

The Development and Psychometric Properties of the Arab Children's Self-Concept Scale

**Hessa Abdel Rahman Fakhroo, University of Qatar
Sabry Abd-El-Fattah, University of South Australia
Anwar Abd-El-Rahem, El-Minia University, Egypt**

Abstract

The Arab Children's Self-Concept Scale (ACSCS) was developed using a sample of 460 children (220 males and 240 females) enrolled in a number of primary schools in the State of Qatar. An exploratory factor analysis of the ACSCS identified three factors: (a) Academic self-concept (12 items), (b) Non-academic self-concept (13 items), and (c) General self-concept (6 items). A confirmatory factor analysis revealed that the ACSCS could be described by three first-level factors that were linked by a higher order factor of overall self-concept. The correlation among the three factors of the ACSCS ranged from 0.30 to 0.33. The factorial structure of the ACSCS held invariant across children males and females groups. The ACSCS showed satisfactory reliability using Cronbach alpha and the congeneric model. Items of the ACSCS fitted the Rasch model adequately. Implications of these findings for measuring children's self-concept are discussed.

Keywords: Children, self-concept, Arab, Qatar, factor analysis, Rasch analysis

Introduction

"The fact that everyone can use the term "self" with such ease and familiarity suggests that the concept of selfhood is rooted in some simple, universal human experience" (Baumeister, 1999, p. 1). Your "self" is the essence of who you are and how you appear to other people. Self lends continuity to your identity over time, tying together your early childhood experiences, your unique approach to the people and events around you, and your aspirations and apprehensions about how you may develop in the future.

Self-concept features prominently in several social-cognitive learning theories, including theories of expectancy-value (Eccles, Barber, Updegraff, & O'Brien, 1998), self-worth (Covington, 1992), self-efficacy (Bandura, 1989), and explanatory styles (Seligman, Peterson, Kaslow, Tanenbaum, Alloy, & Abramson, 1984). In its most generic terms, self-concept refers to one's self-beliefs and self-evaluations, and presumably to how one answers the questions "Who am I?", and "How do I feel about myself?" It can also be described as a mental representation of those personal qualities that individuals use for the purpose of defining themselves and regulating their behaviours (Marsh, 1990).

The development of self-concept is thought to contain information about the characteristic features of the self in specific situations, and temporal contexts and moods (Greenwald & Pratkanis, 1984; Marsh, 1990; Marsh & Ayotte, 2003). Individuals tend to develop and maintain their self-concept through the process of taking actions, reflecting on these actions, and receiving feedbacks and judgments on these actions from others. Therefore, it is possible to think of self-concept as a personality trait that can be developed and modified through situations and events (Byrne, 1990).

Researchers have highlighted that as children grow, their self-concept expands, develops, and becomes more differentiated (Harter, 1982, 1986; Shavelson,

Hubner, & Stanton, 1976). For example, Hung (1993) stated that “Children’s self-concepts become more complicated as they mature.” Similarly, Harter (1983, 1989) proposed that self-concept becomes increasingly abstract with age, shifting from concrete descriptions of behaviour in early childhood, to trait-like psychological constructs (e.g., popular, smart, and good-looking) in middle childhood, to more abstract constructs during adolescence. In addition, Marsh and his colleagues (Marsh, 1989; Marsh, Barnes, Cairns, & Tidman, 1984) emphasized that with increasing life experiences, children’s self-concept becomes more differentiated and highly correlated with external indicators of competence (e.g., skills, accomplishments, and self-concepts inferred by significant others). Marsh and his colleagues (Marsh, 1989; Marsh, Craven, & Debus, 1991) have studied the developmental patterns of increasing self-concept differentiation by examining the factor structure of the self-concept from childhood through adolescence. The results of their studies have revealed a steadily decreasing level of inter-factor correlations in older children/adolescents relative to young children. These findings suggest, in agreement with previous research and theory, that there is an increasing differentiation in individuals’ self-structure with age increase (Harter, 1982, 1986; Shavelson, et al., 1976).

Several researchers have been concerned with measuring children’s self-concept (see, Marsh, 1989, 1990; Marsh et al., 1991). For example, Wylie (1979) reviewed research on measuring children’s self-concept prior to 1977 and marshalled several weaknesses of these research, including (a) the use of idiosyncratic self-concept instruments with questionable psychometric properties, (b) the use of uninterpretable discrepancy scores, (c) small sample sizes, (d) the non-equivalence of groups in cross-sectional studies, (e) the effects of attrition in longitudinal designs, and (f) the lack of research combining both cross-sectional and longitudinal comparisons.

Harter and Pike (Harter, 1983, 1986; Harter & Pike, 1984) emphasised that measuring children’s self-concept may entail simplified item contents or pictorial representations, simplified response formats, and individually based interviews rather than the conventional paper-and-pencil tests that are group-administered. Similarly, Stipek and Mac Iver (1989) referred to the inadequacy of the existing measurement instruments of children’s self-concept. They recommended the use of more appropriate assessment procedures. In addition, Wylie (1989) noted several measurement problems associated with measuring children’s self-concept even though children apparently have some descriptive and evaluative self-conceptions. She stated that “The progress in theory and research on children’s self-concept could be stimulated through the development of better multidimensional measurement instruments” (p. 32)

Researchers have proposed several instruments to measure children’s self-concept and reviews can found in (Harter, 1986; Marsh, Byrne, & Shavelson, 1988; Marsh & Shavelson, 1985; Silon & Harter, 1985). For example, Wylie (1989) reviewed Joseph Pre-School and Primary Self-Concept Screening Test (Joseph, 1979). She described it as a test that relies on items from a variety of different domains to infer a global and undifferentiated self-concept. In addition, Wylie described a proposed instrument that was developed by Harter and Pike (1984) to measure several domains of young children’s self-concept, including: (a) physical, (b) cognitive, (c) peers, and (d) maternal. Each domain was defined by six bipolar items represented by parallel verbal statements and pictures.

Examinees are shown two pictures, one in which the target child appears with one other child and one in which the target child appears with five other children. The

examinee is advised that the first target child does not have many friends to play with, and that the other target child has many friends to play with. The examinee selects the target child who is most like the examinee. Then, the examinee indicates whether he or she is a lot like or just slightly like the chosen target child. There are more specific prompts for each item, including “*Do you have a whole lot of friends to play with or pretty many?*” This two-stage response format, consisting of two dichotomous responses, results in a 4-point response scale.

