

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

**Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600**

UMI[®]



Université d'Ottawa • University of Ottawa

**Preadolescent Self-concept and Self-concept/Academic Achievement Relations:
Investigating Multidimensional and Hierarchical Structures
Within and Across Gender**

**Lisa Larocque
School of Psychology**

**Dissertation submitted to the School of Graduate Studies and Research
of the University of Ottawa in partial fulfillment of the requirements
for the degree of Doctor of Philosophy**

© Lisa Larocque, Ottawa, Canada, 1999.



National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-46529-2

Canada

This dissertation is dedicated
to my mother and father
for their steadfast love, acceptance, support, encouragement, and assistance
and
for their belief in me and my belief in myself

ACKNOWLEDGEMENTS

The completion of this dissertation would not have been possible without the contribution of numerous people. Prominent among these is my supervisor, Dr. Barbara Byrne, whom I thank for her invaluable guidance throughout this project. I would also like to thank my committee members, Drs. Gail Crombie, Catherine Lee, Luc Pelletier, and Richard Shavelson, for their support and feedback. Special thanks are extended to Drs. Richard Shavelson and Herbert Marsh for the work that formed the basis for this project, and to Drs. Herbert Marsh and Susan Harter for the use of their instruments.

I would also like to thank the students, teachers, and school officials whose participation helped render this project possible.

I am grateful to the Social Sciences and Humanities Research Council of Canada, the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche, and the University of Ottawa for funding support.

I would like to thank Moira Costello for her assistance in data collection, entry, and verification.

On a more personal note, I wish to express my appreciation to the many friends and family members who walked beside me throughout this journey. I am grateful for the support and assistance they so generously provided. To my mother, for our "middle of the night" conversations, and to my parents for always being willing to lend a helping hand. To my grandmother, Reina Larocque, and to Leonor and Arthur, for their prayers and good wishes. I also thank my sisters - Natalie, Julie, Nicole and Estelle - for their support and assistance. Thanks are also extended to Natalie and Doug for sharing their computer. I would like to thank Bernie Routhier for his willingness to share his knowledge, and for providing invaluable assistance. To my dearest friends, Alison Lee, Sarah Russell, and Karen Ricciuti, for their assistance, support, and encouragement.

Finally, to Wally, my partner, thank you for your support, encouragement, and understanding. And to my dog, Chewy, for her constant companionship.

ABSTRACT

Using a confirmatory factor analytic approach, the aim of the study was to examine the construct validity of an academic expansion of the Marsh/Shavelson model of self-concept (SC; Shavelson & Marsh, 1986; Marsh, 1990d), and its relations with academic achievement (AA) within and across gender. Participants were Grades 5 and 6 girls ($n = 220$) and boys ($n = 205$) drawn primarily from middle-class suburban communities in and around Ottawa, Canada. Four nonacademic SCs (i.e., physical appearance, physical ability, parents, and peers), one general-academic SC, and 11 subject-specific academic SCs were measured with the Self-Description Questionnaires (Marsh, 1990d, 1992b) and the Self-Perception Profile for Children (Harter, 1985). Specific academic SCs were measured for four Language Arts (i.e., Listening, Reading, Writing and Speaking), two mathematics (i.e., Arithmetic and Measurement), two science (i.e., Science and Social Studies), and three non-core (i.e., Religion, Art, and Physical Education/Gym) subjects. Grades and self-reported grades in eight of these subjects served as measures of AA. The hypothesized multidimensional SC measurement and structure was supported for both girls and boys. Tests of competing hierarchical SC structures provided support for five higher-order SCs (i.e., Nonacademic, Non-core Academic, and Verbal-, Mathematics-, and Science-Academic). Gym SC was better represented hierarchically under Nonacademic, rather than under Non-core Academic SC. Gender differences in the baseline hierarchical SC models related only to Parent SC, which was defined by only Non-core Academic SC for girls, but Verbal-Academic SC for boys. Reading, Writing, and Speaking in Language Arts AAs were not distinct, and thus,

were represented as a single Language Arts Skills AA construct. A multidimensional pattern of relations in which SC and AA constructs were more highly correlated for corresponding than for non-corresponding academic subject areas was supported, albeit with two exceptions. A multidimensional pattern of SC/AA relations was not consistently found for the Language Arts constructs for both girls and boys, and was not found for Art AA for girls. An hierarchical pattern of SC/AA relations in which AA constructs were more highly correlated with first- than with second-order SCs was not consistently demonstrated for either girls or boys. Tests for the gender invariance of SC measurement and structure, and of SC/AA relations were generally tenable. Gender differences that were of the largest magnitude indicated that, in comparison to girls, Art AA was more highly related to Art SC, but less correlated with both Speaking in Language Arts and Verbal-Academic SCs for boys.

Table of Contents

	Page
DEDICATION	ii
ACKNOWLEDGEMENTS	iii
ABSTRACT	iv
INTRODUCTION	1
Overview of Self-concept Literature	1
Limitations of Previous Self-concept Research	2
Recent Advances in Self-concept Research	3
Limitations of Current Self-concept Research	7
Overview of the Present Study	9
Theoretical Framework: The Shavelson Model of Self-concept	10
Construct Validity Research: Preadolescent Self-concept	15
Within-network Self-concept Research	16
Empirical Research with the Multidimensional Self Concept Scale	17
Empirical Research with the Self-Perception Profile for Children	17
Empirical Research with the Self-Description Questionnaire-I	21
The Marsh/Shavelson Model of Self-concept:	
A Revision of the Shavelson Model	26
The Internal/External Frame of Reference Model	29
The Expanded Marsh/Shavelson Model:	
An Expansion of Academic Self-concept	31
Between-network Research: Self-concept/Academic Achievement Relations	37
Multidimensional Self-concept/Academic Achievement Relations	40
Hierarchical Self-concept/Academic Achievement Relations	44
Construct Validity Research: Gender Differences in Preadolescent Self-concept	46
Within-network Research: The Invariance of a	
Multidimensional/Hierarchical Self-concept	48
Between-network Research: Self-concept/Academic Achievement Relations	51
Purpose of the Present Study	53
Importance of the Present Study	55

METHOD	56
Sample	56
Procedure	57
Instrumentation	58
Self-concept Measures	59
The Self-Description Questionnaire-I	61
The Academic Self-Description Questionnaire-I	63
The Self-Perception Profile for Children	63
Academic Achievement Measures	66
Grades	67
Self-report of Grades	68
Academic Importance Ratings	69
Data Analyses and Research Hypotheses	70
Preliminary Analyses	70
Research Hypotheses	74
Hypothesis 1	74
Hypothesis 2	75
Model H1	83
Model H2	84
Model H3	84
Model H4	85
Hypothesis 3	86
Hypothesis 4	86
Hypothesis 5	88
Test of Hypotheses: Model Evaluation and Modification	89
Model Evaluation	92
Model Modification	95
RESULTS	99
Preliminary Analyses	99
Data Preparation	99
Descriptive Statistics	101
Instrument Validity: The Self-Description Questionnaires and the Self-Perception Profile for Children	109
Cronbach Alpha Coefficients: Self-concept Instruments	109
Confirmatory Factor Analyses: Self-concept Instruments	111

Self-concept Measurement and Structure: Within-network Analyses	113
Preadolescent Girls: Testing for the Multidimensional Structure of Self-concept	114
Preadolescent Boys: Testing for the Multidimensional Structure of Self-concept	118
Testing for the Multidimensional Structure of Preadolescent Self-concept for Girls and Boys: A Summary	120
Testing for the Hierarchical Structure of Self-concept: An Overview	120
Preadolescent Girls: Testing for the Hierarchical Structure of Self-concept	121
Preadolescent Boys: Testing for the Hierarchical Structure of Self-concept	132
Testing for the Hierarchical Structure of Preadolescent Self-concept for Girls and Boys: A Summary	137
Testing for the Invariance of Self-concept Measurement and Structure Across Gender	146
Between-network Analyses: Self-concept/Academic Achievement Relations	147
Preadolescent Girls: Testing for Self-concept/Academic Achievement Relations	148
Preadolescent Boys: Testing for Self-concept/Academic Achievement Relations	153
Testing for the Multidimensional and Hierarchical Structure of Self-concept/ Academic Achievement Relations for Preadolescent Girls and Boys: A Summary	159
Testing for the Invariance of Self-concept/Academic Achievement Relations Across Gender	160
DISCUSSION	165
Preliminary Analyses	165
Self-concept Multidimensional and Hierarchical Measurement and Structure:	
Within-network Issues	169
Self-concept Multidimensional Measurement and Structure	170
Self-concept Hierarchical Structure	172
Gender Invariance: Self-concept Multidimensional Measurement and Hierarchical Structure	178
Between-network Issues: Self-concept/Academic Achievement Relations	182
Gender Invariance: Self-concept/Academic Achievement Relations	188
Strengths and Limitations of the Present Study	193
Sampling Frame	193
Use of Non-independent Self-concept Instruments	194
Sample Size: Cross-validation and Replication	194
The Breadth of Self-concept Dimensions	194
Academic Achievement Indicator Variables	195

Implications and Directions for Future Research and Clinical Intervention Efforts	195
REFERENCES	198
APPENDICES	214
APPENDIX A: INFORMED CONSENT ORAL PRESENTATIONS, AND PARENTAL PERMISSION LETTER AND FORMS	214
Student Oral Presentation: Request to Obtain Student Consent and to Forward Parental Permission Letter and Forms	214
First Administration Session: Oral Presentation to Students	216
Parent Permission Letter	217
Parent Permission Form	219
APPENDIX B: ACADEMIC SUBJECT AREA DESCRIPTIONS OR GRADING CATEGORIES PROVIDED BY THE SCHOOL BOARDS	220
APPENDIX C: INSTRUMENT STANDARDIZED INSTRUCTIONS, RESPONSE FORMAT, AND ITEMS	225
Standardized Instructions and Response Format for the Self-Description Questionnaire-I (SDQ-I) and the Academic Self-Description Questionnaire-I (ASDQ-I)	225
Standardized Instructions and Response Format for the Self-Perception Profile for Children (SPPC)	228
Standardized Instructions and Response Format for the Harter Importance Rating Scale	230
The Self-Description Questionnaire-I (SDQ-I) Items	232
The Academic Self-Description Questionnaire-I (ASDQ-I) Items	234
The Self-Perception Profile for Children (SPPC) Items	235
Children's Report of School Grades	237

Children's Ratings of the Importance of School Subjects	239
Harter Importance Rating Scale	241
APPENDIX D: LATENT FACTOR CORRELATIONS AMONG FIRST-ORDER SELF-CONCEPT (SC) CONSTRUCTS FOR MODEL 16G FOR GIRLS (<u>N</u> = 220) AND MODEL 16B FOR BOYS (<u>N</u> = 205)	242
APPENDIX E: COPYRIGHT PERMISSION LETTERS	243

LIST OF FIGURES

	Page
Figure 1	The representation of the multidimensional and hierarchical model of self-concept proposed by Shavelson and colleagues 12
Figure 2	The Marsh/Shavelson model of self-concept 27
Figure 3	A pictorial representation of the academic expansion of the Marsh/Shavelson model 34
Figure 4	A pictorial representation of the core academic portion of the first-order 16-factor self-concept (SC) model 76
Figure 5	A pictorial representation of the core academic portion of the first-order 15-factor self-concept (SC) model 77
Figure 6	A pictorial representation of the core academic portion of the first-order 13-factor self-concept (SC) model 78
Figure 7	A pictorial representation of the structural portion of hierarchical self-concept (SC) Model H1 79
Figure 8	A pictorial representation of the structural portion of hierarchical self-concept (SC) Model H2 80
Figure 9	A pictorial representation of the structural portion of hierarchical self-concept (SC) Model H3 81
Figure 10	A pictorial representation of the structural portion of the hierarchical self-concept (SC) Model H4 expansion 82
Figure 11	First-order estimates for the nonacademic and non-core academic portions of the final hierarchical self-concept (SC) model for girls (Model H4G ²) 138
Figure 12	First-order estimates for the core academic portions of the final hierarchical self-concept (SC) model for girls (Model H4G ²) 139
Figure 13	Second-order estimates for the final hierarchical self-concept (SC) model for girls (Model H4G ²) 140

Figure 14	First-order estimates for the nonacademic and non-core academic portions of the final hierarchical self-concept (SC) model for boys (Model H4B ³)	141
Figure 15	First-order estimates for the core academic portions of the final hierarchical self-concept (SC) model for boys (Model H4B ³)	142
Figure 16	Second-order estimates for the final hierarchical self-concept (SC) model for boys (Model H4B ³)	143

LIST OF TABLES

	Page
Table 1	Sample Mean Age Within and Across Grade and Gender 56
Table 2	Summary of the Specific Academic Areas Measured in Each of the Three School Boards and the Specific Academic Constructs Common Across the School Boards 60
Table 3	Final Sample Mean Age Within and Across Grade and Gender 100
Table 4	ASDQ-I and SDQ-I Scale Score Mean, Standard Deviation, Skewness, and Kurtosis Values for the Sample of Girls ($N = 220$) and of Boys ($N = 205$) 102
Table 5	SPPC Scale Score Mean, Standard Deviation, Skewness, and Kurtosis Values for the Sample of Girls ($N = 220$) and of Boys ($N = 205$) 103
Table 6	ASDQ-I Indicator Variable Score Mean, Standard Deviation, Skewness and Kurtosis Values for the Sample of Girls ($N = 220$) and of Boys ($N = 205$) 104
Table 7	SDQ-I Indicator Variable Score Mean, Standard Deviation, Skewness, and Kurtosis Values for the Sample of Girls ($N = 220$) and of Boys ($N = 205$) 105
Table 8	SPPC Academic Self-concept (SC) Indicator Variable Score Mean, Standard Deviation, Skewness, and Kurtosis Values for the Sample of Girls ($N = 220$) and of Boys ($N = 205$) 106
Table 9	SPPC Nonacademic Self-concept (SC) Indicator Variable Score Mean, Standard Deviation, Skewness, and Kurtosis Values for the Sample of Girls ($N = 220$) and of Boys ($N = 205$) 107
Table 10	Academic Achievement (AA) Indicator Variable Score Mean, Standard Deviation, Skewness, and Kurtosis Values for the Sample of Girls ($N = 220$) and of Boys ($N = 205$) 108
Table 11	Cronbach Alpha Coefficients for Girls ($N = 220$) and for Boys ($N = 205$) for Each of the (A)SDQ-I and SPPC Scale Scores, and the Mean Across Scales for Each Instrument 110

Table 12	Summary of the Goodness-of-fit Indices for the Sample of Girls: First-order 16-, 15-, and 13-factor Self-concept (SC) Models	115
Table 13	Summary of the Goodness-of-fit Indices for the Sample of Boys: First-order 16-, 15- and 13-factor Self-concept (SC) Models	116
Table 14	Hierarchical Models: Summary of the Goodness-of-fit Indices for the Sample of Girls ($N = 220$)	122
Table 15	Hierarchical Models: Summary of the Goodness-of-fit Indices for the Sample of Boys ($N = 205$)	124
Table 16	Latent Factor Correlations Among Academic Achievement Constructs and First-order Self-concept Constructs for the Sample of Girls ($N = 220$)	149
Table 17	Latent Factor Correlations Among Academic Achievement Constructs and Second-order Self-concept Constructs for the Sample of Girls ($N = 220$)	153
Table 18	Latent Factor Correlations Among Academic Achievement Constructs and First-order Self-concept Constructs for the Sample of Boys ($N = 205$)	155
Table 19	Latent Factor Correlations Among Academic Achievement Constructs and Second-order Self-concept Constructs for the Sample of Boys ($N = 205$)	159

PREADOLESCENT SELF-CONCEPT
AND SELF-CONCEPT/ACADEMIC ACHIEVEMENT RELATIONS:
INVESTIGATING MULTIDIMENSIONAL AND HIERARCHICAL STRUCTURES
WITHIN AND ACROSS GENDER

Overview of Self-concept Literature

Interest in self-concept (SC), broadly defined as individuals' perceptions of themselves (Shavelson, Hubner, & Stanton, 1976), has a long history among theorists and researchers in psychology, education, and educational psychology. Indeed, one of the earliest psychology texts devoted an entire chapter to the self (James, 1890/1950). Since James, theorists and researchers with widely disparate interests have examined the importance of SC as an outcome variable in a multitude of contexts (e.g., the home environment; Song, 1982), and a moderator variable of a variety of other outcomes (e.g., academic achievement; AA; Shavelson et al., 1976). Thus, SC has been assumed to be both an important outcome variable and an important moderator variable in educational and psychological research.

