Goerge 2001).

One approach that combines the focus of emphasizing the positive aspects of individuals’ lives with the focus on current well-being is the field of subjective well-being research. Subjective well-being is considered to consist of positive and negative affect, life satisfaction, and domain satisfaction (e.g., Diener et al. 1999). One of the most commonly used instruments to assess satisfaction with life in adults is the Satisfaction with Life Scale (Diener et al. 1985). This scale has been adapted by subject matter experts for children in grades 4–7 and the psychometric properties of this scale, the Satisfaction with Life Scale adapted for Children (SWLS-C), have been shown to be favourable with a sample of children in grades 4–7 (Gadermann et al. 2010). However, in order to develop and/or validate an instrument, it is recommended to use experiential experts (i.e., members of the target population) to investigate the cognitive processes that respondents use to answer questions (Collins 2003; Willis 2005). This is especially important for measures developed for children, as the conceptualization of the adult test developers can potentially be quite different from the one of children. Therefore, in the present study we used think-aloud protocols, a cognitive interviewing technique, with children for evaluating the items of the SWLS-C. This technique has shown to be useful to investigate the cognitive processes of children in previous studies (e.g., Cremeens et al. 2006a; Fox et al. 1983; Lodge et al. 1998, 2000; Rebok et al. 2001). We used this technique to investigate the cognitive processes of children when answering the items of the SWLS-C, in order to explore how the children arrive at their specific response. The investigation of cognitive processes during a measurement task is one way to evaluate the substantive aspect of construct validity (Messick 1995). In the following, we will first provide a brief overview of the importance of cognitive interviewing techniques for the validation of self-report measures before describing our study.

1.1 The Importance of Validating Self-Report Measures Using Cognitive Interviewing

Self-report measures, such as questionnaires and surveys, are commonly used in the social sciences to collect data on psychological constructs, such as subjective well-being. The information from such questionnaires and surveys is used for a variety of reasons; for example, to evaluate intervention programs, to describe societal conditions, and to inform public policy. Accordingly, self-report measures can have far-reaching consequences. However, the data collected with such measures are obviously only as meaningful as the questions that are asked and the responses that participants provide (Schwarz 1999). Therefore, the thorough development and ongoing validation of questionnaires and surveys is of special relevance. In this regard, it is of interest to investigate the substantive aspect of

construct validity (Messick 1995). Specifically, it is of interest to investigate how and why respondents arrive at their answers and how this is influenced by characteristics of the respondent and the questionnaire (and their potential interactions). In other words, one needs to ask the question: What are the underlying cognitive processes that result in respondents providing responses to self-report questions? In the last three decades, this topic has become of increasing interest for researchers in areas such as psychology and survey methodology. In the 1980s, the Cognitive Aspects of Survey Methodology (CASM) initiative started, an interdisciplinary movement with the aim to improve the quality of self-report data and “to bridge the communication gaps between survey research and the cognitive and social sciences, and to initiate CASM research that would benefit survey applications as well as basic cognitive research” (Sirken and Schechter 1999, p. 1). CASM research investigates the cognitive processes that underlie self-reports in order to understand how these processes function. CASM research can thereby influence questionnaire design (e.g., by suggesting how to redesign a questionnaire if the items do not perform/function as expected) as well as stimulate basic research on cognition (Sirken and Schechter).

This is in line with contemporary views of measurement validity, in that cognitive processes or models are investigated in the validation process to support the inference one makes from the scale scores. As Messick (1989) stated, “validity is an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the *adequacy* and *appropriateness* of *inferences* and *actions* based on test scores or other modes of assessment” (p. 13). In Messick’s unified view of validity, construct validity lies at the core and “comprises the evidence and rationales supporting the trustworthiness of score interpretation in terms of explanatory concepts that account for both test performance and score relationships with other variables” (Messick 1995, p. 743). As mentioned above, one aspect of construct validity is the substantive aspect, which highlights the importance of theories and process modeling in examining the processes that are involved in the measurement task, and which can be investigated using different approaches such as cognitive interviewing or modeling response times. Evidence based on response processes is also listed as one of the five sources of validity in the *Standards for Educational and Psychological Testing* (American Educational Research Association et al. 1999). The research question ‘What does a score on a self-report measure provided by a participant mean?’ is also very much in line with what has been described as a strong form of construct validity, which “should provide an *explanation* for the test scores, in the sense of the theory having explanatory power for the observed variation in test scores” (Zumbo 2009, p. 69; see also Zumbo 2007).

Although this illustrates the importance of investigating the substantive aspect of construct validity, this aspect is often not investigated in the development or evaluation of measures. It is worth noting that much of the validation research is about correlations with other variables and hence is not explanatory. For example, a study by Cizek et al. (2008) investigated (among other things) the types of validity evidence reported in the current edition of the *Mental Measurements Yearbook*. The authors report that response processes were only investigated for 1.8% of the measures, whereas criterion-related (correlational) validity evidence was provided for 67.2% of the measures. None of the personality/psychological measures or social measures reported on response processes as sources of validity evidence, whereas it was reported in 5.9% of the cases of the developmental measures, 4.0% of the behavioural measures, and 3.7% of the achievement measures. Similarly, in Cremeens et al.’s (2006b) review of health-related measures, including quality of life measures, for children aged 3–8 years, cognitive processes are not reported

at all as source of validity evidence. (It should be noted, however, that children were consulted during the process of item development for 40% of the measures, e.g., through interviews, and that some form of pilot testing with children was conducted for 47% of the measures.)

In line with this, McColl et al. (2003) state that cognitive techniques have rarely been applied to well-being or quality of life research, although in recent years there has been a development in this direction as indicated by the formation of the 'International Study Group on Cognitive Aspects of Quality of Life Research' (Barofsky et al. 2003). Nonetheless, with regard to investigating the cognitive processes underlying self-reports of children, there are few studies that employed think-aloud protocols with children in the area of quality of life and subjective well-being (Cremeens et al. 2006a; see also Riley 2004, for the area of health). It is noteworthy that children and adolescents are often not included in the evaluation of measures, given that several studies have indicated that having children as subject matter experts adds a critical component to the development and evaluation of measures for children and adolescents (e.g., Cremeens et al. 2006a; Rebok et al. 2001; Schilling et al. 2007; Stewart et al. 2005).

One of the few studies using think-aloud protocols with children in the area of quality of life and well-being was conducted by Cremeens et al. (2006a). In their study, children aged 5–9 were asked to think aloud while responding to the TedQL, a generic quality of life measure. The authors report that children used several strategies for responding to the items, namely (1) social comparisons; (2) stable character references; (3) concrete examples; (4) other reasons; and (5) no reason given. The type of strategy utilized was related to the age of children and type of item (ability and social items). Specifically, older children were more likely to use the social comparison and concrete examples strategies than younger children, whereas younger children were more likely to provide no reason than older children. Furthermore, concrete examples was the most frequently used strategy and there was no statistically significant difference in the use of this strategy by type of item. In contrast, the social comparison strategy was used more frequently for ability than social items, whereas the stable character references strategy was used more frequently for social than ability items.

Similar to Cremeens et al.'s study, we were interested in investigating the cognitive processes of children when responding to the items of the SWLS-C and whether response strategies used would be associated with demographic characteristics of the children.

2 Method

This section is structured into the following parts: (1) Sample; (2) measure; (3) think-aloud protocols; (4) procedure; (5) development of coding categories; and (6) quantitative analysis.

2.1 Sample

The study was conducted in one elementary school in the Lower Mainland of British Columbia. Ethics approval was obtained from the University of British Columbia and the school board of the district of that school. The school is located in an urban, low income environment. The median family income (ca. CAN\$ 30,000) in the neighbourhood surrounding the school is approximately one standard deviation (12,000) below the median family income of the entire province of BC (CAN\$ 43,000).

Seventy percent of parents of the children who returned the signed parental consent form provided consent; all of these students provided their assent. These 61 students in grades 4–7 provided think-aloud protocols. Because 6 of them had strong problems with the language due to English as a second language, these were excluded from the analysis; therefore, the total sample size for the analysis was 55. Two of the participants asked the interviewer to read out the items for them as they had reading problems. The students were from six classrooms: Two grade 4/5 classrooms, one grade 5/6 classroom, and three grade 6/7 classrooms. Fifty-eight percent of the students were girls. The mean age was 11.0 years (with a standard deviation of 1.2) ranging from 8.8 to 12.8 years. In terms of the grades, 16% of the children were in grade 4, 24% in grade 5, 29% in grade 6, and 31% in grade 7. The school has 350 students who speak more than 25 languages. In our sample, children reported having learned 20 different languages as their first language at home. Specifically, 31% of the children reported having learned English only, 33% reported having learned another language than English, and 36% reported having learned English and another language at home.¹ With regard to the children who had first learned another language than English only at home, the most frequently learned languages were Farsi, Chinese,² and Korean. With regard to children who had learned English and another language, the most frequently learned languages were Chinese, Punjabi, Spanish, French, and Korean.