Harter and Pike (1984) highlighted a number of merits of their scale, including: (a) the items of the scale seem to be appropriate to children’s developmental level, (b) the scale excludes the notion of global self-concept, (c) the scale is based upon pictorial format, and (d) the scale uses a 4-point response scale that encouraged a greater range of responses than do the dichotomous responses typically used with young children, and (e) the scale can be administered individually to each child instead of using a group-administration procedure. Harter and Pike argued that this individual-administration procedure may help to ensure that the child understands an item and enable the administrator to clarify the meaning of an item for the child as well as the child’s response.

However, the psychometric properties and theorization of Harter and Pike’s instrument were vulnerable to criticism. Firstly, Harter and Pike failed to support their prior four-factor structure (i.e., physical, cognitive, peers, and maternal), presumably because they used only exploratory factor analysis to identify the factorial structure of their instrument (Marsh & Hocevar, 1985). Specifically, the exploratory factor analysis identified two factors: (a) Competence (incorporating the physical and cognitive scales) and (b) Social acceptance (incorporating the peer and maternal scales). Secondly, Harter and Pike failed to separate the physical and cognitive components of the competence scale.

Thirdly, the correlations among the physical and cognitive components ranged from 0.43 to 0.56, and failed to approach 1.0, even after correction for unreliability. Fourthly, although Harter and Pike claimed that their factor structure of young children’s self-concept was less differentiated than the factor structure found with older children, they neither administered their instrument to older children nor offered any evidence that self-concept became more differentiated within the 4-7 age range that they considered. Instead, Harter and Pike suggested that the factor structure was similar for their preschool/kindergarten sample and for their Year 1/Year 2 sample (see, Marsh et al., 1991). Marsh and Hocevar (1985) concluded that whereas the Harter and Pike’s instrument is important; it may be premature to conclude that children can only differentiate two broad components of self.

In several studies, Marsh and his colleagues (Marsh & Hocevar, 1985; Marsh, Smith, Barnes, 1983; Marsh et al., 1984) have developed the Self-Description Questionnaire (SDQ I). The SDQ I is based upon a multidimensional hierarchical model to measure children’s self-concept (Shavelson et al., 1976). Two characteristics of Shavelson et al.’s model can be identified. Firstly, it is multi-dimensional, with each dimension reflecting an individual or groups’ self-referent category system. Secondly, it is hierarchical, with personal behaviour in specific situations at the base, broad inferences about the self in the middle, and global/general self-concept at the apex (see, Marsh, et al., 1988; Marsh & Hocevar, 1985). The highest level is general self-concept and is subsequently divided into non-academic self-concept and academic self-concept. The non-academic self-concept includes social, emotional, and physical self-concepts. Each of these categorizations is divided into several domains that are more concrete and specific. Finally, the most detailed, discrete level under each

domain includes the performance of specific behaviours. Self-concepts of different fields are correlated but can also be individually explained (Byrne & Shavelson, 1996; Hou, 2001; Byrne 2002).

The SDQ I (Marsh, et al., 1983; Marsh et al., 1984) is a 76-items scale that intends to measure three areas of children's self-concept: (a) Academic Self-Concept (32 items), including reading (10 items), mathematics (10 items), and general school (12 items), (b) Non-Academic Self-Concept (36 items), including physical ability (10 items), physical appearance (9 items), peer relationships (8 items), and parent relationships (9 items), and (c) General Self-concept scale (8 items).

Four total scores can be generated on the basis of these scales: (a) Academic Self-Concept (i.e., the average of scores on reading, mathematics, and general school self-concept scales), (b) Non-Academic Self-Concept (the average of scores on physical ability, physical appearance, peer relationships, and parent relationships self-concept scales), (c) Overall Self-Concept (the average of scores on Academic and Non-Academic Self-Concept scales), and (d) General Self-Concept (the average of scores on the items of the General Self-Concept Scale). Each item of the SDQ I is scored on a 6-point response scale (i.e., *False; Mostly false; Sometimes false, sometimes true; Mostly true; and True*). The SDQ I was subsequently identified in several factor analyses (see, Marsh, 1989; 1990; Marsh & Craven, 1997).

Context of the study

The present study is based upon data collected from a sample of primary school children in the State of Qatar. It is conceivable that entry to primary school represents a critical life event for the development and shaping of many children's self concept. Two reasons can be identified. Firstly, entry to primary school can represent an opportunity for young children to free themselves from the egocentric subjective self-concept, while it promotes them to develop an objective self-image to form a complete self-concept. Secondly, entry to primary school can expose children to different sources of information (e.g., peers, teachers, and others) that could help them develop their self-concept. It is possible that children assimilate and internalize this information in a way that helps them shape their personality, self-worth, and self-concept as they grow.

Goal of the study

In the present study, our first goal was to develop a new scale which would further articulate the notion of children's self-concept within an Arabic context. Specifically, we sought to construct a scale that will be of use to future researchers. Using Marsh SDQ I as a starting point, we developed items which reflected academic self-concept, non-academic self-concept, and general self-concept differentially. Hence, the scale was based upon previous questions as used within the literature, as well as incorporating several new items. Accordingly, we developed the Arab Children's Self-Concept Scale (ACSCS), and sought to test its psychometric properties using data collected from sample of primary school children in the State of Qatar.

Our second goal was to test for the invariance of the factorial structure of the ACSCS across Qatari males and females children. A third goal was to test whether the items of the ACSCS would calibrate on the Rasch model. A fourth goal was to explore the internal reliability of the ACSCS using both Cronbach alpha and the congeneric model (Fleishman & Benson, 1987).

Methods

Sample

Subjects of the present study included 460 Qatari students (220 males and 240 females) enrolled in several primary schools in the State of Qatar. The age range was from 10 to 12 years-olds.

Instrument and procedure

The Arab Children's Self-Concept Scale (ACSCS) was developed for use within the present study. The ACSCS consisted of 31 items. Several items were adopted from the SDQ I (Marsh et al., 1983), and were translated from English to Arabic. Other items were reworded to match the culture and educational context of the State of Qatar. In general, we felt that the other items from the SDQ I did not translate easily, and new items were generated based upon our understanding of children's self-concept notions. It was intended that 12, 13, and 6 items reflected the academic, non-academic and general self-concepts respectively. Respondents rated the applicability or inapplicability per item on a 3-point Likert type scale that ranged from 1 (*Absolutely Not Applicable*) to 3 (*Absolutely Applicable*). Children were recruited to participate in data collection voluntarily during their normal classes at their schools.