Prior to the 1980s, however, researchers were primarily interested in the SC construct because of its assumed relevance to variables specific to their area of research. In fact, there was a proliferation of studies examining the relation between SC and a multitude of other constructs (see Wylie, 1974, 1979). Despite a plethora of such studies, these earlier research efforts yielded inconsistent and indeterminate findings. The lack of clarity in this research was attributed to the manner in which the study of SC was approached (see, e.g., Byrne, 1984; Shavelson et al., 1976; Wells & Marwell, 1976; Wylie, 1974, 1979). That is, researchers generally examined SC in terms of its relation to variables relevant to their own particular area of interest without first investigating the SC construct itself.

According to measurement theory, the valid use of a construct requires its thorough investigation and delineation; a process referred to as the validation of a construct's nomological network (Cronbach, 1971; Cronbach & Meehl, 1955). Specifically, construct validation involves locating a construct within a theoretical framework in which both within- and between-network characteristics are examined. The within-network component of

construct validation establishes the nature of the construct under study such that the dimensions or facets of the construct, as well as the relations among them are identified. For example, within-network SC research might involve an examination of the relations among academic and social dimensions of SC. In contrast, between-network research focuses on the relations between the facets of one construct and those of some other construct with which it is theoretically related. Between-network investigations of the SC construct, for example, might focus on relations among various SC facets and other constructs with which it is posited to be related but from which it is conceptually distinct (e.g., between mathematics SC and mathematics achievement). The value of between-network research, however, is limited by the extent to which within-network applications are valid; that is, the extent to which the definition, measurement, and interpretation of a construct have been addressed (Shavelson et al., 1976). Thus, within-network research relevant to both the theory and the measurement of a construct is a necessary prerequisite to valid between-network research.

Limitations of Previous Self-concept Research

In systematic reviews of SC research conducted prior to the 1980s, studies were criticized for conducting between-network research before within-network issues had been addressed (for reviews of this research, see, e.g., Burns, 1979; Byrne, 1984; Shavelson et al., 1976; Wells & Marwell, 1976; West, Fish, & Stevens, 1980; Wylie, 1974, 1979). As a result, these earlier studies were plagued with conceptual and methodological shortcomings. Conceptually, SC definitions were vague and imprecise, and an explicit and empirically testable theoretical model of SC was not available in the literature. In the absence of an explicit model of SC, researchers did not have a clear conceptual framework to guide instrument development and investigations of construct validity. The lack of clarity at the conceptual level resulted in confusion at the methodological level. Specifically, SC measures were criticized for their psychometric inferiority (Shavelson et al., 1976; Wells & Marwell, 1976; Wylie, 1979).

Prior to the 1980s, little attempt was made to develop and validate SC instruments based on a priori-specified and empirically-derived scales. Instruments typically consisted of

an idiosyncratic collection of self-descriptive statements. Despite the widespread assumption that SC is multidimensional, SC instruments (e.g., The Coopersmith Self-Esteem Inventory; Coopersmith, 1967) generally yielded a total global score reflective of a unidimensional model of SC. In addition, many researchers constructed instruments for use in single studies. Attempts to examine the construct validity of SC instruments occurred after their development and were neither frequent nor promising. Factor analyses of these instruments, although frequently yielding multiple factors consistent with a multidimensional model of SC, typically did not identify consistent and readily interpretable facets (see Marsh & Smith, 1982; Shavelson et al., 1976). Similarly, multitrait-multimethod (MTMM) analyses characteristically found little evidence for discriminant/divergent validity, although support for convergent validity was generally stronger (e.g., Marx & Winne, 1978; Shavelson et al., 1976; Winne, Marx, & Taylor, 1977; a reanalysis of the Marx & Winne data can, however, be found in Shavelson & Bolus, 1982). The poor quality of SC instruments was linked to the fact that they were not developed based on (1) an explicit conceptual model of SC and, therefore, a priori-specified scales; and (2) examinations of their construct validity and, therefore, empirically-derived scales (Shavelson et al., 1976).

In summary, the absence of clarity at the conceptual and methodological levels of the within-network component of the SC construct resulted in between-network investigations that were seriously flawed, and that were characterized by weak, null, or inconsistent findings (e.g., Wylie, 1979). This confusion in the literature was due to the lack of attention to, and integration of theoretical and measurement issues. That is, the lack of consistent and definitive findings in earlier SC research were attributed to an absence of empirically-based conceptual models of SC, and theoretically-based and psychometrically sound measuring instruments (Shavelson et al., 1976; Wells & Marwell, 1976; Wylie, 1974, 1979).

Recent Advances in Self-concept Research

Since the 1970s, theoretical and empirical advances in our understanding of SC have occurred within the context of two research traditions, namely, within a cognitive or social cognitive information-processing perspective (e.g., Honess & Yardley, 1987; Markus, 1983,

1990; Markus & Sentis, 1982; Markus & Wurf, 1987; Nurius, 1986), or within the framework of measurement and/or construct validation (e.g., Marsh & Shavelson, 1985; Marsh, Smith, & Barnes, 1983, 1985). Within the context of information-processing research, SC facets are typically referred to as self-schemas, with self-schemas being conceptualized as multidimensional and dynamic cognitive knowledge systems (Markus, 1977, 1983; Markus & Kunda, 1986; Markus & Nurius, 1986, 1987; Markus & Sentis, 1982; Markus & Wurf, 1987; Singer & Kolligian, 1987). Empirical research within the information-processing tradition has focused on examining contextual influences on self-schemas, and on comparing groups differing in terms of the salience and elaboration of one or a few specific self-schemas. In particular, these research efforts have examined the implications of contextual factors and group differences in self-schemas for cognition, perception, affect, motivation, behavior, and/or clinical status (e.g., Dance & Kuiper, 1987; Hammen, Marks, deMayo, & Mayol, 1985; Higgins, 1987; Higgins, Bond, Klein, & Strauman, 1986; Higgins, Klein, & Strauman, 1985, 1987; Kuiper & Higgins, 1985; Linville, 1985, 1987; Markus & Kunda, 1986; Markus & Ruvolo, 1989; Nurius & Markus, 1990; Pietromonaco, 1985). Prominent among the theories within the information-processing perspective is Markus' notion of possible selves, in which SC is expanded to include both current and future conceptualizations of self (e.g., Markus, 1977, 1983; Markus, Cross, & Wurf, 1990; Markus & Nurius, 1986, 1987; Wurf & Markus, 1991).

In contrast to information-processing research, empirical research conducted within the framework of instrument or construct validation has focused on establishing SC measurement and structure, and the relations among SC facets and other constructs. For example, this research has begun delineating the multidimensional and hierarchical structure of SC (e.g., Byrne & Shavelson, 1996; Marsh, 1990d), and relations between SC and AA constructs (e.g., Byrne & Shavelson, 1986; Byrne & Worth Gavin, 1996; Marsh, 1992a). Thus, information-processing research has focused on different theoretical and empirical issues than has construct validation research. The differing focus of empirical efforts within each of these research traditions has entailed differing theoretical frameworks, as well as

differing strategies at the level of measurement, research design and data analysis. This situation has rendered the information yielded from each of these research areas less amenable to integration. Indeed, information-processing and construct validation efforts have remained separate in the theoretical and empirical literature. The purpose of the present study is to expand current construct validation research. Accordingly, the theoretical and empirical framework utilized is grounded in the tradition of construct validation research. We turn now to a review of theoretical and empirical advances within this research tradition.

Over the course of the last several decades, researchers have conducted within-network SC measurement and construct validation research addressing the conceptual and methodological problems inherent in earlier SC research. Conceptually, Shavelson et al. (1976) proposed a detailed theoretical model of multidimensional SC. This model represented the first explicit and empirically testable model of SC available in the literature. Since its inception, this model has undergone rigorous construct validation research and currently represents the most thoroughly investigated model of SC (Byrne, 1996a; Marsh, Byrne, & Shavelson, 1988). For this reason, the present SC study was conducted within the framework of the Shavelson model.

At the methodological level, a number of researchers have developed theoretically-based and empirically-derived multidimensional SC instruments (for a review of these and other multidimensional SC instruments, see Byrne, 1996b). Of particular relevance to the present study, Marsh developed and validated a series of instruments with which to test the validity of Shavelson's model. These instruments, the Self-Description Questionnaires (SDQs), are multidimensional SC instruments designed to measure SC during preadolescence (SDQ-I, Marsh, 1992b), early adolescence (SDQ-II, Marsh, 1992c), and late adolescence and young adulthood (SDQ-III, Marsh, 1992d). These research efforts have also led to the development of more specific multidimensional instruments designed to measure multiple dimensions of preadolescent and adolescent academic SCs (the Academic SDQ-I and II; ASDQ-I and -II, respectively; Marsh, 1990d), and of adolescent physical SC facets (the Physical SDQ; Marsh, Richards, Johnson, Roche, & Tremayne, 1994). These

instruments were developed based on the theoretical framework provided by the Shavelson model, and empirical procedures were used to validate and refine both the SDQ instruments and the Shavelson model of SC. Thus, the SDQ instruments provided researchers with age-specific theoretically-based and empirically-derived measures of the multiple facets of SC posited within Shavelson's model, as well as a wealth of research relevant to the construct validity of this model.

Although empirical validation of a measuring instrument is theoretically conducted independent of the validation of a theoretical construct, the two processes are, in practice, intertwined (Byrne, 1990b). That is, empirical investigations of one have implications for, and cannot proceed without some validity in the other. Validation of a theory is dependent upon the validity of the instrument(s) used to measure its construct(s), and the validity of a measuring instrument is closely linked to the validity of the theoretical model upon which it is based. Construct validation, therefore, involves the complementary processes of theory construction and instrument development. Accordingly, scale construction and validation efforts, particularly in relation to the SDQ instruments, have important implications for the validity of the Shavelson model of SC (e.g., Marsh et al., 1988; Marsh & Hocevar, 1985; Marsh & Shavelson, 1985).

Methodological advances have also occurred in relation to the statistical procedures used to examine the validity of SC measures and theory. For example, early research with the SDQ instruments utilized exploratory factor analytic techniques and correlations among scale (or factor) scores as the basis for evaluating the SDQ instruments and the Shavelson model of SC. As support for these instruments was demonstrated, the limitations of these statistical procedures were recognized (e.g., Marsh & Shavelson, 1985). That is, exploratory factor analyses and patterns of correlations among observed scale (or factor) scores do not permit researchers to determine the extent to which an hypothesized measurement or theoretical model provides a good fit to the sample data. These procedures are, therefore, not the most appropriate methodology for investigating the validity of measurement and theoretical models (for a discussion of this issue, see, e.g., Marsh, 1987b; Marsh & Hocevar,

1985; Marsh, Smith, & Barnes, 1985). Confirmatory factor analysis (CFA), however, overcomes many of the limitations inherent in correlational and exploratory factor analyses, and in doing so, provides a more powerful test of the validity of theoretical models, as well as the instrument(s) used to measure the constructs of interest. In particular, CFA permits researchers to (1) hypothesize a specific model and test the extent to which it provides a good fit to the data, (2) specify a series of models and compare the extent to which each model provides a good fit to the data, and (3) test the extent to which an hypothesized model is invariant across groups (e.g., across gender, grade, etc.). Thus, CFA procedures offer a more powerful test of the validity of both SC instruments and theoretical models.

The conceptual and methodological advances in SC theory and research achieved since the 1970s have permitted researchers and theorists a greater understanding of the SC construct. In contrast to earlier unidimensional measurement efforts, more recent reviews of theoretical and empirical SC literature emphasized the multidimensional structure of SC (e.g., Bracken, 1996; Harter, 1996), and of academic and nonacademic portions of the SC construct (e.g., Byrne, 1996a; Marsh, 1990a; Marsh & Hattie, 1996; Shavelson & Marsh, 1986). Researchers have also addressed the between-network component of SC validation; in particular, the relations between SC and AA measures (e.g., Byrne, 1986; Byrne & Shavelson, 1986; Marsh et al., 1988; Song & Hattie, 1985). However, the generalizability of the within- and the between-network components of these research efforts is limited for several reasons. The purpose of the present study was to address several of the shortcomings in our current understanding of the SC nomological network.

Limitations of Current Self-concept Research

Three limitations in current SC research are of particular relevance to the present study. First, research examining the structure of SC and its relation with AA has almost exclusively limited the study of the construct validity of the academic portion of the SC nomological network to mathematics and reading SC and achievement indices. As Marsh (1990d) noted, the limited breadth of academic areas examined has precluded tests of the adequacy of the Shavelson model of SC. Research examining the validity of an expanded

academic SC construct is characterized by several limitations. That is, this research has included neither multiple independent SC measures nor measures of nonacademic SC, was limited to samples of Australian preadolescent and adolescent boys (Marsh, 1990d), and between-network analyses of SC/AA relations were reported exclusively for the adolescent sample of boys (Marsh, 1992a). More recent instrument and construct validation efforts relevant to an expansion of the academic portion of SC have not, to the author's knowledge, been conducted; empirical research with the SDQ instruments has more recently focused on the physical component of adolescent SC (e.g., Marsh, 1996a, 1996b). Thus, there is a need to examine the adequacy of the Shavelson model of SC in explaining within- and between-network relations among an expanded set of academic SC and AA subject areas for both girls and boys, particularly among preadolescents.