With regard to how difficult it is for them to read and write in English, 53% of the children reported it as being very easy, 38% as easy, and 9% as hard.

2.2 Measure

2.2.1 *The Satisfaction with Life Scale Adapted for Children (SWLS-C)*

The SWLS-C was adapted from the Satisfaction with Life Scale (SWLS; Diener et al. 1985), a commonly used measure to assess satisfaction with life in adults. The SWLS was adapted for children by three subject matter experts in the area of socio-emotional development of children. The SWLS-C consists of five items addressing the respondents' life satisfaction with a 5-point response scale ranging from 'disagree a lot' to 'agree a lot' (see Table 1). Gadermann et al. (2010) provided psychometric evidence for the construct validity of the SWLS-C in a sample of 1,233 students in grades 4–7. Specifically, the results indicated that the scale was unidimensional, had a high reliability, and measured life satisfaction in the same way across different groups of children (namely, across gender, first language learned at home, and different grades) at the item and scale level as investigated by differential item and scale functioning analyses in that sample. Furthermore, the SWLS-C showed relationships to convergent and discriminant measures as was expected based on previous research.

2.3 Think-Aloud Protocols

According to Messick (1995), there are six aspects of construct validity, one of which is the substantive aspect. The substantive aspect of construct validity highlights the importance of identifying and modeling the processes that respondents employ in completing assessment tasks. Evidence for this construct validity aspect can be provided from different sources, and one is the think-aloud protocol (Ericsson and Simon 1980). In a think-aloud

¹ Two of these children reported having learned English and two other languages at home.

² The language label Chinese includes Mandarin, Cantonese, and Taiwanese.

Table 1 The Satisfaction with Life Scale adapted for Children. For each of the following statements, please circle the number that describes you the best. Please read each sentence carefully and answer honestly. Thank you

	Disagree a lot	Disagree a little	Don't agree or disagree	Agree a little	Agree a lot
1. In most ways my life is close to the way I would want it to be	1	2	3	4	5
2. The things in my life are excellent	1	2	3	4	5
3. I am happy with my life	1	2	3	4	5
4. So far I have gotten the important things I want in life	1	2	3	4	5
5. If I could live my life over, I would have it the same way	1	2	3	4	5

protocol, respondents are typically given the instruction to think aloud while completing a questionnaire, which is called concurrent verbalization. Also, respondents may be asked to describe previous cognitive processes, for example right after having finished a task, and this procedure is called retrospective verbalization (Ericsson and Simon 1980). In both of these think-aloud procedures, the researcher hardly interjects. A related approach is verbal probing, where respondents are probed for specific information by the interviewer; that is, the interviewer utilizes specific verbal probes, such as asking the respondent to reformulate an item, or to define some of the key terms in their own words (Willis et al. 1999). Oftentimes, researchers use a combination of the think-aloud and verbal probing approaches. In the present study we also used a combination of the concurrent verbalization with verbal probing.

2.4 Procedure

During class time, individual students were asked to come to a quiet room for the think-aloud protocol. The think-aloud protocols were audiotaped. Three practice items were used to familiarize the children with the task of thinking-aloud. Specifically, the first practice item was verbally presented to the children (adapted from Cremeens et al. 2006a): “When you are answering the items, I would like you to say out loud all the things that come into your head when you are choosing your answer. For example, I am answering a question about whether I am good at tidying my bedroom...Now what do I think? I don't like to tidy my bedroom, but I do tidy it when my mother tells me to...and I make sure that all my things are put away...so I think I am good at tidying my bedroom, and I point here (i.e., on the high end of the rating scale). Now we are going to answer some more questions, and I want you to remember to talk aloud, and say what you are thinking as you answer.” The children were then asked to respond themselves to this item. Then they were asked to respond to two more practice items “In general, I like to eat vegetables.” and “I enjoy reading books.” For each item, the children were asked to ‘think-aloud’ while they were considering their responses. If a child was silent for more than 10 s, s/he was given up to two prompts such as “Remember to say out loud all the things that come into your head” and “What are you thinking and saying to yourself now?” (Cremeens et al. 2006a, p. 85).

After the three practice items were completed, the children were asked to respond to the SWLS-C while thinking aloud. A subsample of 23 of the children was asked after the think-aloud protocol what they thought about the items and what they thought about giving

these items to other children their age. On average, the session took about $\frac{1}{2}$ h per child. At the end of the interview, the children were given two erasers.

2.5 Development of Coding Categories

All interviews were transcribed by a professional transcriptionist, and then checked for accuracy by the first author. Content analysis was used for deriving a coding scheme of the transcripts (Berg 2004) in order to decipher and interpret the data (Böhm 2004). Content analysis is “the systematic, objective, quantitative analysis of message characteristics” (Neuendorf 2002, p. 1).

The children’s responses were coded for each of the five items of the SWLS-C separately, using the software Atlas.ti 5.0. First, open coding was used for a wide inquiry into the data. After the open coding, codes were assigned to categories. The categories were developed to (1) reflect the research purpose, (2) be exhaustive and mutually exclusive, (3) be grounded conceptually in the theoretical quality of life research literature, and (4) be grounded empirically in the data (see Dey 1993; Holsti 1969). Accordingly, the coding combined an inductive and deductive approach, with a larger reliance on the inductive approach.

Themes were chosen as the unit of analysis for the coding. Themes in its simplest form can be a “simple sentence, a string of words with a subject and a predicate” (Berg 2004, p. 273). Children’s responses were coded according to themes, which were then assigned to the accordant category. Generally, one primary theme was coded for each response to a particular item. Children frequently provided more than one theme in response to a single item (without one being primary). In those cases, the different themes were coded into separate categories. Eventually, for a category to be kept in the overall final coding scheme, a category had to occur at least three times in any of the five items of the SWLS-C (cf. Berg 2004).

The development of our coding categories was guided by three general research purposes. The first purpose was to investigate how the children understand, interpret, and respond to the items of the SWLS-C. Specifically, we were interested in the strategies that children employed to respond to these items. A second purpose was to identify the content the children talked about; in other words, the aim was to investigate on which content children focused when using a certain strategy for making their life satisfaction judgments. A third research purpose was to find out whether children use positive or negative statements in their responses. According to these purposes, the process of developing the coding categories was informed by the following specific questions:

1. What strategies do the children employ when responding to the item?
2. What are the general content topics that come up for the children when responding to the items?
3. Is the valence of these content topics typically positive (i.e., presence of something positive or absence of something negative) or negative (i.e., presence of something negative or absence of something positive)?

Furthermore, it was of interest to examine whether children had any difficulties in terms of understanding the SWLS-C items and/or the response format.

After preliminary categories were developed, the most prominent categories were further elaborated, and a coding scheme was developed with main categories and sub-categories. Based on the coding scheme, the first author went through the transcripts again to check the codes and recode the data. After that, a second rater, who is a researcher in the

area of child development, coded the data based on the coding scheme, and the inter-rater reliability was computed. In case there was a difference in coding, the two raters discussed the code and came to an agreement.

2.6 Quantitative Analysis

The frequencies of the codes were then transferred into SPSS 17.0 for further analysis. Specifically, the frequencies of the use of different categories were investigated. Furthermore, we were interested in whether children used certain strategies significantly more often than another for the respective items. This was investigated by calculating the McNemar test, a paired test of equality of proportions, for each item separately. Moreover, we were interested in whether there were differences in the use of strategies depending on the children's age. In order to investigate this, the sample was divided into two groups: A younger group (with the children in grades 4 and 5) and an older group (with children in grades 6 and 7). The McNemar test was then run separately for the age/grade categories.

In addition, in order to detect potential demographic differences with regard to the use of the strategies, we ran Poisson or binary logistic regression analyses (depending on whether the response variable was a count or binary variable, respectively) with the factors gender, grade, and first language background.

Finally, one additional set of analyses was conducted to examine whether the valence (positive or negative) of children's responses was correlated with their scores on the individual SWLS-C items.

3 Results

3.1 Definitions of Categories and Subcategories

All of the participants' responses to the five items were coded into four levels of categories, according to specific coding definitions. In the following, we present a definition of each category. The coding categories are represented by the tree diagram in Fig. 1.