Results

Exploratory factor analysis

The Mplus 4.0 program (Muthen & Muthen, 2006) was used to run an exploratory factor analysis with oblique rotation of the ACSCS. The analysis identified three factors: (a) Non-Academic Self-Concept (13 items, Cronbach alpha = 0.70), (b) Academic Self-Concept (12 items, Cronbach alpha = 0.84), and (c) General Self-Concept (6 items, Cronbach alpha = 0.67). The percentage of variance explained by a specific factor was 17 per cent for Non-Academic Self-Concept, 13 per cent for Academic Self-Concept, and 11 per cent for General Self-Concept. Table 1 represents the ACSCS items and the factor loadings.

Confirmatory factor analysis

The Mplus 4.0 program was used to run a confirmatory factor analysis of the ACSCS model using the full information maximum likelihood estimation procedure (Bollen, 1989). The analysis showed that a nested model in which the ACSCS items were assigned to three specific correlated first-order factors of Academic Self-Concept (12 items), Non-Academic Self-Concept (13 items), and General Self-Concept (6 items) as well as a general higher order factor (labelled as self-concept) provided the best fitting model, $\chi^2(431, N = 460) = 468, p = 0.11$, Root-Mean-Square Error of Approximation (RMSEA) = .02, Standardized Root-Mean-Square Residual (SRMR) = 0.02, Adjusted Goodness of Fit Index (AGFI) = 0.98, Parsimonious Goodness of Fit Index (PGFI) = 0.30, Parsimony Ratio (PRATIO) = 0.85, Parsimony Normed Fit Index (PNFI) = 0.84.

The test statistic for the significance of the hypothesised path coefficients is the critical ratio. The critical ratio represents the parameter estimate divided by its standard error, as such, it operates as a z-statistic in testing whether the estimate is statistically different from zero. Based on a significance level of .05, the test statistic needs to be $> \pm 1.96$ before the hypothesis that the estimate equals 0.0 can be rejected

(Byrne, 2001). In accordance with these guidelines, Table 2 shows that all the hypothesized regression path coefficients of the ACSCS model were statistically significant.

The disattenuated correlation coefficients were calculated among the three factors of the ACSCS (Bollen, 1989). Non-academic self-concept has a positive relationship with academic self-concept ($r = 0.30$), and general self-concept ($r = 0.33$). There was a positive relationship between academic self-concept and general self-concept ($r = 0.31$).

Table 1

Exploratory factor analysis of the Arab Children's Self-Concept Scale

Item	Statement	Factor loadings
Non-academic self-concept		
1	Most other kids like me.	0.64
2	I get along with kids easily.	0.62
3	My father and mother like me. **	0.60
4	My peers want me to be their friend. **	0.55
5	I enjoy physical education classes. *	0.52
6	I can run along way without stopping.	0.50
7	I spend along time with my father and mother. **	0.48
8	I am a well-dressed person. *	0.48
9	I have a good looking body.	0.42
10	I joke a lot with my father and mother. *	0.39
11	My mother and father understand me. **	0.39
12	I have lots of friends.	0.35
13	I have more friends than most other kids.	0.35
Eigenvalue		3.2
Academic self-concept		
14	All school subjects are easy for me. **	0.76
15	Reading is easy for me. **	0.75
16	I am good at mathematics.	0.71
17	I read a lot. *	0.66
18	I like reading. **	0.65
19	It is easy for me to solve mathematics drills. *	0.61
20	I enjoy solving mathematical drills. *	0.58
21	I am easy to like.	0.56
22	I look forward to all school subjects.	0.55
23	I like Arabic. *	0.52
24	I am good at Arabic. *	0.51
25	I like mathematics.	0.40
Eigenvalue		4.5
General self-concept		
26	A lot of things about me are good.	0.65
27	I can do things as well as most other people.	0.64
28	I have a lot of qualities to be proud of. **	0.62
29	I can do anything right. **	0.60
30	Other people think I am a good person. **	0.58
31	In general, I like being the way I am.	0.55
Eigenvalue		2.2

Note. $N = 460$. * New generated items. ** Reworded items. Items with no asterisks were adopted from the SDQ

Table 2

Standardized path coefficients, standard error, critical ratio, error variance, and R^2 of the second-order confirmatory factor analysis of Arab Children's Self-Concept Scale

Paths	Path coefficient	Standard error	Critical ratio	Error variance	R^2
Non-academic self-concept					
1	0.65	0.11	7.8	0.62	0.38
2	0.62	0.08	5.7	0.68	0.32
3	0.57	0.10	4.2	0.70	0.30
4	0.55	0.13	4.8	0.77	0.23
5	0.48	0.10	4.8	0.68	0.32
6	0.57	0.12	8.8	0.72	0.28
7	0.53	0.06	4.4	0.81	0.19
8	0.44	0.10	6.7	0.78	0.22
9	0.47	0.07	7.5	0.80	0.20
10	0.45	0.06	4.8	0.86	0.14
11	0.38	0.08	5.9	0.83	0.17
12	0.41	0.07	5.8	0.88	0.12
13	0.35	0.06	7.8	0.62	0.38
Academic self-concept					
14	0.74	0.14	5.9	0.58	0.42
15	0.76	0.12	5.3	0.45	0.55
16	0.73	0.13	6.3	0.42	0.58
17	0.60	0.09	5.6	0.47	0.53
18	0.61	0.12	6.7	0.64	0.36
19	0.59	0.13	5.1	0.63	0.37
20	0.55	0.11	4.5	0.65	0.35
21	0.52	0.08	5.0	0.70	0.30
22	0.50	0.09	6.5	0.73	0.27
23	0.57	0.07	5.6	0.75	0.25
24	0.47	0.08	8.1	0.68	0.32
25	0.44	0.10	5.9	0.78	0.22
General self-concept					
26	0.60	0.12	5.0	0.64	0.36
27	0.57	0.13	4.4	0.68	0.32
28	0.64	0.10	6.4	0.59	0.41
29	0.55	0.08	6.9	0.70	0.30
30	0.58	0.07	8.3	0.66	0.34
31	0.51	0.10	5.1	0.74	0.26
Self-Concept					
Non-academic self-concept	0.65	0.13	5.0	0.58	0.42
Academic self-concept	0.67	0.12	5.6	0.55	0.45
General self-concept	0.63	0.10	6.3	0.60	0.40