Second, much of the research examining gender differences in SC has focused on substantive issues related to gender differences in the level (i.e., mean) of SC facets. Researchers have generally assumed an invariant SC measurement and structure, and invariant SC/AA relations across gender. An important aspect of construct validation involves determining the validity of a construct across gender. Unless the measurement and structure of a construct are equivalent across girls and boys, comparisons of observed means are methodologically inappropriate, and conclusions based on combined samples may not be valid (see, e.g., Byrne & Shavelson, 1987). Although researchers have examined empirically the validity of this assumption (e.g., Byrne & Shavelson, 1987; Marsh, 1987b; Marsh, Smith, & Barnes, 1985; Tanzer, Marsh, & Sim, 1992), these research efforts have not included an expanded set of academic SC and achievement indices. Accordingly, there is a need for methodological research that examines the measurement and structure of SC, and of SC/AA relations with an expanded set of SC and AA constructs within and across gender. The purpose of the present study was to examine this methodological aspect of the construct validity of SC for preadolescent (i.e., Grades 5 and 6) girls and boys.

Third, the research examining the validity of the Shavelson model of SC has been conducted almost exclusively with the SDQ instruments. The few research studies that have

examined the validity of the Shavelson model with instruments other than the SDQs have focused on limited portions of this model (e.g., Byrne & Shavelson, 1996; Byrne & Worth Gavin, 1996), and have utilized adolescent or college samples (e.g., Byrne & Shavelson, 1986; Fleming & Watts, 1980; Marsh et al., 1988). For example, recent research with Canadian samples was restricted to social SC facets (Byrne & Shavelson, 1996), or limited academic dimensions of the SC construct (Byrne & Worth Gavin, 1996). As a result, a comprehensive examination of the Shavelson model of SC was precluded. Thus, there is a need for research examining the construct validity of an expansion of the Marsh/Shavelson model with independent multidimensional SC instruments.

Overview of the Present Study

The purpose of the present study was to address current limitations in SC theory and methodology, and thereby, enhance our understanding of the nomological network of the SC construct. Given that construct validation research relevant to preadolescent SC is more limited than that relevant to adolescent SC, the present study focused on preadolescents. Specifically, the construct validity of SC and SC/AA relations within and across gender was examined for preadolescents within one conceptual framework using CFA procedures. The structure of SC, and the pattern of SC/AA relations was examined separately for Canadian girls and boys in Grades 5 and 6 for multiple nonacademic SC facets, and an expanded set of academic SC and achievement constructs. In addition, the invariance of the SC construct and of SC/AA relations across gender was empirically tested. The SC constructs examined included four nonacademic facets (i.e., physical appearance, physical ability, peers, and parents), a general-academic facet, and 11 subject-specific academic facets (i.e., Listening, Reading, Writing, and Speaking in Language Arts, as well as Arithmetic, Measurement, Science, Social Studies, Religion, Art, and Gym). The nonacademic SC constructs were measured using the SDQ-I (Marsh, 1992b) and, with the exception of Parent SC, the Self-Perception Profile for Children (SPPC; Harter, 1985). The academic SC constructs were measured using adaptations of the ASDQ-I (Marsh, 1990d) and the SPPC. Children's grades and self-reported grades in Listening, Reading, Writing, and Speaking in Language Arts, as

well as Arithmetic, Science, Art, and Gym represented the achievement constructs. Thus, the present study was designed to replicate, extend, and integrate previous theory and research relevant to the construct validity of the Shavelson model.

In order to provide the theoretical and empirical framework within which the present study was conducted, within- and between-network SC theory and research relevant to the Shavelson model will be reviewed. The Shavelson model of SC will be presented first as it provides the conceptual basis for the subsequent review of relevant construct validation research. Consistent with the construct validation process, within-network SC research will be reviewed first, and is followed by a review of between-network research examining SC/AA relations. The review of the literature then focuses on research investigating gender differences in the measurement and structure of SC, and of SC/AA relations.

Theoretical Framework: The Shavelson Model of Self-concept

Within the Shavelson model, SC was broadly defined as an individual's self-perceptions of competence in a variety of categories of experience (Shavelson et al., 1976). These self-perceptions were posited to arise out of one's environmental experiences. That is, self-perceptions were thought to be formed on the basis of environmental reinforcements and interactions with significant others (e.g., family members, peers, and teachers). The resulting self-conceptions were postulated to influence behavior, which in turn, influenced self-conceptions. Based on their review of SC theory and research, Shavelson et al. identified seven characteristics as central to the definition of SC. Specifically, SC was posited to be organized, multidimensional, evaluative, hierarchical, stable, developmental, and differentiable. These characteristics refer to the within-network nature of the SC construct, with the last characteristic also referring to the between-network component of SC.

According to the Shavelson model, individuals, in an attempt to make sense of, and simplify their experiences, organize their self-perceptions into categories. The particular categories utilized to organize self-perceptions were hypothesized to parallel the different domains of experience specific to people's lives, and more generally, the cultural group to which individuals belonged. Thus, the category system developed, in terms of both the

specific dimensions of self-conceptualization and their organization, was posited to be shared within groups of people sharing similar categories of experience. For example, elementary school-aged children might have specific academic SCs corresponding to the category system used to classify their academic performance (e.g., Social Studies and Religion). Likewise, paralleling the categories of their social experience, children might have SCs specific to their experiences with their family and with their peers. Thus, experiences in particular situations were hypothesized to yield self-perceptions specific to those areas or situations. Accordingly, the organization of SC was posited to be multidimensional, with the different categories of self-perceptions representing different dimensions, domains, or facets of SC. These multidimensional SCs were posited to include both a descriptive and an evaluative component.

Shavelson et al. (1976) further hypothesized that these multidimensional SCs were organized to form an hierarchical structure. This hierarchical structure was posited to be organized such that self-perceptions became increasingly general as the hierarchy was ascended. That is, self-perceptions in specific situations were thought to form the base of the hierarchy and to lead to situation-specific SCs (e.g., SCs relevant to mathematical ability, to physical ability, etc.). Together, these specific SCs formed an increasingly general hierarchy of inferences about the self in increasingly broader areas of experience (e.g., SCs in academic and nonacademic areas), with general SC at the apex of the hierarchy.

Shavelson et al. (1976) proposed a possible representation of this multidimensional hierarchical model of SC. This SC model, referred to as the Shavelson model, is presented schematically in Figure 1. Within this model, general SC was divided into academic and nonacademic SC at the second level of the SC hierarchy. Nonacademic SC was divided into social, emotional, and physical SCs, with each facet further divided into more specific facets at lower levels of the hierarchy. For example, physical SC was divided into physical ability and appearance SCs. Likewise, academic SC was divided into more specific subject-area SCs (e.g., Mathematics and Science); these subject-area SCs were further divided into specific areas within each subject area.

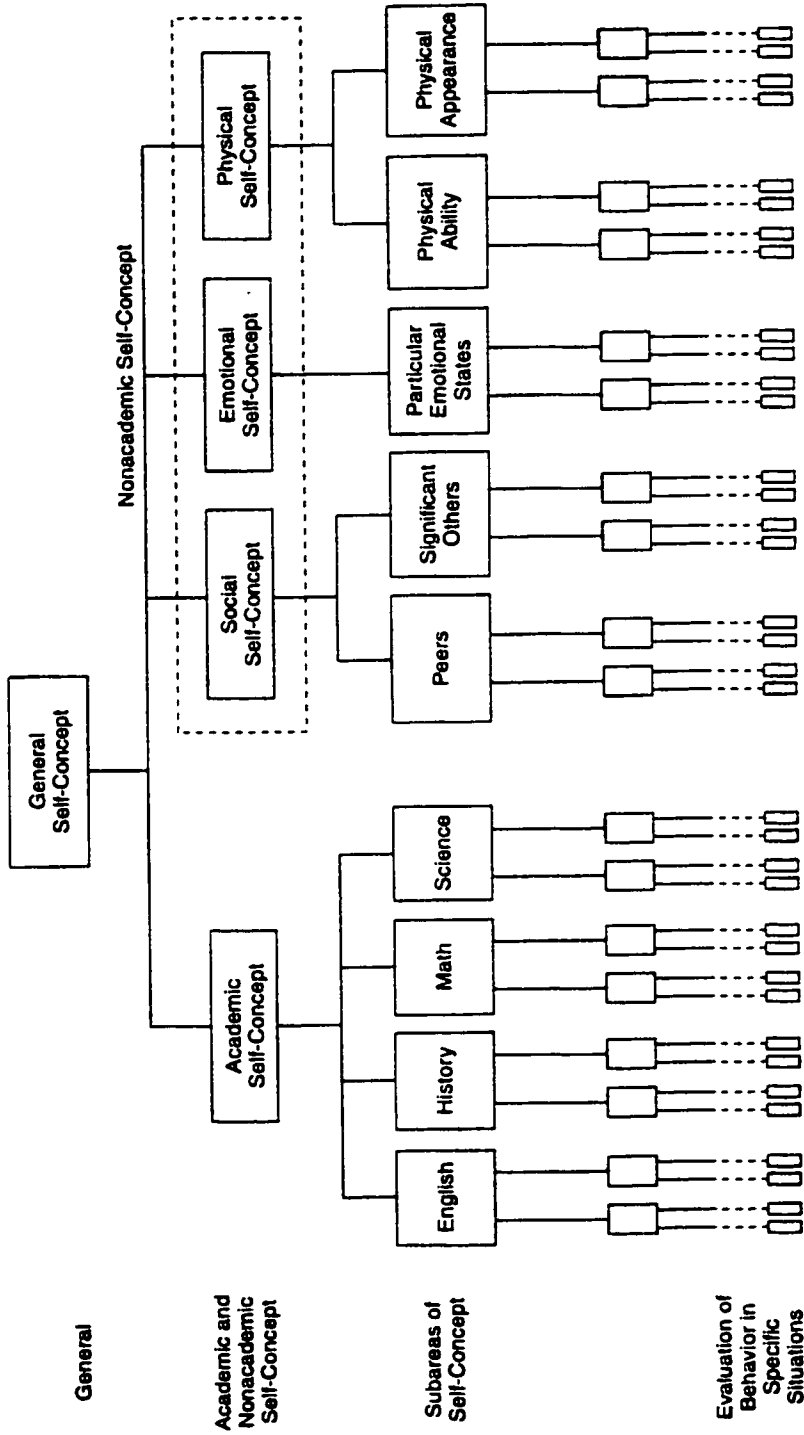


Figure 1. The representation of the multidimensional and hierarchical model of self-concept proposed by Shavelson and colleagues. From “Self-concept: Validation of construct interpretations” by R. J. Shavelson, J. J. Hubner, and G. C. Stanton, 1976, *Review of Educational Research*, 46, p. 413. Copyright 1976 by the American Educational Research Association. Reprinted with permission of the publisher.

In recognition of the greater specificity of SC at lower levels of the hierarchy, SC was posited to be increasingly linked to specific behaviors in particular situations as the base of the hierarchy was approached (Shavelson et al., 1976). Accordingly, the association between SC and behavior (i.e., external criteria) was hypothesized to increase as the hierarchy was descended. In a related vein, general SC was postulated to be stable, with SC facets becoming less stable at lower levels of the hierarchy. SCs at the base of the hierarchy were posited to be the least stable because of their increasing dependence on experiences in specific situations. However, changes in SCs at the base of the hierarchy were thought to be attenuated by SCs at higher levels, thereby enhancing the stability of more specific SCs. Thus, for example, the impact of a single failure experience in science on the science SC of an individual with high science and academic SCs would be expected to be attenuated by their academic SC. Increasingly numerous experiences incongruent with SC were believed to be necessary to change SC at higher levels within the hierarchy. Therefore, general SC was posited to be the most stable and, in contrast to SC facets at lower levels of the hierarchy, to require more experiences incongruent with SC to result in change; however, changes in lower level SCs were believed to be attenuated by SCs at higher levels in the hierarchy.

Shavelson et al. (1976) also posited that the structure of SC was developmental. That is, with increasing age and experience, the category system used to organize SC was posited to become increasingly differentiated. This increasing differentiation of the SC dimensions has been interpreted to mean that there may be increasing independence of SC facets and/or the development of additional facets with age (e.g., Marsh, 1990a; Marsh, Barnes, Cairns, & Tidman, 1984). Thus, the dimensions of SC may become increasingly independent, distinct, and differentiated, with new dimensions or facets possibly emerging with increasing age.

The final characteristic of SC postulated by Shavelson et al. (1976) was that SC is differentiable from other theoretically related constructs. Although Shavelson et al. elaborated on this characteristic of SC in relation to only the between-network component of SC, it has also been interpreted at the within-network level of construct validation (e.g., Marsh, Barnes, et al., 1984; Marsh & Hattie, 1996). Specifically, support for the

multidimensionality of SC requires that the multiple facets of SC be differentiable from each other, as well as differentiable from other theoretically related constructs; the former differentiation is relevant to within-network SC relations and the latter to between-network relations. However, the extent to which the different facets of SC are related to one another would be expected to be congruent with the hypothesized hierarchical structure of SC.

Within an hierarchical model, higher-order facets are meant to represent patterns of relations among lower-order facets, such that lower-order facets represented under the same higher-order facets would be expected to be more highly correlated amongst themselves than with other lower-order facets. Stated differently, correlations between SC dimensions represented under the same higher-order SC facet would be expected to be higher than the correlations between facets represented under different higher-order facets. Following from the Shavelson model, the highest correlations would be expected among the four academic SCs, between the two physical SCs, and between the two social SCs. For example, specific academic SCs would be expected to be more closely related to one another than to specific social, physical, or emotional SCs, and similarly, physical ability SC would be expected to be more strongly correlated with physical appearance SC than with academic SC.

With regard to between-network SC relations, Shavelson et al. (1976) hypothesized that the strength of the association between SC and other constructs paralleled the extent of the similarity of the situations with which each construct was linked. That is, area-specific SCs were expected to be distinguishable from, although related to theoretically relevant variables, such as achievement measures; however, the strongest associations were expected to occur between variables relevant to the same situations. For example, mathematics achievement was expected to be most strongly associated with mathematics SC, less correlated with other specific academic SCs, and least correlated with nonacademic SCs.

A similar principle of association has been hypothesized with regard to the relations between external criteria and SC facets at different levels of the SC hierarchy (e.g., Byrne & Worth Gavin, 1996; Marsh & Hattie, 1996). Consistent with the hypothesized hierarchically-linked specificity of SC, associations between external criteria and SC would be expected to

be higher when content areas are similar and specific than when they are less so. Thus, relations between external criteria (i.e., AA) and SC dimensions would be expected to decrease as the SC dimension of interest is drawn from increasingly higher levels of the SC hierarchy. For example, the correlation between mathematics achievement and mathematics SC would be expected to be higher than the correlation between AA and academic SC. Furthermore, mathematics achievement would be expected to be most highly correlated with mathematics SC, less correlated with academic SC, and least correlated with general SC. Such a pattern of correlations would support both the multidimensional and the hierarchical structure of SC hypothesized within the Shavelson model.