3.1.1 Strategy Categories

As a first step, for each response, it was examined whether the participant used an *absolute* (A) or a *relative* (R) strategy in her/his response thought process. That is, it was examined whether a participant used absolute or relative statements while responding to an item. In this context, an absolute statement indicated the presence or absence of something that was apparently important with regard to a participant's judgment of her/his satisfaction with life (e.g., "I agree a lot with that cause I have very nice parents and a really nice sister."). A relative statement, on the other hand, included a comparative statement with regard to the presence or absence of something that was important with regard to a participant's judgment of her/his satisfaction with life (e.g., "I wish I would get *better* grades."; "I want to have *more* friends."). In addition to these two categories, a category labelled *General positive* was defined to capture all statements that did not include an absolute or relative strategy, but that included a general, positive statement (e.g., "Because it's fun, and I just like it."; "Well, it's not boring, it's kind of fun.") Finally, any responses that could not be coded into any of these three categories were assigned to a category labelled *unclear*.

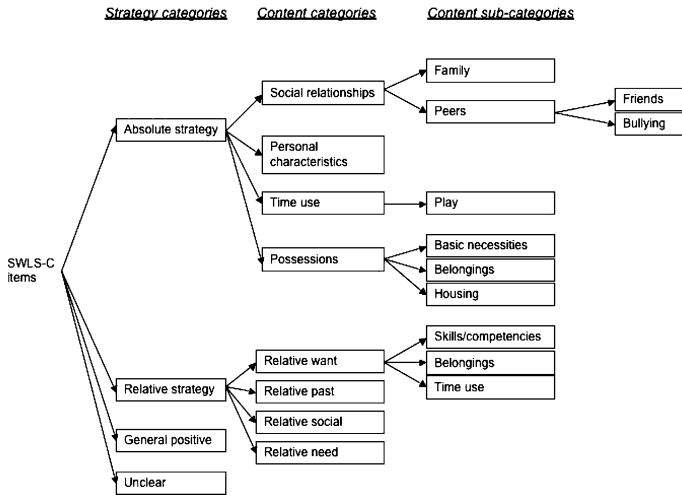


Fig. 1 Tree diagram of the coding categories

(e.g., “It’s because mostly sometimes it happens and sometimes it doesn’t.”; “I don’t exactly know what I want in life”; “It’s just a life. I just live on a daily basis or something.”).

3.1.2 Content Categories

In a second step, two clusters of *content categories* were developed. The first cluster of content categories was assigned to the absolute strategy category, and the second cluster of content categories to the relative strategy category. For the absolute strategy category, four content categories were defined: (A1) *Social relationships*; (A2) *Time use*; (A3) *Personal characteristics*; and (A4) *Possessions*. Similarly, for the relative strategy, the following four categories were defined: (R1) *Relative social (social comparisons)*, (R2) *Relative to one’s wants*; (R3) *Relative to one’s past*; and (R4) *Relative to one’s needs*.

For the four content categories under the absolute strategy category (see Fig. 1), the following definitions were developed. (A1) *Social relationships*: Each response that used an absolute statement referring to a social relationship was coded in this category. In order to capture the diversity of social relationships that were mentioned, this category was subdivided into two subcategories, according to whether the statement referred to *Family* members, such as parents, siblings, or grandparents (e.g., “I am happy with my life because I have a really caring family.... And like my mom is always there for me whenever I’m sad or there’s something that I really want to tell her.” or “My dad gets mad at me for no reason. And he swears a lot.... My dad keeps getting mad because he was on drugs and stuff.”); or *Peers*. The category *Peers* was further subdivided into the subcategories *Friends* (“My friends, they’re very supportive of me and they’re wonderful.”) and *Bullying* (“Because people make fun of me and call me names. Like Big Apple because they think I’m fat. They bully me a lot, like start punching me and kicking me.”). (A2) *Time use*: Any statement referring to an activity was coded into this category (e.g., “I like going shopping on Saturdays.”). The subcategory *Play* was created, which included statements referring to games and play activities (e.g., “Because most of the time I like to play.”). (A3) *Personal*

characteristics: This category included statements referring to personal characteristics, competences, skills, and likes (e.g., “Cause I am not doing that good in school and stuff.”; “I’ve got these good talents in singing and a lot of knowledge, too.”). (A4) *Possessions*: Any references to personal belongings, possessions, or access (or lack of) material things were coded into this category, according to the following subcategories: *Basic necessities* (this subcategory included things that fulfill basic needs, e.g., “We’re not living in poverty and that’s also important. And I have shelter and all the other basics, like, water.”); *Belongings* (this category included material things, such as computer games, as well as pets; e.g., “Since I have lots of Lego.”; “Because I have a cat.”); and *Housing* (statements in this category did not indicate the presence of shelter as a necessity, but referred to the quality of the housing situation; e.g., “Because we have a nice house.”).

The four content categories within the relative strategy category (see Fig. 1) were coded according to the following definitions: (R1) *Relative social*: This category included statements that made social comparisons (e.g., “I have a younger brother, so my parents like him more. And they care for him more.”; “We should realize how it will be to be like other people with struggles in other parts of the world.”). (R2) *Relative to one’s wants*: Any statement indicating a comparison between what a child currently has and what s/he wants was coded in this category. This category was subdivided into three subcategories: *Time use* (e.g., “Because I’ve always wanted to start karate and right now I’m starting my monthly karate.”); *Belongings* (e.g., “We live in an apartment, but I want to live in a house.”; “I agree a little cause I want to get a Wii and a bigger house and a car.”); and *Skills/competencies* (e.g., “I want better grades..., but I don’t usually get good grades, but they’re okay.”; “Because I always wanted to be a good drawer and now I’m a really good drawer.”). (R3) *Relative to one’s past*: If a child made a comparative reference to her/his past, the response was coded into this category (e.g., “Back at the old place, they [my cousins] teased me, but here they don’t tease me.”; “Because I have the things I wanted to happen. Oh, like, one night, when we were in Afghanistan, so when there was a fight, there was a war with Taliban, so we wanted a good life. So I wished that we could go to another country or somewhere else. Or maybe Canada or America. First we went to Uzbekistan, then we lived there for 8 or 9 years and then we came here.”). (R4) *Relative to one’s needs*: Any statements that made a comparison between a child’s status quo and his/her (stated) needs was included in this category (e.g., “I don’t have the best life, I don’t think. But I have one that suits me. I have everything I need right now”).

3.2 Category Frequency Counts

In the Sect. 2, it was described that the criterion for creating (and keeping) a category was that at least three participants used statements referring to this category in response to an item. It must be noted that categories that occurred three or more times for one item (e.g., item 1) were then also used for the coding of the other items (e.g., 2–5). This led to the scenario that, for some of the items, a category was used by less than three participants. In those cases, the codes for these statements were assigned to the next higher level category (e.g., if statements in reference to the *peers* category occurred only twice in response to item 2, these statements were counted towards the higher level category *social relationships*).

In Fig. 1, it should be noted that all categories that occurred for at least one item are shown. However, when we report the results for individual items and present them graphically in the respective items’ tree diagrams (Figs. 2, 3, 4, 5 and 6), only those categories that were used at least three times for that particular item are reported and are shown as individual boxes in the diagram. In those cases, any categories with less than

three codes were collapsed into a general category labelled as *other*, and the count of statements within this general category was simply counted towards the higher level category. For example, if, on item 3, only one participant referred to *basic necessities* and one participant referred to *housing* [in the tree diagram, see: *absolute* (strategy category) → *possessions* (content category) → *belongings* and *housing* (subcategory)], these two codes were collapsed, represented in a box labelled *other*, and counted towards the next higher level category, *possessions*. It must also be noted that there were a few responses, which were coded and assigned to a category, even if the codes did not fit into any of the category’s subcategories—because they occurred less than three times on all five items. For example, a child referred to the relationship with her/his teacher. This reference was, naturally, coded under the category *social relationships*, but the code could not be assigned to any of the subcategories of *social relationships* (that is, *family* or *peers*). In the frequencies reported in our Sect. 3 as well as in the graphic representations of our results, such codes also appear under the generic *other* category or subcategory. This procedure allowed us to maintain the highest level of detail in the descriptive reporting of the data, and allowed us to conduct the statistical analyses of the category frequencies at a methodologically adequate level.

As mentioned above, the objective was to develop categories that reflect the research purpose, and that are exhaustive and mutually exclusive. These requirements were met except for being exhaustive. Specifically, we had one category that was entitled *unclear*; however, as Holsti (1969) points out “even the most carefully designed study is likely to fall short of completely satisfying this requirement” (p. 99).