Note. $N = 460$

Factorial invariance

Separate models. The Mplus 4.0 program was used to test two separate models of the ACSCS for males and females groups (Byrne, 2001). In relation to the males group, the model fitted the data satisfactorily, ($\chi^2(431, N = 220) = 455.8, p = .20$), Root-Mean-Square Error of Approximation (RMSEA) = 0.02, Standardized Root-Mean-Square Residual (SRMSR) = 0.02, Adjusted Goodness of Fit Index

(AGFI) = 0.97, and Parsimonious Goodness of Fit Index (PGFI) = 0.36. In relation to the females group, the ACSCS model fitted the data satisfactorily, (χ^2 (431, $N = 240$) = 453.2, $p = .22$), Root-Mean-Square Error of Approximation (RMSEA) = 0.02, Standardized Root-Mean-Square Residual (SRMR) = 0.01, Adjusted Goodness of Fit Index (AGFI) = 0.98, and Parsimonious Goodness of Fit Index (PGFI) = 0.31.

Baseline model. A baseline model of the ACSCS was tested simultaneously for males and females groups (Joreskog & Sorbom, 1985). Results of the analysis revealed that the baseline model fitted the data satisfactorily, (χ^2 (860, $N = 460$) = 909, $p = .12$), Root-Mean-Square Error of Approximation (RMSEA) = 0.01, Standardized Root-Mean-Square Residual (SRMR) = 0.01, Adjusted Goodness of Fit Index (AGFI) = 0.98, and Parsimonious Goodness of Fit Index (PGFI) = 0.30.

It is important to note that the χ^2 statistic and its degrees of freedom are additive. That is the sum of the χ^2 values, derived from the baseline model, reflects the extent to which the underlying structure fits the data across groups with no cross-group constraints imposed. The baseline model, without invariance constrained, represents an absolute upper bound, an optimum or target, for the alternative models that impose equality constraints (Abd-El-Fattah, 2006).

In the present analysis, the χ^2 statistics for any test of invariance across males and females groups must be at least 909 (i.e., 455.8 + 453.2); the sum of χ^2 statistics of the ACSCS model across males and females groups. No alternative model that requires any or all parameters to be invariant across males and females groups can have χ^2 statistic less than 909. The χ^2 statistics for such an alternative model would only approach 909 if parameters estimates in the unconstrained model for males and females groups are similar (Abd-El-Fattah & Barnes, 2006)

Constrained model. A constrained model of the ACSCS was constructed with factor loadings and factor variance being constrained equal across males and females groups. Results of the analysis showed that the constrained model fitted the data satisfactorily, (χ^2 (875, $N = 460$) = 930.7, $p = .09$), Root-Mean-Square Error of Approximation (RMSEA) = 0.02, Standardized Root-Mean-Square Residual (SRMR) = 0.03, Adjusted Goodness of Fit Index (AGFI) = 0.97, and Parsimonious Goodness of Fit Index (PGFI) = 0.32.

Nested models. The baseline and the constrained models are nested (Bollen, 1989; Joreskog, 1985). A nested model is one with parameter restrictions compared to a full model. One model is nested compared to another if the researcher can go from one model to the other by adding constraints or by freeing constraints. Constraints may include setting paths to zero, that is, making a given variable independent of others in the model. However, the two models will still have the same variables. The statistical significance of the chi-square difference statistics ($\Delta\chi^2$) for two nested models can be evaluated relative to the differences in the degrees of freedom (Δdf) of the two models (Abd-El-Fattah & Yates, 2006). In the present analysis, the $\Delta\chi^2$ between the baseline model and the constrained model of the ACSCS was nonsignificant ($\Delta\chi^2 = 21.7$, $\Delta df = 15$, $p = .12$). This finding suggests that the factorial structure of the ACSCS was invariant across males and females groups.

Rasch analysis

Since the ACSCS represents a new measure of children's self-concept within an Arabic context, it was desirable to investigate further its psychometric properties through the use of item analysis techniques. Specifically, the Rasch analysis has been widely acknowledged as a useful scaling procedure to examine individuals' response patterns. The Rasch model is based on the assumption that is a single underlying

latent trait is being measured (i.e., unidimensionality). In addition, items and individuals can be placed on a continuum that expresses the difficulty of the items relative to the ability of individuals (Wright & Stone, 1979). The basic Rasch scale is a dichotomous response model (Rasch, 1960) that represents the conditional probability of a binary outcome as a function of a person's trait level (β_n) and an item's difficulty (δ_i). When $\beta_n > \delta_i$, $\beta_n = \delta_i$, and $\beta_n < \delta_i$, the chances of a "yes" response is greater than 50 per cent, equal to 50 per cent, and less than 50 per cent, respectively.

The rating scale model is an additive linear model that describes the probability that a specific person (n) will respond to a specific Likert-type item (i) with a specific rating scale step (x) (see, Andrich, 1978). The simple dichotomous response model can be extended to provide an appropriate model for use with polytomous response categories (i.e., 0, 1, 2) by the addition of an additional difficulty parameter; either a second δ parameter or a τ parameter. The Rasch rating scale model is given by

$$p_{nij} = fn \left(\frac{\exp(\beta_n - \delta_i - \tau_j)}{1 + \exp(\beta_n - \delta_i - \tau_j)} \right)$$