Of the seven SC characteristics proposed within the Shavelson model, only three are of central relevance (i.e., form the focus of the hypotheses tested) in the present study. These are that the SC construct is: (a) multidimensional, (b) hierarchically structured, and (c) differentiable. With regard to this latter characteristic, the within-network differentiability of the dimensions of SC, and the between-network differentiability of multidimensional SC and AA constructs are examined in the present study. Thus, construct validation research relevant to these SC characteristics are the focus of the literature review. Given that construct validation research has supported a developmentally-linked SC construct (see, e.g., Harter, 1990; Marsh, 1990a; Marsh & Hattie, 1996), the review of the literature was primarily restricted to that bearing on preadolescent SC, the developmental period relevant to the age group which is the focus of the present study (i.e., Grades 5 and 6 children).

Construct Validity Research: Preadolescent Self-concept

When the Shavelson model of SC was first proposed, there was little empirical support for its construct validity. However, subsequent instrument and construct validation research, particularly with the SDQ instruments, has provided a body of research relevant to the within- and the between-network component of the Shavelson SC model. The review of this literature begins with a review of within-network SC research and theory, and is followed by a review of between-network research and theory relevant to SC/AA relations.

Within-network Self-concept Research

The research reviewed in this section was designed to examine the validity of particular SC measuring instruments, and/or the validity of the Shavelson SC model. As previously noted, validation of a measuring instrument has important implications for the validity of the model upon which it is based. In particular, empirical information bearing on the validity of the Shavelson model can be obtained by extrapolating from instrument validation research. Although instrument development and validation efforts with a number of multidimensional SC instruments have provided information bearing on the validity of the Shavelson SC model, research designed to evaluate the Shavelson model of SC has almost exclusively been conducted with the SDQ instruments. In addition, research with the SDQ instruments has examined the validity of the Shavelson SC model with more content-specificity than has research using other SC instruments, particularly with regard to the academic portion of the SC construct. Accordingly, instrument development and construct validation research conducted by Marsh and colleagues with the SDQ-I (Marsh, 1992b), a preadolescent measure of multidimensional SC, is of particular relevance to the present investigation. Given the greater content-specificity of research examining the validity of the Shavelson SC model with the SDQ instruments, however, the review of research conducted by Marsh and colleagues was presented subsequent to the review of research conducted with less content-specific preadolescent SC instruments.

Given the conceptual and methodological difficulties identified with earlier SC research (e.g., Byrne, 1984; Shavelson et al., 1976; Wells & Marwell, 1976), research conducted with multidimensional measures designed and validated on the basis of theoretical models and empirical research will be the primary focus of the literature review. In a review of SC measures, Byrne (1996b) identified several preadolescent multidimensional measures of SC developed since the 1980s. Specifically, these were the Multidimensional Self Concept Scale (MSCS; Bracken, 1992), the SPPC (Harter, 1985), and two Self-Description Questionnaires, the SDQ-I (Marsh, 1992b) and the ASDQ-I (Marsh, 1990d). We turn now to a review of research with these measures.

Empirical Research with the Multidimensional Self Concept Scale

As with the development of the SDQ instruments, the MSCS (Bracken, 1992) was based on the model of SC proposed by Shavelson et al. (1976). Specifically, Bracken posited that SC is characterized by the seven features hypothesized in the Shavelson SC model. His model, however, more explicitly delineated and linked behavioral principles to the formation and maintenance of SC (Bracken, 1996; Crain & Bracken, 1994; Keith & Bracken, 1996). The MSCS measures 9- to 19-year-old children's SCs in six areas: (a) social SC, (b) a more specific family SC, (c) academic SC, (d) physical SC, (e) affective SC, and (f) competence SC in relation to the attainment of goals. Thus, the MSCS was designed to measure multidimensional SC facets, although the more specific academic SC facets identified in the Shavelson model of SC were not targeted.

The MSCS was developed, normed, and validated using a large representative sample ($N=2,501$) of Grades 5 to 12 children in the United States (Bracken, 1992). Bracken (1992) offered empirical support for the factorial, concurrent, convergent and discriminant validity of the MSCS based on the normative sample, as well as several independent research efforts (see, e.g., Crain & Bracken, 1994). Internal reliability estimates for the MSCS and its scales are reported to be high (i.e., in the 80s and 90s), as have test-retest reliability coefficients over a 4-week interval (i.e., values ranged from .73 to .90; see Bracken, 1996; Crain & Bracken, 1994; Keith & Bracken, 1996). Similarly, exploratory factor analyses of the MSCS and four other multidimensional SC instruments (Bracken, Bunch, Keith, & Keith, 1992) supported the six SC facets measured by the MSCS (see Crain & Bracken, 1994). Taken together, research efforts with the MSCS have supported the multidimensionality of preadolescent SC, albeit the SC dimensions measured have not included the specific academic SC facets delineated in the Shavelson model.

Empirical Research with the Self-Perception Profile for Children

Although not developed on the basis of the Shavelson model, scale construction and validation efforts with the SPPC instrument were conducted within a multidimensional developmental framework (Harter, 1982, 1983, 1985). That is, Harter hypothesized that

children, in forming their self-perceptions of competence, make distinctions among the different domains of experience in their life. Similar to the Shavelson model, Harter hypothesized that as children develop from childhood to adolescence, there is increasing differentiation among the domains of self-conception. Specifying these domains a priori, Harter (1982) developed the Perceived Competence Scale for Children (PCSC) to measure 8- to 12-year-old preadolescent self-perceptions of general competence, as well as SCs in the social, physical, and cognitive domains. On the basis of empirical research examining the validity of the PCSC, this instrument was revised, expanded, and renamed the SPPC (Harter, 1985). The SPPC contains additional scales measuring physical appearance and behavioral conduct SCs, with the PCSC physical scale designated a physical ability scale. Thus, Harter developed these instruments using principles and categories similar to those specified by Shavelson et al. (1976).

Construct validation research with the PCSC and SPPC has supported the hypothesized multidimensional character of SC. Separate exploratory factor analyses of the PCSC identified the social, physical, academic, and general SC facets for Grades 3 to 9 American children (Harter, 1982), as well as Grades 7 to 9 Australians (Marsh & Gouvenet, 1989). Exploratory (Marsh & MacDonald Holmes, 1990) and confirmatory (Marsh, 1990b) factor analyses of a Grade 5 Australian sample, as well as separate Grades 5 and 8 samples of normal and gifted Canadian children (Byrne & Schneider, 1988) identified the four facets of SC measured by the PCSC. Exploratory factor analyses also identified the five SPPC scales for three American samples of Grades 5 to 8 children (Harter, 1985), as did exploratory and confirmatory factor analyses of a Dutch translation of this instrument (Van Dongen-Melman, Koot, & Verhulst, 1993). Also consistent with the hypothesized developmental nature of SC, analyses suggested that preadolescent SC was less differentiated than for older children (Byrne & Schneider, 1988; Harter, 1985), and more differentiated than for younger (Harter, 1985) and less cognitively-able (Silon & Harter, 1985) children. Thus, scale construction and validation efforts with these instruments have supported a multidimensional and developmentally-linked SC structure, and research with

the SPPC has supported the separation of physical appearance and ability SCs.

Researchers have also examined the concurrent, convergent, and discriminant validity of the facets of SC measured by the PCSC. Using the SDQ-I scales corresponding to the PCSC scales, Byrne and Schneider (1988) demonstrated support for concurrent validity with both Grade 5 and Grade 8 gifted Canadian children, and Marsh and Gouvenet (1989) found support for convergent and discriminant validity with Grades 7 to 9 Australian children. Using an Australian sample of Grade 5 students, Marsh examined the construct validity of three preadolescent multidimensional SC instruments: the PCSC, the SDQ-I, and the Piers-Harris Children's Self-Concept Scale (PHSC; Piers, 1984). MTMM analyses, within both an exploratory (Marsh & MacDonald Holmes, 1990) and a confirmatory mode (Marsh, 1990b), supported the convergent and discriminant validity of multidimensional SC, and in particular, of general, social, academic, physical ability, and physical appearance SCs.

Research with the PCSC and SPPC instruments also supported the differentiability of the dimensions of SC. That is, the strength of the correlations among the SC facets for these studies were not so high as to suggest that the facets were indistinguishable, and therefore, non-differentiable. For example, for a sample of Grades 5 and 6 children, correlations ranged from .08 to .62 (Harter, 1985). Thus, these studies supported the multidimensionality and differentiability of SC, as well as the construct validity of the PCSC and SPPC as measures of preadolescent multidimensional SC.

The pattern of correlations among the SC facets measured by the PCSC and SPPC instruments for the studies reviewed above also provided information relevant to the hierarchical organization of SC. In particular, the pattern of correlations suggested a clustering of the physical, social and general SC facets, and a separate clustering of the academic and general SC facets. That is, the largest correlations were generally obtained among the general, social and physical SC facets, and among the general and academic SC facets (Byrne & Schneider, 1988; Harter, 1982, 1985; Marsh, 1990a; Marsh & Gouvenet, 1989; Van Dongen-Melman et al., 1993). Thus, these studies supported an hierarchical (i.e., higher-, or second-order level) differentiation of academic SC from the social and physical

SCs, with less support for the same differentiation between the social and physical SC facets.

The pattern of correlations among observed scale scores or factors derived through exploratory or confirmatory factor analyses, however, are not the most appropriate procedure for determining the validity of hierarchical models (Marsh, 1987c, 1990d; Marsh & Hocevar, 1985). A more powerful test of the adequacy of hierarchical models is provided by hierarchical CFA. Hierarchical CFA is conceptually similar to the application of CFA on the correlations among more specific lower-order (i.e., first-order) facets (e.g., physical ability and appearance SCs). Consistent with the advantages offered by CFA procedures, hierarchical CFA allows researchers to examine empirically (i.e., test statistically) alternate conceptualizations of the manner in which first-order SC facets combine to form more general second-order SC facets. Thus, although the pattern of correlations among first-order SC facets may suggest alternate hierarchical SC structures, hierarchical CFA offers a more appropriate and powerful test, and comparison of hierarchical structures. This form of analysis has not, however, been conducted with the PCSC and SPPC instrument scales.

Results of empirical research have suggested that a caution is in order with respect to the interpretation of results with Harter's SC instruments. On the basis of exploratory and confirmatory analyses of MTMM data of the SDQ-I, PCSC, and PHSC instruments for a sample of Grade 5 Australian children, Marsh (1990b; Marsh & MacDonald Holmes, 1990) concluded that there was evidence of a moderate method effect associated with the PCSC, although the largest method effect was found for the PHSC; there was little indication of a method effect with the SDQ-I. Marsh suggested that the complexity of the PCSC response format, a 4-point structured-alternative response format, might be related to the method effect found for this instrument. Consistent with this, a number of children did not respond appropriately to the PCSC, and analyses indicated that these children had modestly lower AA scores than did children who responded appropriately. As a result, Marsh suggested that the method effect might be related to verbal abilities and thus, that it might be larger among younger and less able students. Accordingly, he recommended that interpretations of the structure and level of differentiation of SC on the basis of these instruments be made

cautiously with children who are younger or less cognitively developed than preadolescents.

In summary, scale construction and validation efforts with the PCSC and SPPC have supported the multidimensionality of SC hypothesized within the Shavelson model and, in particular, the differentiability of general, academic, social, and physical appearance and ability SCs. Thus, in contrast to earlier SC research, but consistent with suggestions made in reviews of this research (e.g., Shavelson et al., 1976; Wylie, 1974, 1979), support for the multidimensionality of SC was demonstrated when instruments developed on the basis of a priori and empirically-derived scales were used. Research findings with these instruments also supported an hierarchically structured SC, with a greater level of differentiation being evident between the academic and nonacademic SC facets, than within the nonacademic social and physical SC facets; more powerful hierarchical CFA have not, however, been conducted. Finally, although not directly relevant to the present study, research with these instruments were consistent with the hypothesized developmental character of SC.

Research with the PCSC and SPPC instruments, as with that for the MSCS, has limited applicability to the validity of the Shavelson model. These instruments have been used to measure SC facets with less specificity than is necessary to more fully examine the multidimensionality of SC hypothesized within the Shavelson model; in particular, they have not been used to measure more specific academic SC facets. Therefore, research with these instruments, unlike that with the SDQ instruments, do not permit any conclusions regarding the validity of additional and more specific academic SC facets.

Empirical Research with the Self-Description Questionnaire-I

The SDQ-I was designed to measure multiple dimensions of SC consistent with the more specific components of the academic, social, and physical SC facets of Shavelson's model. It includes a total of eight scales, namely, a general SC scale, four nonacademic scales, and three academic scales (Marsh, 1992b). The four nonacademic scales were designed to measure specific SCs related to two physical (i.e., physical ability and physical appearance) and two social (i.e., relations with peers and relations with parents) SC facets. The three academic scales measure reading, mathematics, and general-academic (the All

School scale) SC facets. General SC is measured with the General-Self scale, a scale which was later added to the SDQ-I (Marsh, 1992b; Marsh, Smith, & Barnes, 1984). Thus, the design of the SDQ-I was consistent with Shavelson's model of SC in that it measures general and general-academic SCs, as well as specific nonacademic and academic SCs; the specific academic SC facets measured are restricted to reading and mathematics subject areas.

In a series of studies, Marsh and colleagues examined the factor structure, and the convergent and discriminant validity of the SDQ-I scales, as well as the construct validity of multidimensional SC. These studies utilized samples of preadolescent students from Sydney, Australia differing on several characteristics (Marsh, 1985; Marsh, 1992b). The samples differed in terms of their level of academic performance (ranging from below average to above average); socioeconomic status (ranging from lower to upper-middle class); type of school (e.g., coeducational public schools or private Catholic schools where students primarily attended single-gender classes); grade (i.e., Grades 2 to 6); primary language within the home (i.e., English or non-English); and geographic region within Sydney, Australia.

Exploratory and confirmatory factor analyses across Australian samples supported the SDQ-I scales and a multidimensional SC construct. Exploratory factor analyses provided support for the seven SC facets measured by the SDQ-I (Marsh, Barnes, et al., 1984; Marsh, Relich, & Smith, 1983; Marsh, Smith, & Barnes, 1983), in addition to general SC (Marsh, Smith, & Barnes, 1984). These results were replicated in analyses of the combined samples of Grades 2 to 6 children ($N = 3,562$) forming the SDQ-I normative data base (Marsh, 1985, 1992b). CFA reanalyses of earlier data (Marsh, Barnes, et al., 1984) substantiated these findings within and across Grades 2 to 5 samples (Marsh & Hocevar, 1985; also see Marsh & Shavelson, 1985; Shavelson & Marsh, 1986), and supported the invariance of factor loadings across grade levels. However, consistent with the hypothesized developmental feature of SC, CFA tests supported a decrease in the magnitude of the correlations among the facets with increasing grade level. MTMM analyses based on the Campbell-Fiske criteria (Campbell & Fiske, 1959) comparing student SC with peer, teacher, and parent ratings of

student SCs also supported the convergent and discriminant validity of the multiple SC dimensions measured by the SDQ-I (Marsh & Craven, 1991; Marsh, Parker, & Smith, 1983; Marsh, Smith, & Barnes, 1983, 1984). Taken together, these studies provided support for the SDQ-I as a measure of preadolescent multidimensional SC, and consistent with the Shavelson model, a multidimensional, hierarchical, and developmental SC construct.