The inter-rater reliability using Cohen’s Kappa was .84. According to Mayring (2004), Kappa coefficients of .70 are considered to be sufficient.

3.2.1 Item 1: In Most Ways My Life is Close to the Way I Would Want it to be

For item 1, the absolute strategy was used 46 times and the relative strategy 33 times. Furthermore, three responses were coded in the general positive category, and three responses were coded in the unclear category. The frequencies of the strategy and content

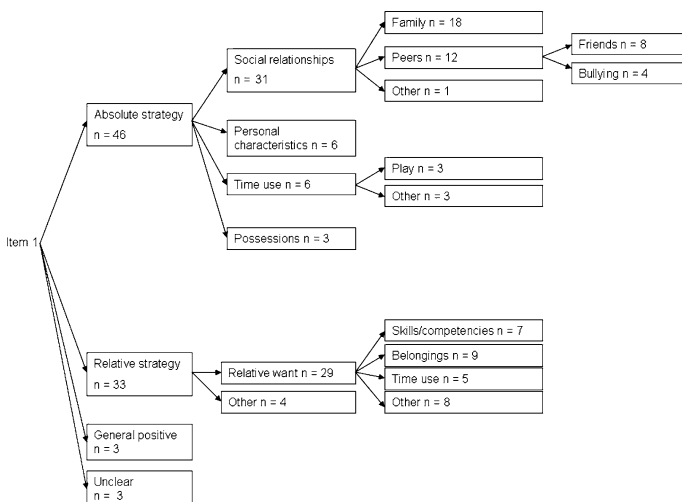


Fig. 2 Tree diagram of the coding categories for item 1

(sub)categories for item 1 are illustrated in Fig. 2. As mentioned above, children frequently used more than one strategy and/or content category in response to one item.³ Therefore, we also report how many children used the relative and absolute strategies and/or content categories: Of the 55 children, 7 children used both the relative and absolute strategy in responding to item 1, 20 children used only the absolute strategy, and 22 children used only the relative strategy in order to respond to the item. Of the children who used the absolute strategy, 11 used one content category, 14 mentioned two content categories, 1 child mentioned three, and 1 child mentioned four different content categories. Of the children who used the relative strategy, 25 mentioned one content category and 4 mentioned two content categories. Seven children had some difficulty responding to the item as they found the item somewhat difficult to understand. With regard to the valence of the statements, 70% of the statements were positive, 22% negative, and 8% were mixed.

3.2.2 Item 2: *The Things in My Life are Excellent*

For item 2, the absolute strategy was used 69 times and the relative strategy was used 15 times. Furthermore, three responses were coded as unclear. The frequencies of the strategy and content (sub)categories for item 2 are illustrated in Fig. 3. Two children used both the relative and absolute strategy, 38 used only the absolute strategy, and 12 children used only the relative strategy. Among those who used the absolute strategy, 20 mentioned one content category, 14 mentioned two content categories, 3 mentioned three content categories, and 3 mentioned four content categories. For the relative strategy, 13 children mentioned one content category, and 1 child mentioned two. With regard to item understanding, 1 child asked to what the word “things” was referring, and 1 child commented that he felt the wording was grammatically incorrect. With respect to the valence of the statements, they were predominantly positive (72%), and only relatively few were negative (21%) or mixed (7%).

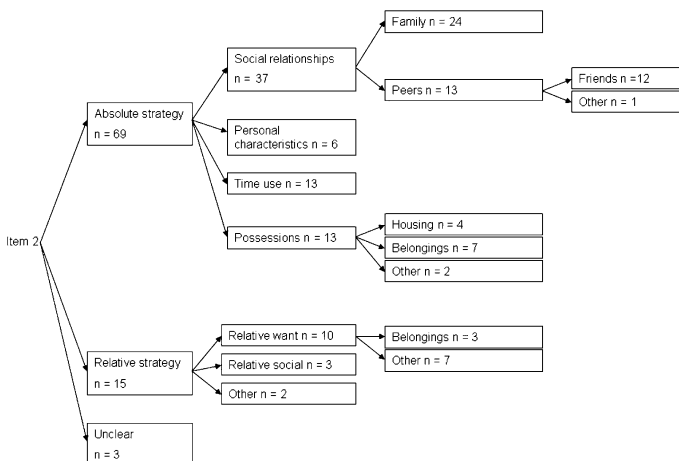


Fig. 3 Tree diagram of the coding categories for item 2

³ As a result, the sum of all frequencies of the categories is not equal to the sample size of 55.

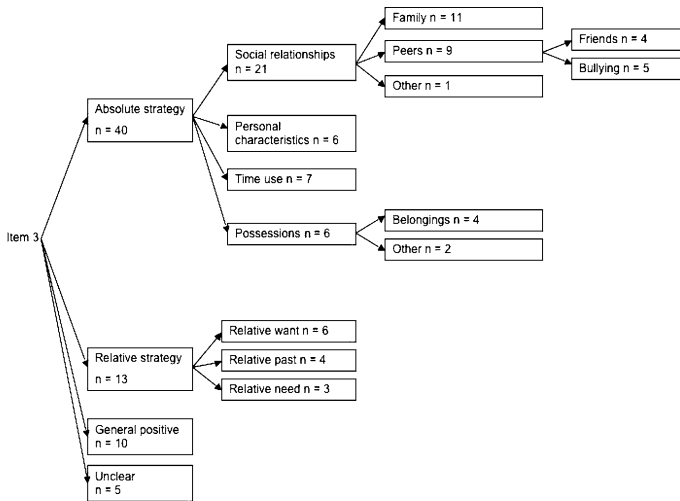


Fig. 4 Tree diagram of the coding categories for item 3

3.2.3 Item 3: *I Am Happy with My Life*

For item 3, the absolute strategy was used 40 times and the relative strategy was used 13 times. Furthermore, ten responses were coded in the general positive category, and five responses were coded in the unclear category. The frequencies of the strategy and content (sub)categories for item 3 are illustrated in Fig. 4. One child did not provide any explanation for his response. One child used both the absolute and relative strategies, 26 used only the absolute strategy, and 12 used only the relative strategy. For the absolute strategy, 19 children mentioned one content category, 5 children mentioned two content categories, 1 mentioned three, and 2 mentioned four content categories. For the relative strategy, all mentioned one content category. Two children commented that the item was similar to the previous ones and one child had problems with the response format to respond to the item. The valence of the statements was mostly positive (80%) with a few negative (11%) and mixed (9%) ones.

3.2.4 Item 4: *So Far I Have Gotten the Important Things I Want in Life*

For item 4, the absolute strategy was used 40 times and the relative strategy was used 47 times. Furthermore, three responses were coded as unclear. The frequencies of the strategy and content (sub)categories for item 4 are illustrated in Fig. 5. Eleven children used both the relative and absolute strategy, 12 only the absolute strategy, and 29 only the relative strategy. For the absolute strategy, 13 children mentioned one content category, 6 mentioned two content categories, 1 mentioned three content categories, and 3 mentioned four content categories. For the relative strategy, 34 children mentioned one content category, 5 mentioned two content categories, and 1 child mentioned three content categories. One child had problems responding to the item. With respect to the valence of the statements, they were predominantly positive (75%), and only relatively few were negative (15%) or mixed (10%).