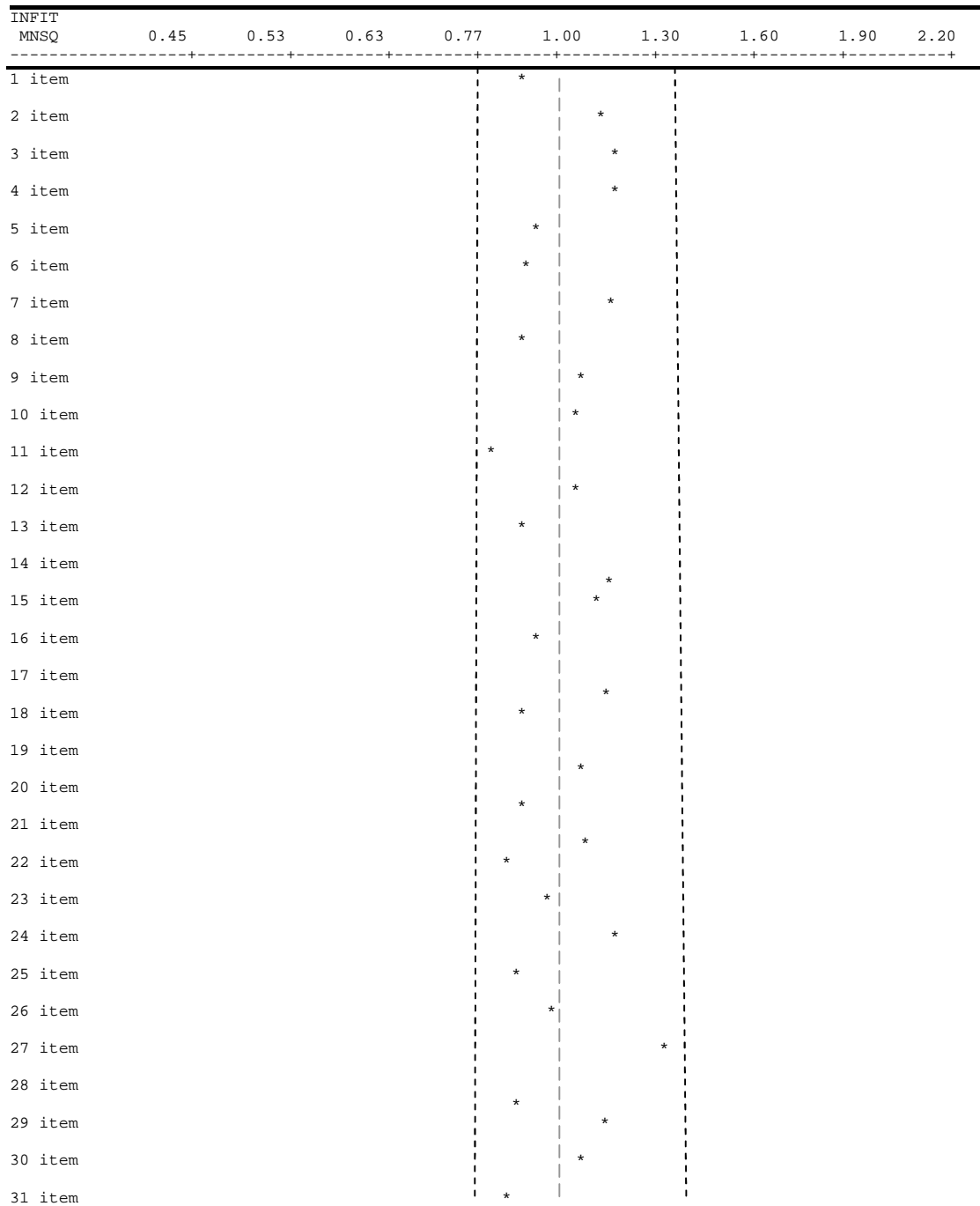
or

$$p_{nij} = \frac{\sum_{k=1}^{mi} \exp \sum_{j=1}^k (\beta_n - \delta_{ij})}{1 + \sum_{k=1}^{mi} \exp \sum_{j=1}^k (\beta_n - \delta_{ij})}$$

where n = subscript for persons, i = subscript for items, and j = subscript for response categories (0, 1, 2).

In the present study, the Quest program (Adam & Khoo, 1993) was used to run Rasch analysis of the ACSCS. However, the RUMM program (Andrich, Sheridan, & Luo, 2004) was used to produce the Item Characteristic Curve (ICC) and the Category Probability Curve (CPC) for selected items of the ACSCS.

Item fit. The infit mean square (Infit MNSQ) statistic measures the consistency of fit of the students to the ICC for each item, with weighted consideration given to those persons close to the 0.5 probability level (Abd-El-Fattah, 2007). The acceptable range of the Infit MNSQ statistic for each item is taken to be from 0.77 to 1.30 (Adams & Khoo, 1993). Values outside this acceptable range; that is above 1.30 indicate that these items do not discriminate well, and below 0.77 indicate that these items provide redundant information. In calibration, items that fall outside the acceptable range should be deleted from the analysis (Wright & Stone, 1979). Applying these guidelines to the present analysis, no item has been deleted from calibration because all values of the infit MNSQ statistic for all items of the ACSCS fall within the accepted range of 0.77 to 1.30. That is, all items of ACSCS fitted the Rasch model satisfactorily. Specifically, the values of the Infit MNSQ for the ACSCS ranged from 0.79 (Item 11) to 1.25 (Item 27). The Infit MNSQ statistics for the ACSCS are presented in Figure 1.



Note. $N = 460$. The values from 0.45 to 2.20 represent the scale of Infit MNSQ. The asterisks represent the Infit MNSQ for each item of the entity belief subscale. The vertical dotted line represents the range of the acceptable Infit MNSQ values

Figure 1. Item fit map for the 31 items of the Arab Children’s Self-Concept Scale

Item Characteristic Curve (ICC). In the RUMM program, the average ability of individuals with each Class Interval (CI) for each item is represented by a dot on the ICC for each item. If an item fits the Rasch model satisfactorily, the dots should fall on or as close as possible to the ICC (Andrich et al., 2004).

In the present analysis, the RUMM program divided the sample ($N = 460$) into 4 CIs, which were represented by 4 dots on the ICC of each item. For illustration purposes, the ICC for Item 21 of the ACSCS is depicted in Figure 2.