Research with Canadian, British, as well as culturally diverse preadolescent samples has similarly provided support for a multidimensional SC construct. For example, exploratory (Smith and Marsh, 1985) and confirmatory factor analyses supported the invariance of the factor structure of preadolescent SC across Australian and British responses to the SDQ-I (Marsh & Smith, 1987). CFA of general, general-academic, English, and mathematics SC facets, conducted separately for Grades 3, 7, and 11 Canadian children, supported the multidimensionality of these SC facets (Byrne & Worth Gavin, 1996). Exploratory (Tanzer & Sim, 1991) and confirmatory (Tanzer et al., 1992) factor analyses of the SDQ-I with a sample of 10- and 12-year-old Singaporean children supported the eight SC facets measured; English was reported to be the language of instruction in school. Exploratory factor analyses with 11- to 13-year-old Nepalese (Watkins, Lam, & Regmi, 1991), Filipino (Watkins & Gutierrez, 1989), Nigerian (Watkins & Akande, 1992), and Zimbabwean (Watkins & Mpofu, 1994) children reported to be fluent in speaking and reading in English supported the six specific facets of SC measured by the SDQ-I; support for general-academic SC was weaker, with support for general SC facet being poor. Presumably, issues related to the primary language of the samples for the Watkins and colleagues studies played a role in the clarity of the obtained SDQ-I factor structure. CFA of an Arabic translation of the seven more specific SDQ-I scales conducted separately, and in combination for Grades 5 to 9 girls and boys also supported these multiple SC dimensions (Abu-Hilal & Aal-Hussain, 1997). Thus, with few exceptions, exploratory and confirmatory factor analytic results supporting the multidimensional SC facets measured with the SDQ-I were replicated across preadolescent samples from a variety of countries.

The correlations observed among the SC scales for the SDQ-I studies reviewed

above also supported the differentiability of these SC constructs. As with the PCSC and SPPC studies, the correlations among the multiple SC scales were not so high as to render them indistinguishable from each other. For example, the correlations among the SC scales ranged from .05 to .43 for a sample of Grade 5 children (Marsh, Barnes, et al., 1984). Thus, the correlations among the SDQ-I scales also supported the differentiability of multidimensional SC.

Marsh and Shavelson (1985) noted, however, that the pattern of correlations among the SC facets across the Australian samples forming the SDQ-I normative data base was somewhat different than that expected on the basis of the Shavelson model. First, the correlations between mathematics and reading SCs were consistently lower than the correlations between general-academic SC and either, or both mathematics and reading SCs. This was particularly so beginning with preadolescent (i.e., Grades 5 and 6) samples in that correlations between mathematics and reading SCs approached zero across Grade 4 to adult samples, with higher correlations having been obtained for younger samples (Marsh, 1986c). For example, the correlations between mathematics and reading SCs ranged from -.06 to .12 compared to correlations between general-academic SC and either mathematics or reading SCs, which ranged from .25 to .56 for three samples of Grades 5 and/or 6 preadolescents (Marsh, Barnes, et al., 1984; Marsh, Relich, & Smith, 1983; Marsh, Smith, & Barnes, 1983; for reviews, see Marsh, 1986c, 1992b). Second, the correlations between peer SC and the two physical SC facets were as high as, or higher than the correlation between the two physical SC facets. Finally, parent SC correlated as highly with some of the academic SC facets as it did with some of the nonacademic facets.

The pattern of correlations obtained across the Australian SDQ-I studies were consistent with an hierarchical SC structure but suggested that the hierarchy was somewhat different than in the Shavelson model. In particular, in comparison to the Shavelson model of SC, these studies suggested: (1) a lower level of differentiation among the nonacademic SC facets, (2) a higher level of differentiation among the academic facets, and (3) an association between parent SC and both the nonacademic and academic SC facets. The

correlations obtained among the SC facets for the British (Marsh & Smith, 1987), Singaporean (Tanzer & Sim, 1991), and Arabic samples (Abu-Hilal, & Aal-Hussain, 1997), as well as for the academic SC facets included in the study of Canadian students (Byrne & Worth Gavin, 1996) were consistent with the pattern observed for the Australian studies reviewed above; correlations were not provided for the studies conducted by Watkins and colleagues. Thus, the pattern of correlations obtained across a diversity of countries were consistent in suggesting that the hierarchical structure of SC might be different than that represented in the Shavelson model.

Taken together, these studies represent a valuable contribution to SC theory. First, they provided support for a multidimensional SC, and for the hypothesis that the multiple SC facets are differentiable from each other. Second, the pattern of correlations among the SC facets were consistent with an hierarchical SC but suggested that the hierarchy was less differentiated for nonacademic SC facets, and more differentiated for academic SCs than that proposed in the Shavelson model. However, the lower correlations between reading and mathematics SCs observed for preadolescents, as compared to younger children, supported the developmental characteristic of SC posited within the Shavelson model; that is, it suggested that SC becomes increasingly distinct across the primary school years (see Marsh, 1990a, for a review of research relevant to SC development with older subjects).

Based on research with the SDQ instruments, and in particular, the unexpected pattern of correlations observed among some of the SC facets, Marsh and colleagues (Marsh & Hocevar, 1985; Marsh & Shavelson, 1985; Shavelson & Marsh, 1986) suggested the need to revise the Shavelson model. However, the empirical basis for the proposed revisions were studies using exploratory factor analytic techniques and correlations among observed scale (or factor) scores. Marsh and colleagues noted that, despite the consistency of the results across samples, these procedures were not the most appropriate methodology for investigating the validity of theoretical models, and in particular, hierarchical models. Accordingly, they examined the validity of a series of alternate hierarchical models of preadolescent SC using a more powerful hierarchical CFA framework.

The Marsh/Shavelson Model of Self-concept: A Revision of the Shavelson Model

Marsh and colleagues (Marsh & Hocevar, 1985; see also Marsh & Shavelson, 1985; Shavelson & Marsh, 1986) compared the adequacy of a series of five hierarchical SC models hypothesized on the basis of the Shavelson model and empirical research. The hierarchical models included the four nonacademic (i.e., Physical Ability, Physical Appearance, Peer, and Parent SCs) and three academic (Reading, Mathematics, and General-Academic SCs) first-order SC facets measured by the SDQ-I. The models differed with respect to only the second-order level of the SC hierarchy, and included second-order SC facet(s) as follows: (a) a single General SC factor, (b) two second-order factors (i.e., Nonacademic and Academic SCs, or (c) three second-order factors (i.e., Nonacademic, and Mathematics-, and Verbal-Academic SCs). Two hierarchical models were specified for the two- and the three-factor models. One of the hierarchical models was consistent with Shavelson's model in that Parent SC loaded on only the second-order Nonacademic SC facet, and the other was consistent with empirical research with the SDQ instruments in that this SC facet loaded on both the nonacademic and the academic portion(s) of the hierarchy. In the two- and three-factor models, the second-order Nonacademic SC facet defined only nonacademic first-order SCs. Except for the single- or dual-loading of the Parent SC facet, the second-order Academic SC facet defined only the first-order academic SC facets in the two-factor model; in the three-factor model, the second-order Verbal- and Mathematics-Academic SC facets defined the first-order General-Academic SC, and its respective reading or mathematics SC.

Hierarchical CFA of the Marsh, Barnes, et al. (1984) data provided support for the superiority of the three-factor hierarchical model allowing the first-order Parent SC facet to be defined by both academic and nonacademic second-order SC facets for both separate and combined samples of Grades 2 to 5 children. Accordingly, Marsh and colleagues concluded that this model more accurately represented the hierarchical structure of preadolescent SC than did the Shavelson model (Marsh & Hocevar, 1985; Marsh & Shavelson, 1985; Shavelson & Marsh, 1986). This revised model of SC, the Marsh/Shavelson model, is presented schematically in Figure 2. As can be seen from Figure 2, the hierarchical structure

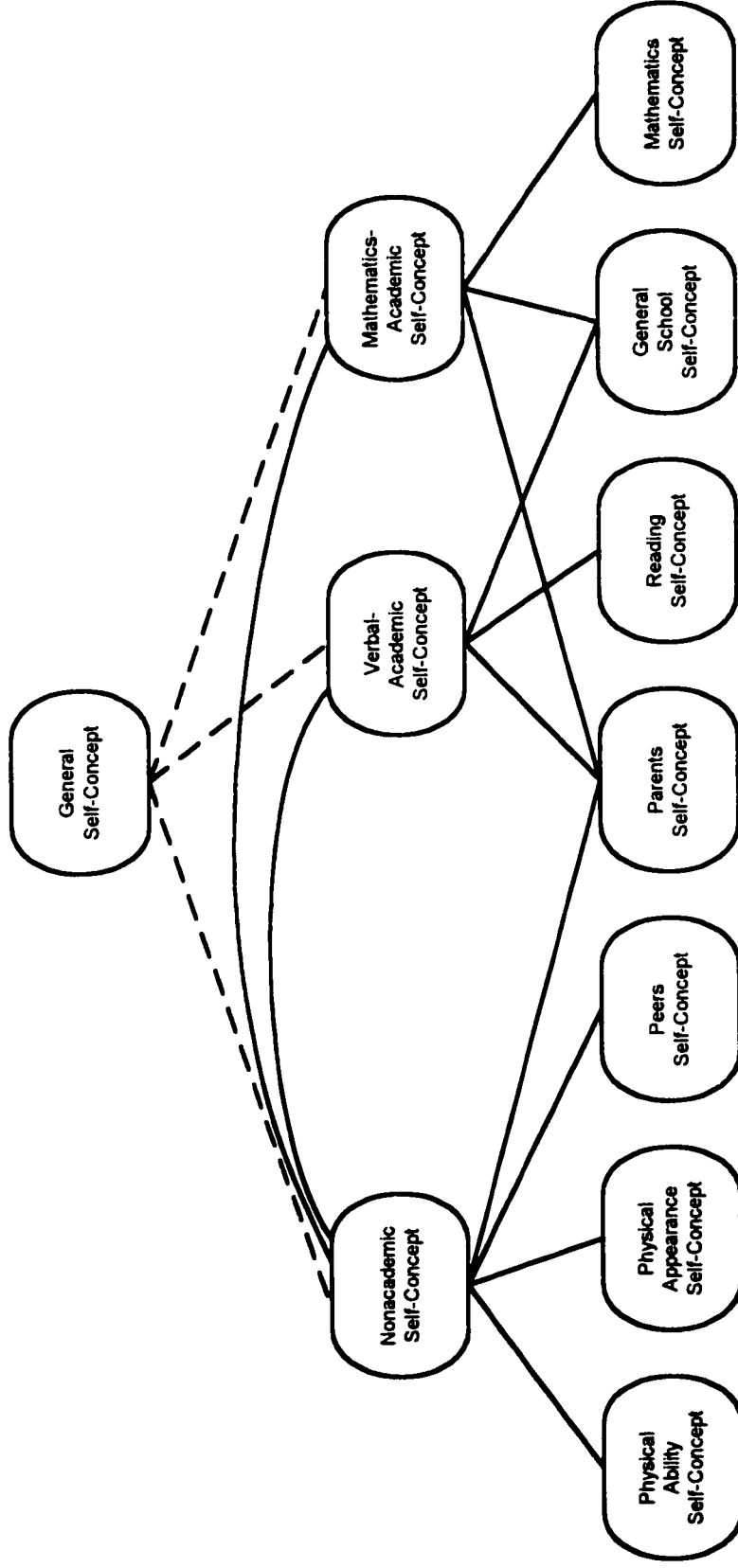


Figure 2. The Marsh/Shavelson model of self-concept. From "Self-concept: Its multifaceted hierarchical structure" by H. W. Marsh and R. J. Shavelson, 1985, *Educational Psychologist*, 20(3), p. 114. Copyright 1985 by Lawrence Erlbaum Associates. Adapted with permission of the publisher.

of SC proposed in the Marsh/Shavelson model is somewhat different than that in the Shavelson model. In particular, the single second-order academic SC facet in the Shavelson model is more differentiated, being represented by two, rather than one second-order academic SC (i.e., Verbal-, and Mathematics-Academic SCs). The second-order academic SC facets are linked to Parent SC, General-Academic (i.e., General School) SC, and their respective academic subject area SC. Nonacademic SC is represented as a single second-order SC facet and defines Physical Appearance, Physical Ability, Peer, and Parent SCs. Thus, this revised model incorporated the low correlations observed empirically between Mathematics and Reading SCs, as well as the higher than expected correlations obtained among the nonacademic SC facets, and between the Parent and academic SC facets.

The research conducted by Marsh and colleagues (Marsh & Hocevar, 1985; Marsh & Shavelson, 1985; Shavelson & Marsh, 1986), however, represented the only attempt to examine alternate conceptualizations of nonacademic and academic SC structure using hierarchical CFA procedures with samples relevant to the present study. That is, to the author's knowledge, research of this nature has been conducted only with an Arabic preadolescent sample (Abu-Hilal & Aal-Hussain, 1997) and with adolescent samples (e.g., Marsh, 1987c). Abu-Hilal and Aal-Hussain (1997) argued for cultural differences in SC between preadolescents from Arabic as compared to Western cultures, particularly in regard to gender differences in SC differentiation. Consistent with this, their study yielded support for a less-differentiated SC structure, particularly for boys, than that obtained by Marsh and colleagues and suggested by SDQ-I research with Canadian (Byrne & Worth Gavin, 1996), or Australian and British (e.g., Marsh & Smith, 1987) samples. Results of hierarchical CFA of the SDQ-III for Australian adults were consistent with those obtained for Australian preadolescents (Marsh, 1987c). That is, support for the second-order separation of mathematics and verbal SC facets was demonstrated, and support for the second-order separation of the nonacademic facets was weak; these results were replicated across data collected at two time periods (for SC research relevant to older subjects, see, e.g., Marsh, 1989b, 1990a; Marsh et al., 1988; Marsh, Parker, & Barnes, 1985). However, the

generalizability of the results of these studies, as compared to those conducted by Marsh and colleagues with preadolescent samples, to the present study are considered less valid given potential cultural or developmental differences.