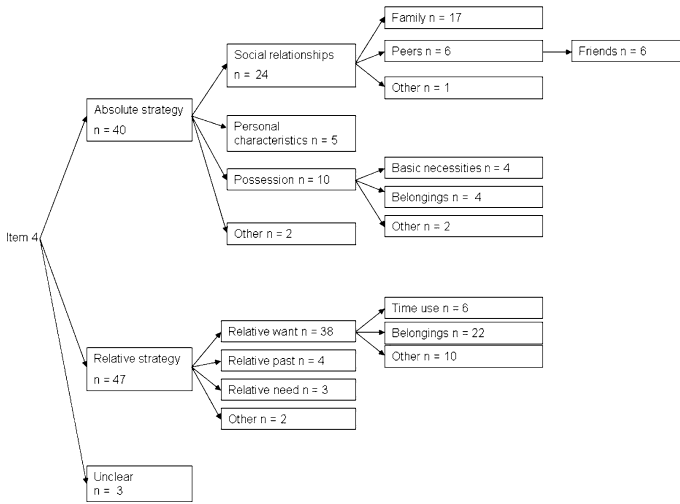
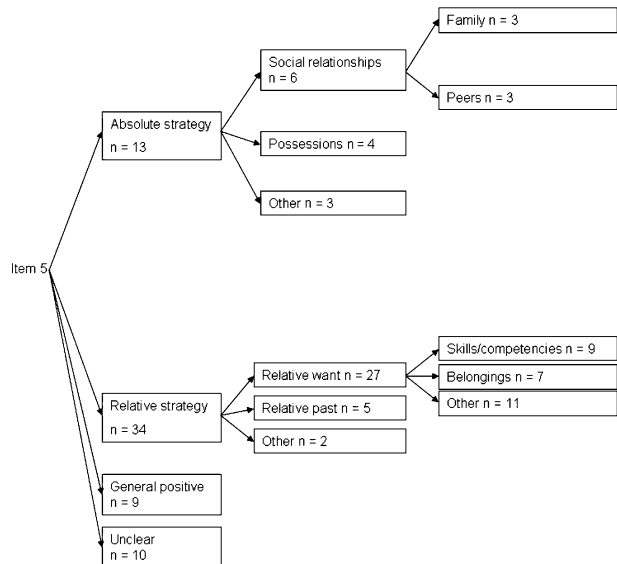


Fig. 5 Tree diagram of the coding categories for item 4

3.2.5 Item 5: *If I Could Live My Life Over, I Would Have it the Same Way*

For item 5, the absolute strategy was used 13 times and the relative strategy was used 34 times. Furthermore, nine responses were coded in the general positive and ten responses in the unclear category. The frequencies of the strategy and content (sub)categories for item 5 are illustrated in Fig. 6. Three children used both the relative and absolute strategy, 7 only the absolute strategy, and 26 only the relative strategy. For the absolute strategy, 7 children mentioned one content category, and 3 mentioned two content categories. For the relative strategy, 24 children mentioned one content category, and 5 mentioned two content

Fig. 6 Tree diagram of the coding categories for item 5



categories. Furthermore, 5 children had problems responding to the item. With respect to the valence of the statements, 52% were positive, 40% were negative, and 8% were mixed.

3.3 Comparison of the Use of Strategies and Content Categories Across Items

There are several ways to look at the patterns of our findings. One way is to compare whether the tree diagrams—that is, the occurrence of categories and subcategories—are similar or different across items. This information is summarized graphically in Fig. 7.

As can be seen, for all items, children used the absolute and relative strategies. Furthermore, the category ‘unclear’ was assigned to responses for all five items. However, comments assigned to the category ‘general positive’ only occurred (three or more times) for items 1, 3, and 5. Among the content categories, ‘social relationships’, ‘possessions’, and ‘relative want’ occur for all five items. The only subcategories that occur for all five items are the social relationship subcategories ‘peers’ and ‘family’. All other (sub)categories occurred for a subset or only one of the items. For example, the content category ‘personal characteristics’ occurred for items 1–4, whereas the content category ‘relative social’ only occurred for item 2.

A further way to explore the patterns of results is to visualize the frequencies with which the different strategies, categories, and subcategories were used across the items. Figure 8 presents the frequencies for the absolute and relative strategies, the content categories, and the most frequently used subcategories.

In the top panel of Fig. 8, it can be seen how often children used absolute versus relative/comparative strategies, for each item. In the middle panel of the figure, it can be

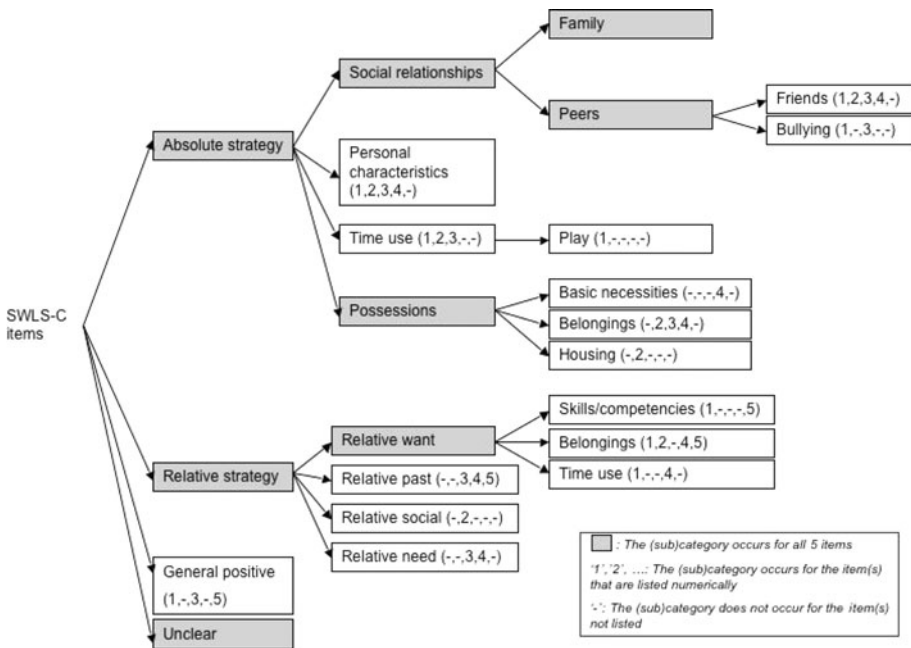


Fig. 7 Summary of the differences and commonalities in strategy and content (sub)categories for the five items of the SWLS-C

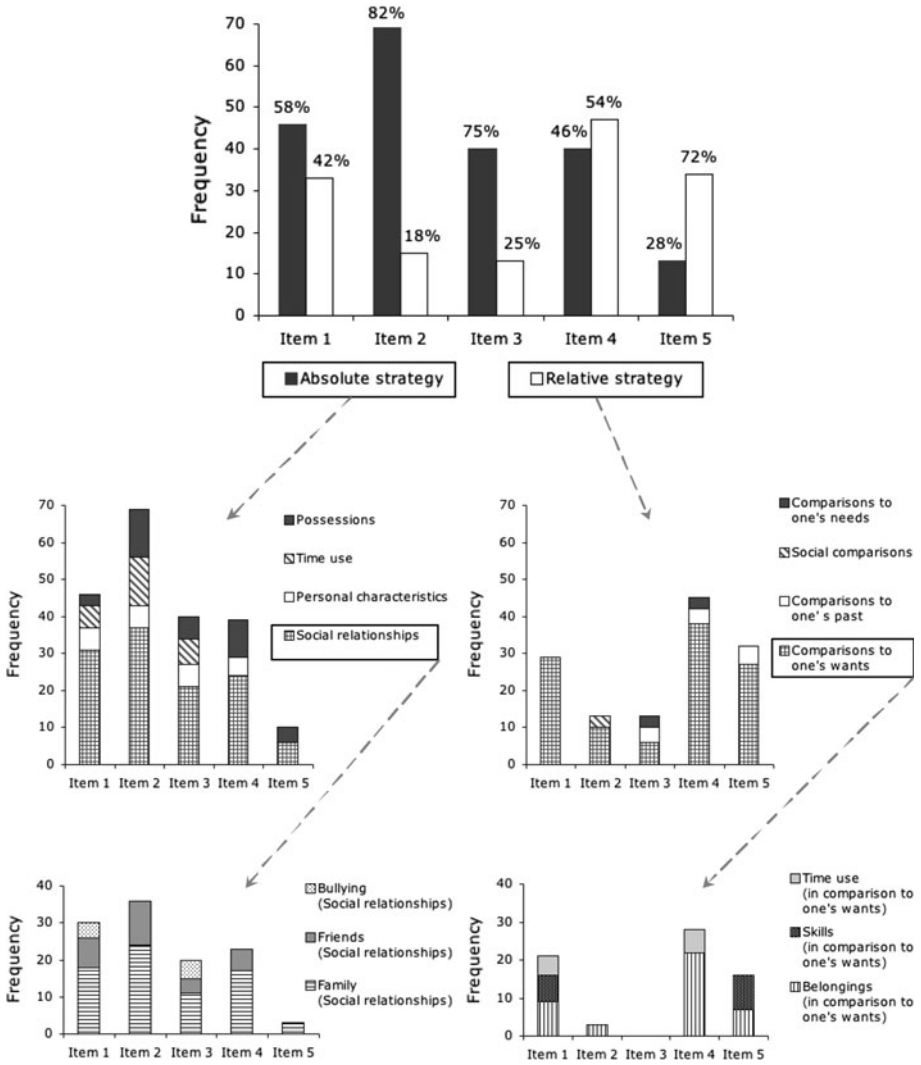


Fig. 8 Frequencies for the five SWLS-C items for *absolute* versus *relative* strategies (top panel), the content categories for absolute strategies (middle panel, left) and relative strategies (middle panel, right), and the subcategories for *social relationships* (bottom panel, left) and for *comparisons to one's wants* (bottom panel, right)

seen how often the different content categories of the absolute strategy (left) and the relative strategy (right) were used. As can be seen, within each of the five items, the content category 'social relationships' was used most frequently in the absolute strategy. In the relative strategy, the content category 'comparison to one's wants' was used most frequently for each of the five items. In the bottom panel of the figure, the subcategories for the 'social relationship' content category (left) and the subcategories for the 'comparison to one's wants' content category (right) are displayed, showing how often each of the respective subcategories occurred. (Note: If the total numbers of the bars in the middle and

lower figures do not correspond to their respectively corresponding bars in the figure(s) one level above, it is because the codes that fell under ‘other’ are left out of these figures.)