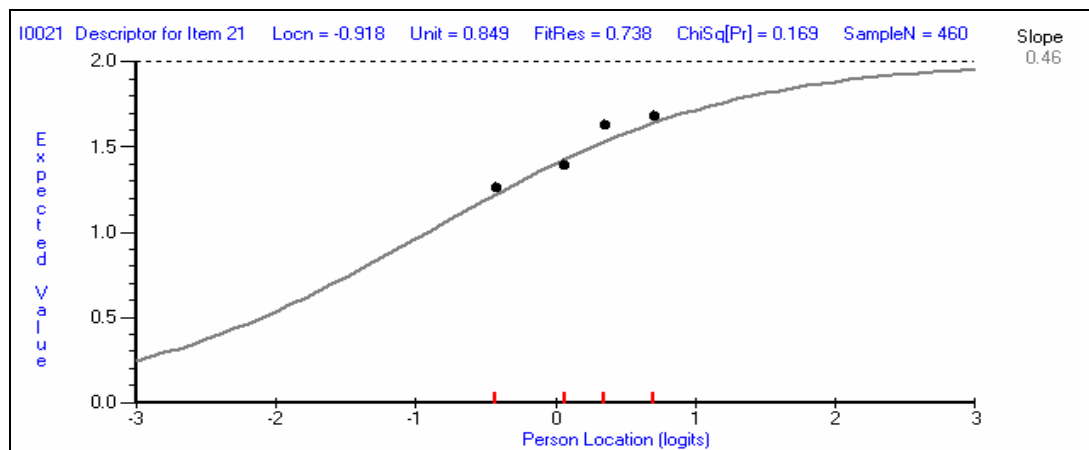


Figure 2. Item Characteristic Curve for Item 21 of the ACSCS

Thresholds. The RUMM program produces the threshold values that represent the difficulty level of each item. According to Bond and Fox (2001), threshold is “the level at which the likelihood of failure to endorse a given response category (below the threshold) turns to the likelihood of endorsing the category (above the threshold)” (p. 234).

In the present analysis, there are two thresholds marks the boundaries between the three response categories: ANA (*Absolutely Not Applicable*), SA (*Somewhat Applicable*), and AA (*Absolutely Applicable*) and all are ordered. That is, the data are regarded as ordinal and the Rasch model transform the counts of the endorsement of these ordered Likert categories into interval scales (Bond & Fox, 2001). For illustration purposes, the threshold for Item 21 of the ACSCS is depicted in Figure 3.

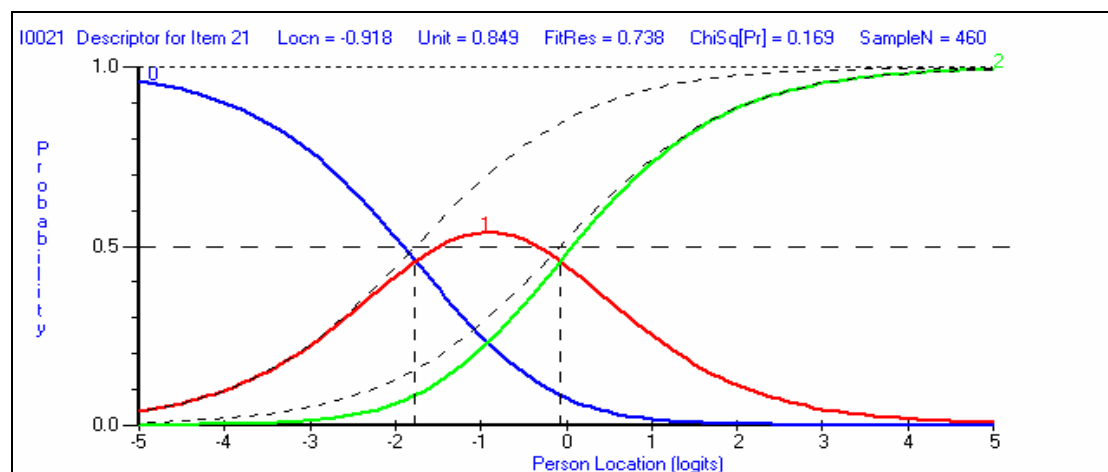


Figure 3. Category Probability Curve for Item 21 of the ACSCS

Case estimates. The other way of investigating the fit of the Rasch scale to data is to examine the estimates for each case. The case estimates give the performance level of each student on the total scale. In order to identify whether the cases fit the scale or not, it is important to examine the case OUTFIT mean square statistic (OUTFIT MNSQ), which measures the consistency of the fit of the persons to the student characteristic curve for each student, with special consideration given to extreme items (Yim, Abd-El-Fattah, & Lee, 2007). The general guideline used for interpreting t value as a sign of misfit is if $t > 5 \pm$ (Wright & Stone, 1979). Thus, if the

OUTFIT MNSQ value for a person has a t value $> 5 \pm$, that person does not fit the Rasch scale and is deleted from the analysis. In present analysis, however, no person was deleted, because the t values for all cases was less than $5 \pm$.

Congeneric reliability of the ACSCS

Cronbach alpha can be written as a function of the number of test items and the average inter-correlation among the items. Cronbach alpha is given by

$$\alpha = \frac{N \cdot \bar{r}}{(1 + (N - 1) \cdot \bar{r})}$$

where N is equal to the number of items and \bar{r} is the average inter-correlation among the items.

Cronbach alpha assumes that the items of a scale are tau-equivalent, that is, items have the same true scores, but may have unequal error variances. This means that each item is measuring the same construct to the same degree; a restrictive assumption that is unlikely to be met in practice. If the items of a scale were not tau-equivalent, Cronbach alpha could be a conservative estimate of reliability (Novick & Lewis, 1967).

Joreskog (1971) argued that the congeneric model is the least restrictive model based on the classical testing theory. The assumption underlying the congeneric model is that different items do not reflect the same true score.

$$X_j = T_j + E_j$$

Specifically, each observed test score (X) consists of the true score (T) for that test plus error (E). The congeneric model further assumes that the true scores for different items correlate perfectly with each other. This assumption led to a respecification of

$$X_j = M_j + \beta_j T + E_j$$

where β_j refers to the loading of an item on a generic true score and M_j represents the mean. In other words, each observed score reflects the same generic true score but to different degrees, as reflected in different β_j for different items (Alwin & Jackson, 1980). If the congeneric model applies, the degree to which each item reflects the generic true score (as shown in the β_j coefficients) must be considered. If items are summed to form a scale, the reliability estimate of this scale is given by

$$p_{xx} = \frac{(\sum \beta_j)^2}{(\sum \beta_j)^2 + \sum \theta_j^2}$$

where θ_j^2 and β represent the values of error variance and the regression path coefficient for each item, respectively (Fleishman & Benson, 1987).