Although research with the SDQ instruments provided support for the validity of the Marsh/Shavelson model, Marsh and colleagues (Marsh, 1990d; Marsh et al., 1988) argued that the adequacy of this model had not been fully assessed. Specifically, they noted that empirical support for the Marsh/Shavelson model was limited to research examining first-order mathematics, reading, and general-academic SC facets. Accordingly, Marsh and colleagues expressed the need for research examining the validity and adequacy of the academic portion of the Marsh/Shavelson model in explaining correlations among an expanded set of first-order academic SC facets. In addition, Marsh (1990d) hypothesized that the two second-order academic SC facets in the Marsh/Shavelson model were insufficient to explain correlations among an expanded set of first-order academic SC facets, and suggested the need to expand the academic portion of the Marsh/Shavelson model. Prior to a review of theory and research relevant to an expansion of the Marsh/Shavelson model, a model proposed by Marsh (1986c) to explain the unexpected differentiation within the academic portion of the SC hierarchy will be reviewed.

The Internal/External Frame of Reference Model

In an attempt to provide a theoretical explanation of the unexpectedly low correlations obtained between mathematics and reading SCs, Marsh (1986c) proposed the Internal/External Frame of Reference Model (I/E model). According to this model, mathematics and reading SCs are derived from more than self-perceptions of skill and performance in the corresponding academic area. The formation of specific academic SCs was hypothesized to be more complex, and to occur within a broader context than was suggested by a simple direct relation between subject-specific SC and AA indices.

Within the I/E model, external and internal comparisons were posited to be involved in the development of mathematics and reading SCs. Marsh (1986c) hypothesized that students, in forming their specific SCs in these academic areas, engaged in an external

comparison process in which they compared self-perceptions of their abilities in each of these academic areas (i.e., mathematics and reading) against those of other students in the corresponding academic area (i.e., mathematics and reading, respectively); the particular students against whom comparisons were made was referred to as the frame of reference. Thus, mathematics and reading SCs were posited to be influenced by self-perceptions of both one's mathematics and reading abilities, as well as the corresponding ability levels of students in one's frame of reference. Marsh suggested that this external comparison process exerted a positive association between mathematics and reading SCs given that these same AA indices were positively correlated. That is, a correspondence in the ability levels for the two academic areas relative to other students contributed to a linear correspondence between the SCs in the two areas. Thus, on the basis of the external comparison process alone, a positive correlation would be expected between mathematics and reading SCs.

The internal comparison process, however, was posited to attenuate the positive correlation between mathematics and reading SCs predicted by the external comparison. Specifically, the internal comparison process specified within the I/E model suggested that students also compare self-perceptions of their mathematics and reading ability against each other. As a result, the difference between the two ability areas was expected to contribute to a higher SC in one of the academic subject areas and a lower SC in the other academic area. Accordingly, a negative relation was exerted on mathematics and reading SCs. The greater the perceived difference between reading and mathematics abilities, the greater the magnitude of the negative relation, and thus, the more a high SC in one area was expected to result in a lower SC in the other academic area. Thus, the internal comparison process, on its own, was expected to result in a negative correlation between reading and mathematics SCs.

The operation of both the internal and the external comparison processes was hypothesized to underlie the near-zero correlations obtained empirically between mathematics and reading SCs. The relative strength of each of these comparison processes was expected to determine the magnitude of the correlation between these two SCs. On the basis of this model, Marsh (1986c; Marsh et al. 1988) posited a pattern of relations between

mathematics and reading SCs and achievement indices which predicted that mathematics and reading SCs were uncorrelated, or substantially less correlated than their corresponding achievement indices. The I/E model, therefore, required only that mathematics and reading SCs be less correlated than their corresponding achievements. Thus, the comparison processes delineated in the I/E model provided a theoretical explanation of the lower than expected correlations observed empirically between these two SCs.

Marsh also provided a developmental explanation for the higher correlations between mathematics and reading SCs observed for younger, as compared to preadolescent and older samples. Specifically, he suggested that the younger children had not yet fully developed the cognitive capabilities to engage in, and integrate the information derived from the comparison processes specified within the I/E model (see Marsh, 1986c, 1990a, for a discussion of these issues). Hence, the predictions associated with these comparisons processes were attenuated for younger children.

Although investigation of the I/E model was not the focus of the present study, a body of research has provided support for the hypotheses specified in this model. In a review of SDQ studies involving preadolescent, adolescent, university, and adult samples ranging from 7 years of age to over 35 years of age ($N = 6,010$), Marsh (1986c) demonstrated empirical support for the I/E model (see also Marsh, 1984, 1987a, 1989b, 1990c; Marsh & Parker, 1984). Support for the predictions hypothesized within the I/E model has been demonstrated with preadolescent and/or adolescent samples of Australian (e.g., Marsh, 1986c), American (e.g., Marsh, 1989b), Canadian (Byrne & Worth Gavin, 1996; Marsh et al., 1988), and Singaporean (Tanzer et al., 1992) children (for reviews of research relevant to this model, see, e.g., Marsh, 1990a, 1992b, 1993a). We turn now to a review of research relevant to the academic expansion of the Marsh/Shavelson model.

The Expanded Marsh/Shavelson Model: An Expansion of Academic Self-concept

On the basis of the academic scales of the SDQ instruments, Marsh (1990d) developed the ASDQ-I and II for preadolescent and adolescent children, respectively, in order to tap a range of academic SCs; these instruments were administered to a sample of

Australian boys in Grades 5 and 6, and in Grades 7 to 10, respectively. The ASDQ-I contained 13 scales, one measuring General-Academic SC and 12 measuring subject-specific academic SCs (i.e., SCs in Spelling, Reading, Handwriting, Social Studies, Computer Studies, Science, Mathematics, Physical Education, Art, Music, Religion, and Health). The ASDQ-II for adolescents contains 16 scales measuring General-Academic SC and 15 subject-specific academic SC facets (i.e., the last eight SC facets measured by the ASDQ-I, in addition to SCs in English Language, English Literature, Foreign Languages, History, Geography, Commerce, and Industrial Art). Differentiating between core and non-core academic areas, Marsh considered all subject-specific academic SCs but the Physical Education, Art, Music, Religion, and Health SC facets of the ASDQ-I and II, as well as the Industrial Art SC facets of the ASDQ-II to be core academic SCs.

Separate analyses of the ASDQ-I and II instruments provided support for the multidimensionality of academic SC (Marsh, 1990d). With one exception, exploratory and confirmatory factor analyses of the ASDQ instruments supported the multidimensionality of the a priori SC facets measured with each of these instruments; exploratory factor analysis did not support the differentiation of the English Language and English Literature scales of the ASDQ-II. Thus, results supported a greater level of academic SC differentiation for both preadolescent and adolescent boys than had previously been examined, thereby suggesting that children differentiate among a wide variety of academic subjects areas.

Hierarchical CFA of the core academic first-order SC facets indicated that the two higher-order SCs in the Marsh/Shavelson model provided a better fit to the data than did the single higher-order academic factor in the Shavelson model. However, results indicated that the second-order Mathematics-Academic SC facet was more accurately represented as a science factor; the factor loadings for the Science, Social Studies, and Computer Studies SC factors were higher on this second-order SC facet than was Mathematics SC. Thus, the superiority of the Marsh/Shavelson model was supported over that of the Shavelson model although the second-order Mathematics-Academic factor was relabeled a science factor.

Consistent with expectations, CFA of the complete set of first-order academic SCs

suggested the need for additional second-order SC facets for both the preadolescent and adolescent samples (Marsh, 1990d). In particular, second-order Physical Education and Art SCs were posited to be necessary to more completely represent the hierarchical structure of SC. In positing this expanded model of academic SC, Marsh assumed that the core first-order academic SCs were more central to the second-order Science- and Verbal-Academic SCs than were the remaining first-order academic SCs. Accordingly, only the General-Academic and core first-order academic SC facets were hypothesized to combine to form the two core second-order Science- and Verbal-Academic SC facets. The remaining SC facets were hypothesized to be defined by the less academically central second-order Physical Education and Art SC facets. Thus, the expanded academic portion of the Marsh/Shavelson model of SC contained four, rather than two second-order academic SC facets (see Figure 3 for a pictorial representation of this model for preadolescents). As revealed in Figure 3, the second-order Science-Academic SC was divided into General-Academic, Science, Social Studies, Computer Studies, and Mathematics SCs. Verbal-Academic SC defined General-Academic, Reading, Spelling, and Handwriting SC facets. The second-order Physical Education SC represented Physical Education and Health SCs whereas the second-order Art SC facet defined SCs in Art, Music, and Religion.

Hierarchical CFA of the first-order SC facets substantiated the need for the two additional second-order academic SC facets hypothesized in the expanded Marsh/Shavelson model for both the preadolescent and adolescent samples (Marsh, 1990d). Post hoc sensitivity analyses, however, indicated that the addition of a posteriori parameters allowing some overlap between the core and non-core academic SC facets improved model fit. Specifically, for the preadolescent sample, the a posteriori parameters included allowing (a) the Health SC facet to be represented under the Science-Academic SC facet, (b) the Religion and Music SC facets to be defined by the Verbal-Academic SC facet, and (c) the Handwriting SC facet to be defined by the second-order Art SC facet. These results provided support for a more differentiated multidimensional preadolescent (and adolescent) academic SC than previous research had indicated; children appeared to differentiate among subject

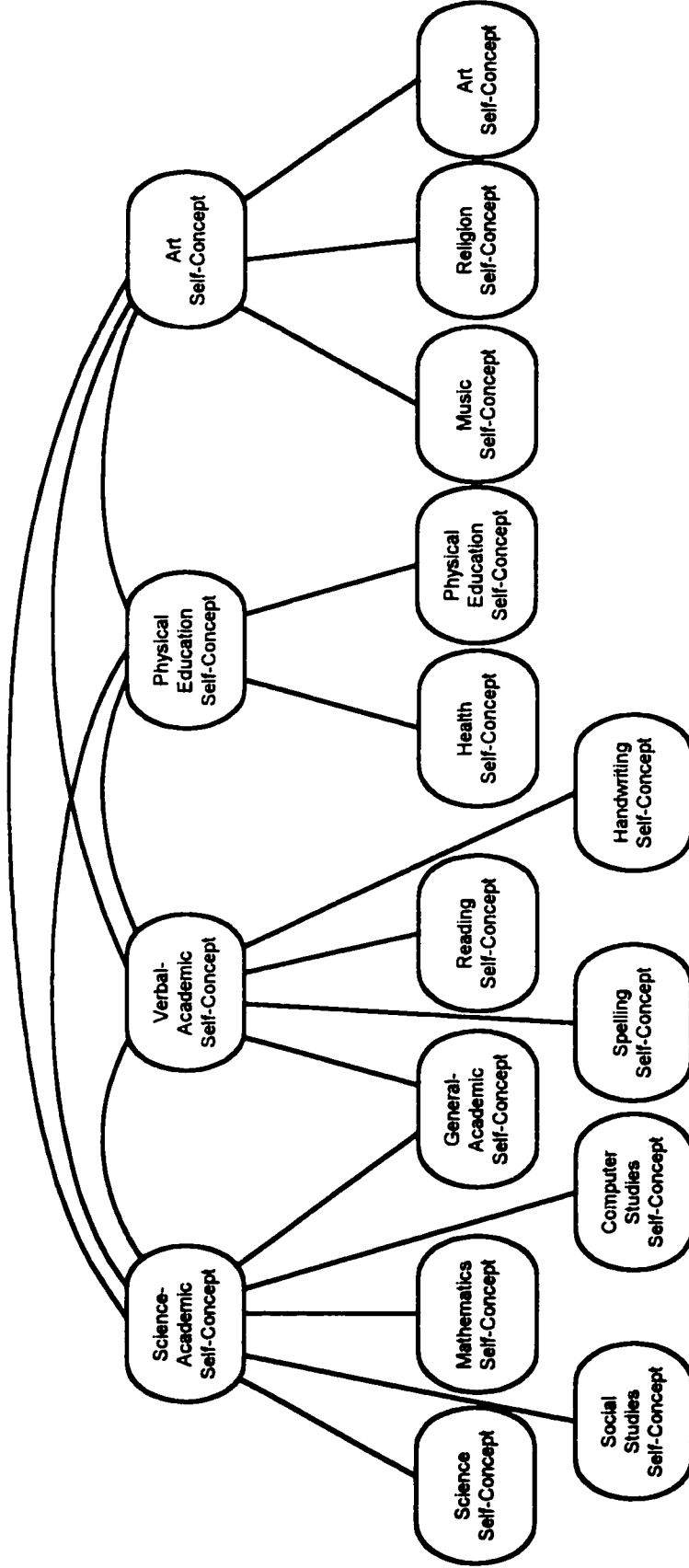


Figure 3. A pictorial representation of the academic expansion of the Marsh/Shavelson model. From "The structure of academic self-concept: The Marsh/Shavelson model" by H. W. Marsh, 1990d, *Journal of Educational Psychology*, 82(4), pp. 623-636.

areas for which they received grades. Moreover, the results indicated the need to expand the Marsh/Shavelson model in order to incorporate additional academic subject areas that had not previously been considered.

The near-zero correlations obtained between mathematics and reading SC facets in previous SDQ research, however, were not replicated (Marsh, 1990d). In particular, for preadolescent boys, Mathematics SC was more highly correlated with Reading and Spelling SC facets ($r_s = .21$ and $.30$, respectively) than in previous research, although these three facets were correlated more highly with General-Academic SC ($r_s = .70$, $.57$, and $.62$, respectively); a similar pattern of results was obtained for the adolescent sample. Nevertheless, these findings were consistent with the I/E model and the second-order differentiation of Mathematics- and Verbal-Academic SC facets. Moreover, Marsh suggested that these higher correlations were due to a generalization of the internal comparison process posited in the I/E model to other academic areas, with the multiple comparisons attenuating the predictions of this model. For example, students may have compared their ability in mathematics with their ability in both reading and spelling. Thus, Marsh suggested the possibility that predictions based on the I/E model may be accentuated when only Reading and Mathematics SC facets are considered.

As Marsh (1990d) noted, the academic expansion of the Marsh/Shavelson SC model must be viewed cautiously. This study represented the only attempt to examine an expanded set of subject-specific academic areas. Research is needed to determine whether the higher correlations obtained between Mathematics and Reading SCs in this study are replicated when an expanded set of academic subject areas are considered. In addition, this construct validation study included only boys. There is a need to examine the construct validity of an expanded model of academic SC for both girls and boys. In doing so, researchers cannot assume that SC is invariant across gender. The structure of SC must be examined for boys and girls separately, and the invariance of the measurement and structure of SC across gender tested.

A further limitation of the study of an expanded academic SC construct (Marsh,

1990d) is that it was restricted to the academic portion of the model. As Marsh (1990d) noted, it may be that some of the less academically central (i.e., non-core) academic SC facets measured by the ASDQ-I (e.g., Physical Education) would more appropriately be represented by a second-order Nonacademic SC facet. If this were the case, the structure of the academic portion of the expanded Marsh/Shavelson model, as well as the nonacademic portion of the SC hierarchy would be different than that suggested in the expanded Marsh/Shavelson model. The examination of nonacademic SCs and an expanded set of academic SC facets within one conceptual framework would allow a more complete test of the hierarchical structure of SC. Thus, there is a need for research examining the validity of the expanded Marsh/Shavelson model within a conceptual framework that includes the nonacademic portion of this model.