Figure 8 illustrates several interesting patterns. First of all, the absolute strategy is used more frequently for items 1, 2, and 3, but the relative strategy is used more frequently for items 4 and 5. The difference in the use of strategies is most pronounced for items 2, 3, and 5. With regard to the content categories for the absolute strategy, ‘social relationships’ were mentioned most frequently in the children’s responses for each of the five items. The content category ‘time use’ only occurred (three or more times) in items 1, 2, and 3; which are the three items that do not make a reference to a time frame (Item 4: ‘*So far* I have gotten ...’; Item 5: ‘If I *could live my live over*, I would ...’). The content category ‘possessions’ occurred most frequently for items 2 and 4, both of which contain the word ‘things’ in it (Item 2: ‘The *things* in my life are excellent.’; Item 4: ‘So far I have gotten the important *things* I want in life.’). For the relative strategy, the content category ‘comparisons to one’s wants’ occurred most frequently for all five items. Comparisons to one’s past were made for items 3, 4, and 5—with items 4 and 5 being the two items that make an explicit reference to a time frame.

Within the content category ‘social relationships’ (bottom panel, left), it can be seen that children most frequently made references to their ‘family’ in their responses, and that ‘friends’ were mentioned with the second-highest frequency. The subcategory ‘bullying’ solely occurred for items 1 and 3. With respect to the content category ‘comparisons to one’s wants’, it can be seen that ‘belongings’ were most frequently mentioned by children in response to item 4, which makes reference to the past (‘*So far* ...’), and to ‘things’.

3.4 Comparison of the Use of Strategies Within Items

We compared the use of the relative versus the absolute strategy overall and separately for the two age groups for each item. For the comparison, each child received a binary code (0 or 1) depending on whether s/he used the absolute or relative strategy or not, and then the McNemar test, a paired test of equality of proportions, was calculated. With regard to item 1, we did not find significant differences overall or between the age groups.⁴ Our results for item 2 indicate that there was a statistically significant difference in the use of the absolute versus the relative strategy overall, with the absolute strategy being used more often ($\chi^2(1) = 12.50$; $p = .0001$; OR = 3.0). This difference was only statistically significant in the younger age group; i.e., the younger children used the absolute strategy statistically significantly more often than the relative strategy (exact significance $p = .001$; OR = 6.0). With regard to item 3, there was a statistically significant difference in the overall use of the absolute versus the relative strategy, with the absolute one being used more frequently ($\chi^2(1) = 4.45$; $p = .04$; OR = 2.1), but there were no statistically significant differences within the age groups. With regard to item 4, there was a statistically significant difference in the overall use of the strategies, with the relative one being used more frequently ($\chi^2(1) = 7.23$; $p = .007$; OR = 2.7); this difference was only statistically significant within the older grade category ($p = .04$; OR = 2.9). With regard to item 5, there was a statistically significant difference in the overall use of the strategies, with the relative one being used more frequently ($\chi^2(1) = 9.82$; $p = .002$; OR = 3.6); this difference was statistically significant only within the older group ($p = .0001$; OR = 9.2).

⁴ According to Fig. 8, it appears as if there should be a statistically significant difference in favour of the absolute strategy. Please note that in contrast to the data in Fig. 8, the data for the McNemar test were recoded into binary data, which explains why the test was statistically non-significant.

3.5 Relationship to Demographic Variables

In the next step, we were interested in investigating whether there are statistically significant differences with regard to demographic variables when using the absolute or relative strategies. Therefore, we ran Poisson or binary logistic regression analyses (depending on whether the data were counts or binary) with the factors gender, grade, and first language background. The results indicate that for the relative strategy there was only one statistically significant result, namely for gender on item 3. Specifically, girls used the relative strategy significantly more often than boys (Wald $\chi^2(1) = 4.24$; $p = .04$; OR = 5.5). With regard to the absolute strategy there was also only one statistically significant result. Specifically, girls used the absolute strategy significantly more often than boys on item 4 (Wald $\chi^2(1) = 6.67$; $p = .01$; expected rate for girls = .97; expected rate for boys = .35).

3.6 Correlations Between Valence of the Responses and SWLS-C Item Scores

In order to calculate the correlations between the valence of children's responses and the SWLS-C scores the positive statements were coded as +1, the negative statements as -1, and the neutral statements as 0, and a sum score was calculated for each item. This score was then correlated with the children's respective item scores. For all five items, statistically significant ($p \leq .001$), positive Spearman rank correlations were found (the SWLS-C item mean and standard deviations (SD) are provided in parentheses): Item 1: $r = .56$ (mean = 4.1; SD = .92); Item 2: $r = .43$ (mean = 4.2; SD = .83); Item 3: $r = .44$ (mean = 4.6; SD = .63); Item 4: $r = .48$ (mean = 4.4; SD = .87); Item 5: $r = .66$ (mean = 4.0; SD = 1.28).

3.7 Feedback on the Items

Of the 23 children who were asked for their feedback on the items, 22 responded that they thought it is important to give these items to children and that they enjoyed responding to them. Specifically, several children said that it was a good way to find out how children their age are feeling, for example: "So you can know how they're feeling in life and—like how they're feeling with their families, friends, teachers and stuff like that."; "Because it's easier then to understand how at our age people think. And what's happening at home and their life, if they're stressed out or not."; "I think you should know what's going on in their heads, because a lot of kids have problems. And they don't talk about it. So, you need to know this stuff."; "It really helps them just [to] get their feelings out. Instead of holding all their feelings inside."

Furthermore, several children said that this would be a good way to get information that would be important to help children, for example: "Because if you wanted to change something and if most people say it, then you could change it."; "So then people can help us more."

In addition, several children mentioned that they enjoyed answering the items, for example: "It's good, because I never even thought about these questions before in my life."; "Because you're asking them what they like the most. And what they do or they don't like the most. So they're encouraging."; "Because then you can think of your life a bit. And see that maybe you made a mistake in your life and then you said it in here, realizing that you did make a mistake, so that you can fix the mistake over in your life if it ever happens again."

One child also mentioned that it would be good to give this scale to older students in high school “because... they have too much homework. They’re stressing out and stuff. They have lots of problems in their life.”

4 Discussion

The purpose of the present study was to investigate the cognitive processes of children when responding to the items of the SWLS-C to provide evidence for the substantive aspect of construct validity. Our study showed that children used two main strategies to answer the items on life satisfaction, namely an absolute strategy and a relative or comparative strategy. In the former, children referred to the presence or absence of something that was of relevance for their satisfaction with life. In the latter, children made comparisons of their current state to what they want, what others have, what they had in the past, and what they need to rate their life satisfaction. Our findings are in line with the *multiple discrepancies theory* (MDT; Michalos 1985) in several regards. MDT makes several propositions about the processes used by individuals to make judgments on their life satisfaction and domain satisfaction. The first proposition of MDT postulates that reported net satisfaction is a function of perceived discrepancies between what an individual currently has compared to (1) what s/he wants (‘self-want’), (2) what relevant others have (‘self-others’), (3) the best s/he has had in the past (‘self-best past’), (4) what s/he expected to have 3 years ago at this point in life (‘self-progress’), (5) what s/he expects to have after 5 years (‘self-future’), (6) what s/he deserves (‘self-deserves’), and (7) what s/he needs (‘self-needs’).