Table 3 shows that the reliability estimates of the ACSCS subscales obtained based on the congeneric model were higher compared to the reliability estimates based on the tau-equivalent model (i.e., Cronbach alpha). This finding is consistent with the suggestion that Cronbach alpha works as a lower bound estimate of reliability when the assumption of the tau-equivalence model is not met (Novick & Lewis, 1967; Shevlin, Miles, Davies, & Walker, 2000).

Table 3

Reliability estimates of three subscales of the Arab Children's Self-Concept Scale based on the tau-equivalent model (i.e., Cronbach alpha) and congeneric model

Scale	Tau-equivalent model	Congeneric model
ACSCS		
1. Non-Academic Self-Concept	0.70	0.76
2. Academic Self-Concept	0.84	0.89
3. General Self-Concept	0.67	0.73

Note. $N = 460$ for all instances.

Mean testing

A series of one-way analysis of variance (ANOVA) was run to test the differences between males and females students in Non-Academic Self-Concept, Academic Self-Concept, and General Self-Concept. Results of the analysis are shown in Table 4.

Table 4

Mean differences across males and females groups in three subscales of the Arab Children's Self-Concept Scale

Variables /Statistic	<i>M</i>	<i>SD</i>	<i>F</i> (1, 458)	<i>d</i> ⁽¹⁾
1. Non-Academic Self-Concept			1.9	
Males	19.4	3.9		
Females	18.9	3.4		
2. Academic Self-Concept			7.3*	0.26
Males.	29	4.9		
Females	30.2	4.3		
3. General Self-Concept			1.4	
Males.	14.7	2.2		
Females	15	2.1		

Note. $N = 460$, males = 220, and females = 240 for all instances. * $p < .01$. $df = (1, 458)$. ⁽¹⁾ *d* represents Cohen's effect size (Cohen, 1988).

Discussion

The purpose of the present study was to develop the Arab Children's Self-Concept Scale (ACSCS). We sought to develop an index of children's self-concept that holds satisfactory psychometric properties and can be used within an Arab context.

A confirmatory factor analysis showed that the ACSCS could be described by three first-level factors (i.e., Non-Academic Self-Concept, Academic Self-Concept, and General Self-Concept) that were linked by a higher order factor of self-concept in a manner that is both meaningful and highly similar across children males and females groups. The Rasch analysis revealed that none of the ACSCS items appeared to infringe the conventional calibration points. Specifically, all items of the ACSCS fall within an acceptable range of children's ability, as presented by the use of a Rasch type scaling procedure. In addition, the ACSCS factors showed acceptable reliability using Cronbach alpha and the congeneric model.

At the outset, we were unsure if the three intended factors (i.e., Non-Academic Self-Concept, Academic Self-Concept, and General Self-Concept) could be defined clearly, as earlier studies did not suggest that this would be the case. However, the current data do suggest that it is meaningful to describe scores of every factor separately or sum the scores on the three factors to obtain an overall score which represents a participant's overall self-concept.

It appears more appropriate to employ the ACSCS to describe scores along traditional psychometric dimensions. Within the current dataset, responses to ACSCS items conform to the standard distributional assumptions, and the coherent and normal properties appear to be evident.

Were there any meaningful differences in the way Qatari male and female children responded to the ACSCS? We constructed tallies of the three factors of the ACSCS (i.e., Non-Academic Self-Concept, Academic Self-Concept, and General Self-Concept) for the male and female samples independently. Results from mean testing showed that levels of Non-Academic Self-Concept and General Self-Concept were similar across both samples. However, females showed higher levels of academic self-concept than males (see Table 4).

Since the ACSCS consisted of 31-items, and can be completed within 5 to 10 minutes, we believe that the ACSCS is an instrument that will prove useful in further studies into Arab children's self-concept. The original intention was for the ACSCS to be serviceable, brief, and easy to administer and score.

References

- Abd-El-Fattah, S. M. (2007). Is the Aggression Questionnaire bias free? A Rasch analysis. *International Education Journal*, 8, 237-248.
- Abd-El-Fattah, S. M. (2006). Effects of family background and parental involvement on Egyptian adolescents' academic achievement and school disengagement: A structural equation modelling analysis. *Social Psychology of Education*, 9, 139-157.
- Abd-El-Fattah, S. M., & Barnes, A. (2006). *Confirmatory factor analysis of Bath County Computer Attitude Scale within an Egyptian context: Testing competing models*. Paper presented at the Australian Association for Research in Education Conference, Adelaide, South Australia.
- Abd-El-Fattah, S. M., & Yates, G. C. R. (2006). *Implicit Theory of Intelligence Scale: Testing for factorial invariance and mean structure*. Paper presented at the Australian Association for Research in Education Conference, Adelaide, South Australia.
- Baumeister, R. F. (1999). The nature and structure of the self: An overview. In R. Baumeister (Ed.), *The self in social psychology* (pp. 1-20). Philadelphia, PA: Psychology Press (Taylor & Francis).
- Adams, R. J., & Khoo, S. T. (1993). *Quest-the interactive test analysis system*. Hawthorn, Victoria: ACER.
- Alwin, D. F., & Jackson, D. J. (1980). Measurement models for response errors in surveys. In K. F. Schuessler (Ed.), *Sociological methodology* (pp. 68-119). San Francisco: Jossey-Bass.
- Andrich, D. (1978). Rating formulation for ordered response categories. *Psychometrika*, 43, 561-573.
- Andrich, D., Sheridan, B., & Luo, G. (2000). RUMM2010: A Windows interactive