In addition, several of the first-order SC facets examined in the study conducted by Marsh (1990d) were not well represented by the second-order facets. In this regard, three aspects of the expanded Marsh/Shavelson model are of particular relevance to the present study. First, Handwriting SC was not well represented under the second-order Verbal-Academic SC facet, although an a posteriori change that also represented this facet under the Art SC facet led to a better representation of the first-order Art SC facet. Perhaps the Handwriting SC facet measured was different than a Writing in Language Arts/English SC facet. In particular, the results of the Marsh (1990d) study suggested the possibility that Handwriting SC was less a reflection of self-perceived verbal writing skills, and more a reflection of concrete/manual aspects related to neatness, and perhaps some component of artistic expression through this medium. A Writing in Language Arts SC scale may more appropriately tap verbal SC aspects related to writing skills than did a Handwriting SC scale, and thus, may be well represented under the second-order Verbal-Academic SC facet.

The Music and Religion SC facets were also not particularly well represented under the second-order Art SC facet (or the second-order Verbal-Academic SC facet in the a posteriori expanded Marsh/Shavelson model). The appropriateness of the representation of Religion SC under a social/moral second-order SC facet needs to be explored given that

religious education also occurs within, and has implications for social contexts, particularly the family; this altered representation may similarly lead to an enhancement in the representation of the first-order SC facets represented under the second-order Art SC facet. In addition, although the first-order Science and Social Studies SC facets were well represented under the second-order Science-Academic SC facet, the first-order Mathematics and Computer Studies SC facets were less well represented. Perhaps the second-order separation of science and mathematics SC facets, as with the separation between mathematics and verbal SC facets, would yield a better fitting hierarchical model. Each of these possibilities was examined with the hierarchical models tested in the present study.

A third limitation of the construct validation study of an expanded academic SC construct (Marsh, 1990d) is that indices of AA were not included in analyses. Although between-network SC/AA data were later reported for the adolescent sample (Marsh, 1992a), construct validation issues examining SC/AA relations among an expanded set of academic SC facets have not, to the author's knowledge, been conducted with preadolescents. Thus, there is a need for between-network research examining the pattern of SC/AA relations among a broader set of subject-specific indices. The present study was designed to address this need for preadolescents. We now turn to a review of between-network research bearing on the validity of the Marsh/Shavelson model of SC.

Between-network Research: Self-concept/Academic Achievement Relations

The assumption that SC is related to achievement was a primary impetus to the plethora of earlier SC research. In fact, AA is one of the constructs that has most frequently been examined in between-network studies of SC. As previously noted, however, between-network research conducted prior to the recent conceptual and methodological advances in SC research yielded disappointing results, and contributed little to our understanding of the SC construct (e.g., West et al., 1980; Wylie, 1979). These earlier studies focused on general or academic SC, and reviews (e.g., Byrne, 1984; Wylie, 1979) and meta-analyses (e.g., Hansford & Hattie, 1982) of this research revealed wide discrepancies in the magnitude of SC/AA relations, as well as null and contradictory findings.

With the recent conceptual and methodological advances in SC research, there is general consensus that relations between SC and AA cannot be understood without attention to the multidimensionality of academic SC (e.g., Byrne, 1996a; Marsh, 1990a; Marsh & Hattie, 1996). Accordingly, between-network research has increasingly examined the relations between more specific multidimensional AA and SC facets. However, preadolescent research examining relations between subject-specific indices of both academic SC and achievement have been conducted exclusively with the SDQ instruments. As a result, much of the between-network SC research to be reviewed in relation to AA indices is limited to studies using the SDQ instruments.

In between-network investigations, support for the construct validity of multidimensional SC, and in particular, multidimensional academic SC imposes two requirements upon the nature of SC/AA relations. First, empirical investigations must demonstrate that measures of SC can be discriminated from measures of AA. That is, unless subject-specific academic SC facets are differentiable from indices of AA, a counterinterpretation of these SC facets is that they are merely an alternative measure of AA (e.g., Byrne & Shavelson, 1986; Marsh, 1986c; Shavelson & Bolus, 1982). Thus, support for the discriminability of subject-specific academic SCs from subject-specific AA indices is crucial in demonstrations of the construct validity of multidimensional academic SC. Second, support for a multidimensional academic SC requires that corresponding subject-specific SC and AA constructs be more positively correlated than non-corresponding SC and AA indices. For example, mathematics achievement should be most correlated with mathematics SC, and reading achievement should be most correlated with reading SC. Thus, support for the construct validity of the Marsh/Shavelson multidimensional model of SC requires empirical support for the differentiability of multidimensional SCs and academic SCs, as well as the content-specificity of subject-specific SC/AA relations.

Support for the between-network construct validity of an hierarchical SC, and in particular, academic SC, imposes the additional requirement that SC facets demonstrate a consistent and logical/theoretically defensible pattern of relations with indices of AA. Within

the context of the expanded Marsh/Shavelson model, support for the construct validity of hierarchical SC requires that SC/AA relations be consistent with its hypothesized structure. The pattern of SC/AA relations can be examined at two levels. First, support for the hierarchical structure of SC hypothesized within the expanded Marsh/Shavelson model can be examined in relation to the pattern of relations between AA indices and first-order SC factors. First-order between-network support for this model requires that subject-specific AA indices be (a) most positively correlated with their corresponding academic SCs; (b) less correlated with other specific academic SC facets hypothesized to be represented under the rubric of the same second-order SC facet; (c) even less, or uncorrelated with other academic SC facets represented under the rubric of other second-order SC facets; and (d) least correlated, uncorrelated, or negatively correlated with nonacademic SCs. For example, mathematics achievement should be most correlated with mathematics SC, less correlated with science SC, even less or uncorrelated with reading SC, and least correlated, uncorrelated, or negatively correlated with nonacademic SCs (e.g., physical appearance SC).

Support for the hierarchical structure of SC hypothesized within the Marsh/Shavelson model can also be examined by comparing the magnitude of the relations between AA indices and first-order SC factors against the magnitude of the relation between AA indices and SC facets at increasingly higher levels of the SC hierarchy. Higher-order between-network support for the Marsh/Shavelson model of multidimensional hierarchical SC requires that subject-specific AA indices be (a) most positively correlated with their corresponding first-order academic SC, (b) less correlated with their corresponding second-order SC facet, (c) even less or uncorrelated with their corresponding third-order SC facet (i.e., General SC in the Marsh/Shavelson model). Empirical research on the basis of these data, however, has not, to the author's knowledge, been conducted. However, correlational data are available between AA and increasingly general SC facets measured as either scale (or factor) scores, or latent SC factors (i.e., factors derived on the basis of CFA). Accordingly, empirical support for the hypothesized hierarchical structure of SC was examined with these data by contrasting the magnitude of correlations between AA indices

and increasingly general SC facets. Within this framework, support for the hierarchical structure would be demonstrated if subject-specific AA indices were most correlated with their corresponding subject-specific SC facet, less correlated with a more generic academic SC facet, and least correlated with the most generic of SC facets (i.e., general SC).

We turn now to a review of research bearing on the between-network component of the validity of multidimensional SC, and in particular, multidimensional academic SC. Empirical research measuring a non-specific academic SC facet will be reviewed first, followed by a review of research examining subject-specific academic SC facets. Although the magnitude of relations between subject-specific SC/AA indices, as compared to those between subject-specific SC facets and more general academic SC facets bear on the validity of multidimensional subject-specific academic SCs, these indices are most relevant to the hierarchical structure of SC. Accordingly, research relevant to this aspect of construct validation will be reviewed under the rubric of research bearing on the validity of an hierarchically structured SC.

Multidimensional Self-concept/Academic Achievement Relations

Between-network research with the SPPC has provided support for the differentiability of academic SC from subject-specific AA indices. For example, Sink, Barnett, and Pool (1993) examined relations between the academic scale of the SPPC and multiple subject-specific AA measures in mathematics, reading, science, and social studies, as determined by standardized tests, for an American Grade 6 sample. Correlations between academic SC and subject-specific AA indices were lower ($r_s = .42$ to $.60$, mean $r = .50$) than were correlations among AAs ($r_s = .48$ to $.83$, mean $r = .75$). These results provided support for the differentiability of academic SC and AA, and the suggestion that stronger SC/AA relations can be obtained when academic, rather than more generic SC indices are utilized.

Between-network SC/AA research with the PCSC has also provided support for the construct validity of multidimensional SC. Marsh and Gouvernet (1989) found mathematics and reading achievement indices, as measured by standardized tests, to be more highly correlated with academic SC than with social, physical, and general SC facets for a sample

of Grades 7 to 9 Australian children. Only one of the correlations between the AA indices and the three nonacademic SC facets was significant, and this correlation was less than half that of any of the correlations between the academic SC scale and the AA indices. Marsh (1990b) similarly found grades in mathematics and English, both separately and in combination, to be more highly correlated with the academic scales of the PCSC, SDQ-I, and PHSC than with their nonacademic scales. These results provided support for the multidimensionality of SC, and in particular, the separation of academic and nonacademic SC facets, and of SC and AA constructs.

SDQ research has provided support for the multidimensionality and content-specificity of mathematics and reading SC and AA constructs. For example, Marsh (1988, 1990a, 1992b) reviewed eleven SDQ studies based on seven samples of Australian preadolescents, and one sample of Grades 7 to 9 Australian children. These studies used standardized tests, and/or teacher ratings of mathematics and/or reading achievement (e.g., Marsh & Gouvernet, 1989; Marsh, Parker, & Smith, 1983; Marsh, Relich, & Smith, 1983; Marsh, Smith, & Barnes, 1984; Marsh, Smith, Barnes, & Butler, 1983). Of 13 correlations between mathematics SC and achievement, all were significant and positive (mean $r = .35$), as were all 17 correlations between reading SC and achievement (mean $r = .38$). Consistent with both the Marsh/Shavelson and I/E models, most of the correlations between reading SC and mathematics achievement, and between mathematics SC and reading achievement were positive, non-significant, and even lower than the correlations described above (mean $r = .12$ for both). In addition, of 136 SC/AA correlations involving the nonacademic SC facets and either the reading or mathematics achievement indices, only two were significantly positive.

One of the SDQ-I studies also examined relations between science achievement, as indicated by teacher ratings, and the seven SC facets measured by the SDQ-I (Marsh, Parker, & Smith, 1983). Separate analyses were conducted on three samples of Australian preadolescents in Grades 5 and/or 6. Science achievement was significantly correlated with general-academic SC in all three samples, and was not significantly, or was negatively correlated with the four nonacademic SCs. In addition, reading and mathematics SCs were

each significantly correlated with science achievement in only two of the three samples. Science achievement was, across the three samples, more highly correlated with mathematics SC (mean $r = .39$) than with reading SC (mean $r = .21$). Thus, the results suggested a stronger association between science and mathematics than between science and English. Unexpectedly, however, mathematics SC was slightly more strongly correlated with science than with mathematics achievement in one sample. Perhaps the mathematics/science association, and a context in which items tapping science SC were not included played a role in the absence of support for the content-specificity of mathematics SC across all three samples. That is, had the children been requested to respond to questions related to their SCs in science, as well as in other subject areas, the content-specificity of science and mathematics SC/AA relations might have been more differentiated.

SDQ-I studies using Canadian (Byrne & Worth Gavin, 1996), Filipino (Watkins & Gutierrez, 1989), Nepalese (Watkins et al., 1991), and Singaporean (Tanzer & Sim, 1991) samples also examined relations between mathematics and reading SC and AA indices. As previously noted, with the exception of the Canadian study, these studies included measures of nonacademic SC; in addition, the Filipino and Nepalese studies measured science AA. These studies provided support for the multidimensionality and differentiability of nonacademic and academic SC facets, as well as that of SC and AA constructs. The content-specificity of SC/AA relations was supported in all studies except for the Nepalese and Filipino samples; support was found for only mathematics and English constructs for the Filipino sample, and for only mathematics for the Nepalese sample. Again, perhaps the results of the Filipino and Nepalese samples were related to (1) interpretation and comprehension issues associated with the differing language of the SDQ-I and the primary language of the students, as well as (2) a context in which science SC was not considered.

In summary, the modest correlations obtained between academic SCs and reading and mathematics achievements supported the differentiability of academic SC and AA indices both from each other and amongst themselves. Furthermore, the lack of association between AA indices and nonacademic SCs provided support for the separation of academic

and nonacademic SCs. Indeed, the results generally demonstrated the importance of distinguishing between not only nonacademic and academic SCs but also among subject-specific academic SCs; such distinctions appeared to be necessary for a more complete understanding of the SC/AA nomological network. Thus, research with the PCSC and SDQ-I instruments has provided support for the multidimensionality of SC, and in particular, the separation of academic and nonacademic SC facets, and research with the SDQ and ASDQ instruments has supported the multidimensionality of academic SC.

Support for the content-specificity of subject-specific SC/AA relations has been less consistent. This inconsistency, however, was based upon research in which an expanded set of subject-specific AA indices were examined within a context in which not all of the corresponding subject-specific SC facets were considered. SDQ-I studies in which only mathematics and reading SCs were considered have found achievement indices in subject-specific areas (e.g., mathematics) to be more highly correlated with SCs in matching content areas (i.e., mathematics) than with the academic SC facet in the non-corresponding academic area (i.e., reading), and not significantly or negatively correlated with SCs in nonacademic areas (e.g., physical ability and peers). These SDQ-I studies have provided strong support for the Marsh/Shavelson model of SC, as well as, the multidimensionality and content-specificity of mathematics and reading SC/AA relations. Accordingly, research was needed in which the content-specificity of SC/AA relations were examined within a context wherein an expanded set of subject-specific SC facets were considered.

To the author's knowledge, only one study has examined the between-network pattern of SC/AA relations within a context in which an expanded set of subject-specific indices of both SC and AA constructs were considered. As previously noted, this study was an extension of an earlier within-network study (Marsh, 1990d), and was based on only the sample of Australian adolescent boys (Marsh, 1992a). CFA of eight subject-specific core academic SC and AA indices (i.e., English, Mathematics, Science, History, Commerce, Geography, Computer Studies, and Foreign Languages) substantiated the differentiability and content-specificity of these multidimensional constructs. That is, the correlations among

the eight corresponding SC/AA facets ($r_s = .45$ to $.70$; mean $r = .57$) were, in all instances, higher than those among non-corresponding SC/AA facets ($r_s = .17$ to $.54$; mean $r = .33$). Accordingly, the differentiability, multidimensionality, and content-specificity of a wide range of core academic SC and achievement constructs was supported.