The MDT also proposes that the discrepancy between what an individual currently has and what s/he wants is a mediating variable between the other discrepancies and life satisfaction (Michalos 1985, pp. 347–348).⁵ Even though the mediation could not be tested with our data, it is of interest to note that the children used the self-want comparison with the highest frequency. This finding suggests that children assign a particular importance to the self-want category in their judgement of life satisfaction. Furthermore, our findings show parallels to findings from previous studies that tested the MDT with university students (Michalos 1985, 1991). In particular, Michalos (1985) tested how successfully the MDT could be used to predict/explain life satisfaction in a Canadian undergraduate student sample. In that study, the discrepancies that were most salient with regard to predicting/explaining variance in the students’ life satisfaction ratings were—in order—self-want, self-others, self-needs, self-best past, self-deserved, self-progress, and self-future. Similarly, in a study that investigated the relative importance of the discrepancies with regard to the prediction of life satisfaction in a large sample of undergraduates from 39 countries, the self-wants and the self-others discrepancies had the largest impact (Michalos 1991). Our findings of the relative strategy show that children in grades 4–7 use some of the same discrepancies to make evaluations of their satisfaction with life when responding to the items of the SWLS-C. Particularly, the four discrepancies that were used by the children are the ones that were most successfully predicting life satisfaction in those previous studies, namely the self-want, self-past, self-need, and self-other discrepancies (ordered according to frequency of occurrence in children’s responses). It needs to be pointed out that the self-past discrepancy was used differently by the children than it is conceptualized in MDT. In MDT it is the discrepancy between what one currently has and the best one has

⁵ For the other propositions, please see Michalos (1985).

had in the past. In contrast, the children were mostly making comparisons with the past, where their lives or a specific occurrence in the past was considered to be negative, and they were commenting on the improvement in their lives since then.

Furthermore, our findings show some similarities with Cremeens et al. (2006a) findings, regardless of the fact that the items of the measure in Cremeens et al.'s study and the SWLS-C are quite different. The items of the TedQL used by Cremeens et al. are quite specific (addressing abilities, such as children's reading ability, or social aspects, such as having friends at school), whereas the items of the SWLS-C are more general (pertaining to overall evaluations of children's lives). Also, the children in the study by Cremeens et al. were younger than the ones in our study (mean age of 7.1 vs. 11.0 years). These differences notwithstanding, there is some overlap in the strategies that children used in responding to the respective measures. Specifically, Cremeens et al. report that children used social comparisons for answering the items, which we also found for item 2. Furthermore, they report that children used stable character references, which in the present study was coded under the absolute strategy and the content category personal characteristics and was used for items 1–4. In addition, Cremeens et al. report on children using concrete examples as a strategy, which we also saw in children's responses to the SWLS-C, but which was not coded as a strategy in itself, as it occurred within the different strategies when children used concrete examples for illustrative purposes. Lastly, they report on other reasons or no reason given, which is similar to the strategies we termed general positive strategy and unclear strategy.

In regard to the demographic variables gender, first language background, and grade, our regression analyses with the relative and absolute strategies as dependent variables did not show any systematic patterns across the five items of the SWLS-C, but it would be of interest to investigate these relationships in future studies with a larger sample size, and a larger range of age/grades.

In a separate set of analyses, we examined whether children's use of the absolute or relative strategy was associated with the (wording of the) items. Our results indicate that the relative strategy was used more frequently than the absolute one when children responded to the two items that make reference to the past (items 4—'So far ...'—and item 5—'If I could live my life over, ...'), whereas the absolute strategy was used more frequently for the two items that make reference to the present (items 2 and 3). (There were no statistically significant differences in the use of strategies for item 1). When looking at the response strategies children used within the respective age groups of younger (grade 4 and 5) and older (grade 6 and 7) children, we found that older children used the relative strategy significantly more often than the absolute strategy for items 4 and 5, whereas there was no such difference for the younger children. For item 2, the younger children were more likely to use the absolute than the relative strategy, whereas there was no difference for the older children. It goes beyond the scope of this study to speculate about the reasons for this. It might be the case that the reference to the past in items 4 and 5 is more likely to elicit a relative strategy rather than an absolute strategy in children, and particularly for older children. It would be of interest to examine in future studies with a larger sample and age range whether age-related cognitive development is associated with specific response strategies in response to the SWLS-C items. In fact, an age-related pattern could be expected based on developmental theories that propose that children's understanding of self becomes increasingly specific during middle childhood because their cognitive skills become more complex (Stone and Lemanek 1990; De Civita et al. 2005), and because their self-descriptions become more comparative (Bee 1989).

With regard to the content topics that came up for the children when responding to the items, the content category of the absolute strategy that was used most frequently was 'social relationships', which was used by the children for all five items, with social relationships with family members being especially prominent. This indicates the importance of social relationships, especially with family members, for children's life satisfaction, which is in line with previous empirical research (Huebner et al. 2004). Huebner (1991) reports that the strongest association between global life satisfaction and domain satisfaction ratings was with the domain family, and the relationship to the domain peers was also significant for children in grades 5–7. Similarly, Man (1991) found that parent orientation had a stronger relationship to life satisfaction than peer orientation with adolescents. In the present study, children often mentioned parental support during the think-aloud procedure. Young et al. (1995) also report that perceived parental support was positively correlated with adolescents' ratings of life satisfaction.

For the relative strategy, the content category that was used most frequently was 'comparisons to one's want', which was utilized by the children in responding to all five items (this was also the discrepancy with the highest success rate in the test of MDT). Within the self-want category, children most frequently referred to belongings. Similarly, 'possessions' was a content category of the absolute strategy that was frequently used. The school in which the study was conducted is located in a neighborhood with relatively low socio-economic status, and several children said that their families do not have enough money to buy them certain things. At the same time, many children were also commenting on the things they (or their family) owned and which they considered important. Furthermore, several children were comparing the things they owned with the ones they would have liked to have had, typically arriving at a positive judgment of their life satisfaction. Empirical findings on the relationship between socio-economic status and life satisfaction for children and adolescents have been ambiguous, with some studies reporting a statistical significant association of moderate effect size (e.g., Dew and Huebner 1994), and other studies reporting a statistically non-significant correlation of negligible effect size (e.g., Huebner 1991). It would be of interest to investigate whether children from a different socio-economic background also mention belongings or possessions as frequently in think-aloud protocols.

With regard to the valence of the children's statements, children predominantly talked about positive experiences and aspects of their lives (i.e., the presence of something positive or the absence of something negative). In fact, 70% of the statements were of positive valence across the items. In contrast, 22% of the statements were of negative valence (i.e., the presence of something negative or the absence of something positive) and 8% were of mixed valence (both positive and negative). This suggests that (most) children are predominantly thinking about positive experiences and aspects of their lives when making judgments on their life satisfaction. In addition, we found that the valence of children's responses to the items was positively related to the respective item scores. Also, the item mean scores were all equal to or above 4.0, indicating that the children in our sample, on average, rated their life satisfaction as positive. These findings are in line with previous empirical research that has shown that most children and adolescents rated their life satisfaction positively (e.g., Gadermann et al. 2010; Greenspoon and Saklofske 1997; Huebner and Alderman 1993; Huebner et al. 2000).

The majority of the children did not have any difficulties with the item content or the response format of the SWLS-C. However, several children found it difficult to respond to items 1 and 5 (7 and 5 children, respectively). These children were slightly younger than the overall sample (mean age of 10.1 years) and, with the exception of two children, all

bilingual. We are hesitant to recommend any changes to the item wording based on the children's feedback as this was quite diverse. Furthermore, these items were performing well in previous pilot studies with focus groups of children and in a psychometric analysis with a larger sample (Gadermann et al. 2010). However, we recommend that future studies validating the SWLS-C should have a special focus on these two items.

In a previous study, the SWLS-C showed favourable psychometric properties in terms of reliability, factor structure, differential item and scale functioning, and correlations to convergent and discriminant measures (Gadermann et al. 2010). The aim of the present study was to add to the validity evidence by evaluating the substantive aspect of construct validity. Toward this end, this study provides insights into the strategies that children used to respond to the SWLS-C. We gained an understanding that children's item responses are governed by strategies that are meaningful and reflect ideas in the quality of life literature. Specifically, the strategies and content of the children's responses were theoretically in line with MDT and also converge with previous empirical findings in the quality of life literature with children and adolescents. Additionally, our results indicate that the majority of the children did not have any difficulties in understanding the items. From a practical, applied perspective, it is also important to highlight the finding that the children enjoyed responding to the SWLS-C and thought it was important to ask children their age these questions.

Messick (1995) stated that "validity is an evolving property and validation a continuing process" (p. 741). In future studies it would be of interest to investigate and compare the cognitive processes that are employed by children and adolescents of different age groups, with different socio-economic background, and of diverse ethno-cultural background when responding to the items of the SWLS-C (i.e., to investigate the generalizability aspect of construct validity). Furthermore, it will be important for future research to monitor for which purposes the SWLS-C is administered and to critically investigate the intended and unintended consequences of the use and interpretation of the SWLS-C scores with regard to these purposes (i.e., to investigate the consequential aspect of construct validity).