- program for analysing data with Rasch unidimensional Models for Measurement [Computer Program]. Perth, Western Australia: RUMM Laboratory.
- Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist*, *44*, 1175-1184.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: John Wiley.
- Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Byrne, B. M (1990). Self-concept and academic achievement: Investigating their importance as discriminators of academic track membership in high school. *Canadian Journal of Education*, *15*, 173-182.
- Byrne, B. M., & Shavelson, R. J. (1996). On the structure of social self-concept for pre-, early, and late adolescents: a test of the Shavelson, Hubner, and Stanton's (1976) model. *Journal of Personality and Social Psychology*, *70*, 599-613.
- Byrne, B. M (2001). *Structural equation modelling with AMOS: Basic concepts, applications, and programming*. New Jersey: Lawrence, Erlbaum Associates, Inc.
- Byrne, B. M. (2002). Validating the measurement and structure of self-concept: Snapshots of past, present, and future research. *American Psychologist*, *57*, 897-909
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Erlbaum publications
- Covington, M. V. (1992). *Making the grade: A self-worth perspective on motivation and school reform*. Cambridge: Cambridge University Press.
- Eccles, J. S., Barber, B. L., Updegraff, K., & O'Brien, K. M. (1998). *An expectancy-value model of achievement choices: The role of ability self-concepts, perceived task utility and interest in predicting activity choice and course enrolment*. In A. K. L. Hoffmann, K. A. Renninger & J. Baumert (Eds.), *Interest and learning: Proceedings of the Seeon Conference on Interest and Gender* (pp. 267-280). Kiel, Germany: Institute for Science Education at the University of Kiel.
- Fleishman, J., & Benson, J. (1987). Using LISREL to evaluate measurement models and scale reliability. *Educational and Psychological Measurement*, *47*, 925-939.
- Greenwald, A. G., & Pratkanis, A. R. (1984). The self. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 3, pp. 129-178). Hillsdale, NJ: Lawrence Erlbaum.
- Harter, S. (1982). The Perceived Competence Scale for Children. *Child Development*, *53*, 87-97.
- Harter, S. (1983). Developmental perspectives on the self-system. In P. H. Mussen & E. M. Hetherington (Eds.), *Handbook of child psychology* (Vol. 4 Socialization and personality development). New York: Wiley.
- Harter, S., & Pike, R. (1984). The Pictorial Scale of Perceived Competence and Social Acceptance for Young Children. *Child Development*, *55*, 1969-1982.
- Harter, S. (1986). Processes underlying the construction, maintenance, and enhancement of the self-concept in children. In J. Suls & A. G. Greenwald (Eds.), *Psychology perspectives on the self* (Vol. 3, pp. 136-182). Hillsdale,

NJ: Erlbaum.

- Harter, S. (1989). Causes, correlates, and the functional role of global self-worth: A life-span perspective. In J. Kolligian & R. Sternberg (Eds.), *Perception of competence and incompetence across the life-span* (pp. 67-97). New Haven, CT: Yale University Press.
- Hou, Y. L. (2001). The development of the self-concept scale for elementary school children. *Psychological Testing*, 48, 141-166.
- Hung, R. H. (1993). The relationships study of elementary school children's self-concept. *The Guide Report*, 1, 243-279.
- Joreskog, K. G. (1971). Statistical analysis of sets of congeneric test. *Psychometrika*, 36, 109-133.
- Joreskog, K., & Sorbom, D. (1985). *LISREL VI: Analysis of linear structural relationships by the method of maximum likelihood*. Mooresville, IN: Scientific Software, Inc.
- Joseph, J. (1979). *Joseph Pre-School and Primary Self-Concept Screening Test*. Chicago, IL: Stoelting.
- Marsh, H. W., Smith, I. D., & Barnes, J. (1983). Multitrait-multimethod analyses of the Self-Description Questionnaire: Student-teacher agreement on multidimensional ratings of student self-concept. *American Educational Research Journal*, 20, 333-357.
- Marsh, H. W., Barnes, J., Cairns, L., & Tidman, M. (1984). The Self Description Questionnaire (SDQ): Age effects in the structure and level of self-concept for pre-adolescent children. *Journal of Educational Psychology*, 76, 940-956.
- Marsh, H. W., & Hocevar, D. (1985). The application of confirmatory factor analysis to the study of self-concept: First and higher order factor structures and their invariance across age groups. *Psychological Bulletin*, 97, 562-582.
- Marsh, H. W., & Shavelson, R. J. (1985). Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist*, 20, 107-123.
- Marsh, H. W., Byrne, B. M., & Shavelson, R. J. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. *Journal of Educational Psychology*, 80, 366-380.
- Marsh, H. W. (1989). Age and sex effects in multiple dimensions of self-concept: Preadolescence to early-childhood. *Journal of Educational Psychology*, 81, 417-430.
- Marsh, H. W. (1990). The structure of academic self-concept: The Marsh/Shavelson model. *Journal of Educational Psychology*, 84, 623-636.
- Marsh, H. W., Craven, R. G., & Debus, R. L. (1991). Self-concepts of young children aged 5 to 8: Their measurement and multidimensional structure. *Journal of Educational Psychology*, 83, 377-392.
- Marsh, H. W., & Craven, R. G. (1997). Academic self-concept: Beyond the dustbowl. In G. Phye (Ed.), *Handbook of classroom assessment: Learning, achievement and adjustment*. New York: Academic Press.
- Marsh, H. W., & Ayotte, V. (2003). Do multiple dimensions of self-concept become more differentiated with age? The differential distinctiveness hypothesis. *Journal of Educational Psychology*, 95, 687-706.
- Muthen, L., & Muthen, B. (2006). Mplus 4.0 [Statistical Program]. Los Angeles, CA: Muthen & Muthen.
- Novick, M. R., & Lewis, C. (1967). Coefficient alpha and the reliability of composite measurements. *Psychometrika*, 32, 1-13.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*.

- Chicago: University of Chicago Press.
- Seligman, M. E. P., Peterson, C., Kaslow, N. J., Tanenbaum, R. L., Alloy, L. B., & Abramson, L. Y. (1984). Attributional style and depressive symptoms among children. *Journal of Abnormal Psychology, 93*, 235-238.
- Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research, 46*, 407-441.
- Shevlin, M. E., Miles, J. N. V., Davies, M. N. O., & Walker, S. (2000). Coefficient alpha: a useful indicator of reliability? *Personality and Individual Differences, 28*, 229-238.
- Silon, E., & Harter, S. (1985). Assessment of perceived competence, motivational orientation, and anxiety in segregated and mainstreamed educable mentally retarded children. *Journal of Educational Psychology, 77*, 217-230.
- Stipek, D. J., & Mac Iver, D. (1989). Developmental change in children's assessment of intellectual competence. *Child Development, 60*, 521-538.
- Yim, H.Y. B., Abd-El-Fattah, S., M., & Lee L. W. M (2007). A Rasch analysis of the Teachers' Music Confidence Scale. *International Education Journal, 8*, 260-269.
- Wright, B. D., & Stone, M. H. (1979). *Best test design: Rasch measurement*. Chicago: Mesa Press.
- Wylie, R. C. (1989). *Measures of self-concept*. Lincoln, NE: University of Nebraska Press.