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- Barofsky, I., Meadows, K., & McColl, E. (2003). Cognitive aspects of survey methodology and quality of life assessment: Summary of meeting. *Quality of Life Research, 12*, 281–282.
- Bee, H. (1989). *The developing child* (5th ed.). New York: Harper & Row Publishers.
- Ben-Arieh, A. (2005). Where are the children? Children's role in measuring and monitoring their well-being. *Social Indicators Research, 74*, 573–596.
- Ben-Arieh, A. (2006). Is the study of the "state of our children" changing? Re-visiting after 5 years. *Children and Youth Services Review, 28*, 799–811.
- Ben-Arieh, A., & Goerge, R. M. (2001). Beyond the numbers: How do we monitor the state of our children? *Children and Youth Services Review, 23*, 603–631.
- Ben-Arieh, A., & Goerge, R. M. (Eds.). (2006). *Indicators of children's well-being: Understanding their role, usage and policy influence*. Dordrecht, The Netherlands: Springer Academic Publishers.
- Ben-Arieh, A., Kaufman, H. N., Andrews, B. A., Goerge, R. M., Lee, B. J., & Aber, J. L. (2001). *Measuring and monitoring children's well-being*. Dordrecht, The Netherlands: Kluwer Academic Press.
- Berg, B. L. (2004). *Qualitative research methods for the social sciences* (5th ed.). Boston: Pearson Education.
- Böhm, A. (2004). Theoretical coding: Text analysis in grounded theory. In U. Flick, E. von Kardoff, & I. Steinke (Eds.), *A companion to qualitative research* (pp. 270–275). London: Sage Publications.

- Cizek, G. J., Rosenberg, S. L., & Koons, H. H. (2008). Sources of validity evidence for educational and psychological tests. *Educational and Psychological Measurement*, *68*, 397–412.
- Collins, D. (2003). Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research*, *12*, 229–238.
- Cremeens, J., Eiser, C., & Blades, M. (2006a). A qualitative investigation of school-aged children's answers to items from a generic quality of life measure. *Child: Care, Health and Development*, *33*, 83–89.
- Cremeens, J., Eiser, C., & Blades, M. (2006b). Characteristics of health-related self-report measures for children aged three to eight years: A review of the literature. *Quality of Life Research*, *15*, 739–754.
- 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*, 659–685.
- Dew, T., & Huebner, E. S. (1994). Adolescents' perceived quality of life: An exploratory investigation. *Journal of School Psychology*, *32*, 185–199.
- Dey, I. (1993). *Qualitative data analysis: A user-friendly guide for social scientists*. London: Routledge.
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment*, *49*, 71–75.
- Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin*, *125*, 276–302.
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*, 215–251.
- Fox, J. E., Houston, B. K., & Pittner, M. S. (1983). Trait anxiety and children's cognitive behaviors in an evaluative situation. *Cognitive Therapy and Research*, *7*, 149–154.
- Frones, I. (2007). Theorizing indicators. *Social Indicators Research*, *83*, 5–23.
- Gadermann, A. M., Schonert-Reichl, K. A., & Zumbo, B. D. (2010). Investigating validity evidence of the Satisfaction with Life Scale adapted for Children. *Social Indicators Research*, *96*, 229–247.
- Greenspoon, P. J., & Saklofske, D. H. (1997). Validity and reliability of the Multidimensional Students' Life Satisfaction Scale with Canadian children. *Journal of Psychoeducational Assessment*, *15*, 138–155.
- Holsti, O. R. (1969). *Content analysis for the social sciences and humanities*. Reading, MA: Addison-Wesley Publishing Company.
- Huebner, E. S. (1991). Correlates of life satisfaction in children. *School Psychology Quarterly*, *6*, 103–111.
- Huebner, E. S., & Alderman, G. L. (1993). Convergent and discriminant validation of a children's life satisfaction scale: Its relationship to self- and teacher-reported psychological problems and school functioning. *Social Indicators Research*, *30*, 71–82.
- Huebner, E. S., Drane, W., & Valois, R. F. (2000). Levels and demographic correlates of adolescent life satisfaction reports. *School Psychology International*, *21*, 281–292.
- Huebner, E. S., Suldo, S. M., Smith, L. C., & McKnight, C. G. (2004). Life satisfaction in children and youth: Empirical foundations and implications for school psychologists. *Psychology in the Schools*, *41*, 81–93.
- Lodge, J., Harte, D. K., & Tripp, G. (1998). Children's self-talk under conditions of mild anxiety. *Journal of Anxiety Disorders*, *12*, 153–176.
- Lodge, J., Tripp, G., & Harte, D. K. (2000). Think-aloud, thought-listing, and video-mediated recall procedures in the assessment of children's self-talk. *Cognitive Therapy and Research*, *24*, 399–418.
- Man, P. (1991). The influence of peers and parents on youth life satisfaction in Hong Kong. *Social Indicators Research*, *24*, 347–365.
- Mayring, P. (2004). Qualitative content analysis. In U. Flick, E. von Kardoff, & I. Steinke (Eds.), *A companion to qualitative research* (pp. 266–269). London: Sage Publications.
- McCull, E., Meadows, K., & Barofsky, I. (2003). Cognitive aspects of survey methodology and quality of life assessment. *Quality of Life Research*, *12*, 217–218.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 12–103). New York: Macmillan.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from person's responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*, 741–749.
- Michalos, A. C. (1985). Multiple discrepancies theory. *Social Indicators Research*, *16*, 347–413.
- Michalos, A. C. (1991). *Global report on student well-being: Life satisfaction and happiness* (Vol. 1). New York: Springer.
- Neuendorf, K. A. (2002). *The content analysis guidebook*. Thousand Oaks, CA: Sage Publications.
- Rebok, G., Riley, A., Forrest, C., Starfield, B., Green, B., Robertson, J., et al. (2001). Elementary school-aged children's reports of their health: A cognitive interviewing study. *Quality of Life Research*, *10*, 59–70.
- Riley, A. W. (2004). Evidence that school-age children can self-report on their health. *Ambulatory Pediatrics*, *4*, 371–376.

- Schilling, L. S., Dixon, J. K., Knafl, K. A., Grey, M., Ives, B., & Lynn, M. R. (2007). Determining content validity of a self-report instrument for adolescents using a heterogeneous expert panel. *Nursing Research, 56*, 361–366.
- Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist, 54*, 93–105.
- Sirken, M. G., & Schechter, S. (1999). Interdisciplinary survey methods research. In M. G. Sirken, D. J. Herrmann, S. Schechter, N. Schwarz, J. M. Tanur, & R. Tourangeau (Eds.), *Cognition and survey research* (pp. 1–10). New York: Wiley.
- Stewart, J. L., Lynn, M. R., & Mishel, M. H. (2005). Evaluating content validity for children's self-report instruments using children as content experts. *Nursing Research, 54*, 414–418.
- Stone, W. L., & Lemanek, K. L. (1990). Developmental issues in children's self-reports. In A. M. La Greca (Ed.), *Through the eyes of the child: Obtaining self-reports from children and adolescents*. Boston: Allyn & Bacon.
- Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks, CA: Sage Publications.
- Willis, G. B., DeMaio, T. J., & Harris-Kojetin, B. (1999). Is the Bandwagon headed to the methodological promised land? Evaluating the validity of cognitive interviewing techniques. In M. G. Sirken, D. J. Herrmann, S. Schechter, N. Schwarz, J. M. Tanur, & R. Tourangeau (Eds.), *Cognition and survey research* (pp. 133–153). New York: Wiley.
- Young, M. H., Miller, B. C., Norton, M. C., & Hill, E. J. (1995). The effect of parental supportive behaviors on life satisfaction of adolescent offspring. *Journal of Marriage & Family, 57*, 813–822.
- Zumbo, B. D. (2007). Validity: Foundational issues and statistical methodology. In C. R. Rao & S. Sinharay (Eds.), *Handbook of statistics, Vol. 26: Psychometrics* (pp. 45–79). Amsterdam: Elsevier Science B. V.
- Zumbo, B. D. (2009). Validity as contextualized and pragmatic explanation, and its implications for validation practice. In R. W. Lissitz (Ed.), *The concept of validity: Revisions, new directions and applications* (pp. 65–82). Charlotte, NC: Information Age Publishing.