Results of the Marsh (1992a) study also revealed that the specific SC facets were more differentiated than were the specific AA constructs. That is, relations among the eight AA constructs were of a higher magnitude ($r_s = .42$ to $.72$, mean $r = .58$) than were relations among the SC facets ($r_s = .21$ to $.53$; mean $r = .34$). As Marsh noted, these results are consistent with the comparison processes postulated within the I/E model which suggest that SCs are affected by different processes than are AA constructs, and as a result, lead to greater differentiation among the former than among the latter. Finally, non-core academic SCs were relatively distinct from, and in fact, sometimes negatively correlated with specific core AA indices ($r_s = -.17$ to $.29$). Thus, the differentiation of subject-specific core SC/AA indices was demonstrated, as was a greater level of differentiation among academic SC facets than among AA constructs, and between non-core academic SCs and core AA indices.

Hierarchical Self-concept/Academic Achievement Relations

In reviews of SC research, Wylie (1979), Byrne (1984), and West et al. (1980) concluded that AA indices were more strongly associated with academic SC than with general SC. Likewise, Hansford and Hattie (1982), in a meta-analysis examining the relations between SC and achievement/performance measures, reported that, on average, AA indices correlated $.18$ with general SC measures, but $.42$ with academic SC. Research with the PCSC similarly found that AA was more highly correlated with academic SC than with general SC (Marsh, 1990b; Marsh & Gouvenet, 1989). In particular, Marsh and Gouvenet (1989) found mathematics and reading achievement indices to be more highly related to academic SC ($r_s = .40$ and $.35$, respectively) than to general SC ($r_s = .19$ and $.03$, respectively). Marsh (1990b) similarly found AA to be more highly correlated with the academic SC scales of the PCSC and SDQ-I ($r_s = .37$ and $.36$, respectively) than with their general SC scales ($r_s = .11$ and $.13$, respectively). Thus, empirical research has supported a

pattern of SC/AA relations whereby the magnitude of the correlations between AA and academic SC are higher than those between AA and general SC.

The pattern of correlations obtained between reading or mathematics achievement indices and increasingly general SCs for the relevant SDQ-I studies reviewed earlier were also consistent in demonstrating higher correlations between academic SC and subject-specific AA indices than between general SC and subject-specific AA. In addition, a pattern was found in which subject-specific academic SC and AA indices were more highly correlated than were general SC and subject-specific AA indices. That is, mathematics and reading achievements were more highly correlated with general-academic SC than with general SC (e.g., Marsh, Smith, & Barnes, 1984; Marsh & Gouvenet, 1989). Similarly, mathematics and reading achievement indices, as measured by either standardized tests or teacher ratings, were correlated more highly with corresponding academic SCs than with general SC. The same pattern of results was obtained with the Canadian (Byrne & Worth Gavin, 1996), Filipino (Watkins & Gutierrez, 1989), Nepalese (Watkins et al., 1991), and Singaporean (Tanzer & Sim, 1991) samples. Taken together, these results provided support for the multidimensionality of specific academic SC facets, as well as the hierarchical organization of academic and general SCs.

However, comparison of the magnitude of the correlations obtained between subject-specific academic SC and AA indices against those between academic SC and subject-specific AA indices yielded an inconsistent pattern of results. Consistent with an hierarchical structure, some studies found the correlations between the former and more specific indices to be higher than those between the latter and more general measures (e.g., Marsh, Barnes, et al., 1984; Marsh & Gouvenet, 1989; Tanzer & Sim, 1991; Watkins & Gutierrez, 1989). However, in other studies, the former correlations for the mathematics and/or reading subjects areas were higher, equal to, or lower than the latter correlations (e.g., Byrne & Worth Gavin, 1996; Marsh, Parker, & Smith, 1983; Marsh, Relich, & Smith, 1983). Research with older samples of Canadian (e.g., Byrne & Shavelson, 1986), American (e.g., Shavelson & Bolus, 1982), Australian (e.g., Marsh & O'Neill, 1984), and Korean (e.g., Song

& Hattie, 1985) students have similarly found inconsistent support for this aspect of the hierarchical pattern of SC/AA relations. Thus, SC research with preadolescent and older samples have not demonstrated consistent support for a multidimensional and hierarchical pattern of SC/AA relations in which subject-specific SC and AA indices are more highly correlated than are subject-specific AA indices and academic SC.

In summary, empirical research relevant to an hierarchical pattern of SC/AA relations has supported a pattern of relations in which AA criteria correlated more strongly with specific and general academic SCs than with general SC. In contrast, comparisons of the association between external criteria and subject-specific academic SCs against those between external criteria and general academic SCs revealed inconsistent results. That is, the prediction that the former correlations would be higher than the latter has not always been supported. Although this inconsistency in the empirical research is puzzling, and remains unaddressed in the theoretical literature, the comparisons were based on SC facets derived from scale (or factor) scores, or latent constructs derived through first-order CFA. A more appropriate test of the between-network hierarchical structure of SC would involve a comparison of correlations between AA and first-order latent SC constructs against those between AA constructs and second-order SC factors derived on the basis of second-order (i.e., hierarchical) CFA. Accordingly, in the present analyses, empirical support for an hierarchical pattern of SC/AA relations was examined by comparing AA relations with first-order latent SC constructs against those with second-order SC constructs derived on the basis of first-, and second-order CFA, respectively. We turn now to a review of research relevant to gender differences in SC and SC/AA relations.

Construct Validity Research: Gender Differences in Preadolescent Self-concept

Much of the research examining gender differences in SC has focused on substantive issues related to differences in mean SC between girls and boys. As with SC research in general, prior to the 1980s, empirical efforts focused on general or academic SCs. Not surprisingly, reviews of this research emphasized the lack of conclusive results (e.g., Hansford & Hattie, 1982; West et al., 1980; Wylie, 1979). For example, Wylie (1979)

concluded that there was no evidence for a relation between gender and overall SC for any age group. However, she suggested the possibility that these results were due to the lack of specificity with which SC was measured; that is, specific SC facets had not been examined, and potential differences in these more specific facets were perhaps being masked with the use of more general measures of SC. With the development of theoretically-based and empirically-derived multidimensional SC instruments, researchers have begun examining gender differences in specific SC facets. However, much of this research has, as with earlier research, focused on observed mean SC scores between girls and boys.

Implicit in the comparison of observed mean scores across gender is the assumption that the measurement (i.e., factor loadings) and factor structure (i.e., factor covariances) of SC is equivalent across girls and boys. Unless this assumption is valid, comparisons of observed means across gender are unjustified (see, for example, Byrne & Shavelson, 1987; Marsh, 1987b, 1993b; Marsh, Smith, & Barnes, 1985). Indeed, Byrne and Shavelson (1987), in their review of research examining gender differences in SC, argued that much of the inconsistency in reported findings derived from the violation of the assumption of equivalent SC measurement and structure. Tests of this assumption bear on the construct validity of SC, and are an important methodological issue that must be addressed prior to substantive issues related to gender differences in specific SC means. Accordingly, the focus of the present review is on issues relevant to the gender invariance of the measurement and structure of preadolescent SC, and in particular, the Marsh/Shavelson model of SC.

Despite the importance of an underlying invariant SC measurement and structure in multigroup comparisons of observed means, few researchers have examined this aspect of the construct validity of SC across gender. Separate exploratory factor analyses of samples of girls and boys can provide information relevant to differences in the factor loading pattern, the magnitude of the factor loadings, and the relations among SC facets across gender. Such comparisons, however, do not permit researchers to test, nor do they address the invariance of the factor loadings and covariances across groups. Moreover, exploratory factor analytic procedures assume that an instrument measures a construct with complete

reliability; that is, error variance is not measured. In contrast, CFA accounts for error variance, and enables researchers to test for factorial invariance by providing for a simultaneous analysis of the data across the comparative groups. CFA is, therefore, a methodologically more rigorous and statistically sophisticated test of multigroup equivalence, and is the most appropriate procedure available for testing measurement and structural invariance across groups. Accordingly, only CFA studies relevant to the gender invariance of SC measurement and structure will be reviewed.

Investigations of factorial invariance using CFA involve two key issues, one related to the measurement of the construct under study (measurement invariance), the other related to its structure (structural invariance; for a more detailed discussion, see, e.g., Byrne, 1989, 1994, 1998; Byrne, Shavelson, & Muthén, 1989). Questions of measurement invariance address whether a particular measuring instrument(s) is operating in the same way across groups. From a statistical perspective, measurement invariance focuses on the equivalence of factor loadings across groups; the central issue being the equality of item or subscale measurements across groups. Questions of structural invariance centre around the equality of relations among the latent constructs (i.e., factor covariances). We turn now to a review of research relevant to the equivalence of multidimensional and hierarchical preadolescent SC measurement and structure across gender; this review is presented under the rubrics of within- and between-network SC research.

Within-network Research: The Invariance of a Multidimensional/Hierarchical Self-concept

In order to examine the assumption of an invariant SC factor structure across gender, Marsh and colleagues (Marsh, Smith, & Barnes, 1985) tested the invariance of a series of increasingly restrictive models of multidimensional SC across fifth-grade preadolescent Australian girls and boys. Results supported the gender invariance of the measurement and structure of the eight SC facets measured by the SDQ-I. Marsh (1987b) further substantiated the factorial invariance of SC measurement and structure across preadolescent girls and boys. Using a random sample of Grade 5 Australian children from the SDQ-I normative data base, Marsh tested the factorial invariance of the seven more specific SC facets measured by

the SDQ-I across gender, and across samples using randomly divided groups of the same gender; measures of general SC were not included in analyses. In both the cross-validation and the cross-gender comparisons, the equivalence of all academic and nonacademic SCs was supported. These results provided strong support for the gender invariance of the measurement and structure of SC, as measured by the SDQ-I, for preadolescents.

To the author's knowledge, these studies represent the only CFA research to examine SC gender invariance with samples relevant to the present study. The gender invariance of the measurement and structure of SC was tested for a preadolescent Arabic sample with a translation of the SDQ-I (Abu-Hilal & Aal-Hussain, 1997). Results supported the invariance of the measurement of SC across gender, however, consistent with a priori expectations related to differences between Arabic and Western cultures, the correlations among the first-order SC factors were found to be of a lower magnitude for girls than for boys. Thus, Arabic girls had more differentiated SCs than did Arabic boys. However, as previously noted, the cultural differences between Arabic and Western cultures render the generalizability of the results of this study less applicable than is true of preadolescent Australian research.

CFA research examining gender differences in the SC construct have also been conducted with adolescent Australian (Marsh, 1993b), Singaporean (Tanzer et al., 1992), and Canadian (Byrne, 1990a; Byrne & Shavelson, 1987) samples; analyses of the Australian and Canadian samples were restricted to verbal, mathematics, academic, and general SCs. With few exceptions, results supported the gender invariance of the measurement and structure of first-order SC constructs. CFA results for the Canadian sample, which measured SC facets with three SC instruments, provided support for the gender invariance of SC measurement with the exception that the SDQ-III Mathematics SC facet cross-loaded on the first-order English SC construct for boys. Comparison of the correlations among the latent SC constructs also revealed gender differences in SC structure: for girls, English SC correlated more highly with Academic and General SCs than did Mathematics SC, whereas the reverse was found for boys. These results suggested that English SC had a stronger influence on the formation of more general SCs for girls, whereas Mathematics SC had a

stronger influence for boys. Thus, the Byrne and Shavelson (1987) results suggested that the academic portion of the SC hierarchy differed across gender for adolescents. Although Byrne and Shavelson's results may be due to a developmental phenomenon, they nevertheless highlight the necessity of testing for higher-order SC structural invariance across gender.

Taken together, the results of these research studies suggest that the measurement and structure of SC is likely invariant across gender for preadolescents in Western cultures. That is, the finding of gender invariant multidimensional SC measurement and structure was replicated across preadolescent Australian samples (i.e., Marsh, 1987b; Marsh, Smith, & Barnes, 1985), but was less consistent for adolescent samples. In addition, consistent with a priori predictions specifying gender-related cultural differences, support for the gender invariance of SC among preadolescents from a non-Western culture was supported for the measurement but not the structure of SC (i.e., Abu-Hilal & Aal-Hussain, 1997). The lack of consistency across studies for adolescents, as compared to preadolescents, is congruent with conclusions made in reviews (e.g., Meece, Parsons, Kaczala, Goff, & Futterman, 1982) and meta-analyses (e.g., Hyde, Fennema, Ryan, Frost, & Hopp, 1990) of research examining gender differences in SC constructs. That is, gender differences in SC constructs are larger and more reliably identified among adolescents than among preadolescents, a finding which has been interpreted to suggest that gender differences are developmentally linked, and are more apparent among high school students than among primary school students. Accordingly, differences in sample characteristics related to developmental and cultural factors render the results of research based on adolescent and Arabic preadolescent samples less applicable to the present study than are results based on Australian preadolescents. Thus, the results of the empirical research reviewed above, and in particular, the replication of gender invariant multidimensional SC measurement and structure across Australian preadolescents, suggest that the measurement and structure of SC is invariant across gender for the preadolescent sample relevant to the present study.

Although Australian studies examining the gender invariance of preadolescent SC measurement and structure have provided a valuable contribution to SC research, these

empirical efforts are characterized by several limitations. First, none of the studies examined gender differences in the measurement and structure of SC using an expanded set of academic SC facets. In addition, differences in the second-order structure of SC were not tested within the framework of hierarchical CFA; tests were restricted to an examination of the first-order SC facets. Although tests of the invariance of relations among first-order SC factors can be used to infer second-order structural invariance, hierarchical CFA procedures allow for a direct test of the invariance of hierarchical structure. Accordingly, a primary purpose of the present thesis was to examine the invariance of the measurement and structure of preadolescent multidimensional and second-order hierarchical SC structure within and across gender for multiple nonacademic SCs, and an expanded set of subject-specific core and non-core academic SC facets.

Between-network Research: Self-concept/Academic Achievement Relations

As was typical of early SC research, research examining gender differences in SC/AA relations has historically focused on global SC and achievement constructs, and has yielded contradictory results (for a review, see West et al., 1980). With the recent conceptual and methodological advances in SC, researchers have begun to examine gender differences in specific SC and achievement constructs. However, as with research examining gender differences in SC, much of the focus has been of a substantive nature. Moreover, of the studies examining gender differences in SC/AA relations, few have been conducted using CFA procedures, and none has tested for the invariance of these relations.

Marsh, Smith and Barnes (1985) examined gender differences in SC/AA relations within the context of a study focusing on the validity of the I/E model. In particular, they examined the influence of gender on relations predicted by the I/E model between mathematics and reading SCs and achievement indices for a preadolescent Australian sample. Gender comparisons indicated that girls had higher reading SCs that were accounted for by their higher reading AA, whereas boys had higher mathematics SCs which could not be accounted for by gender differences in mathematics achievement. Nevertheless, results indicated that the inclusion of gender had little practical effect on the value of the path

coefficients linking mathematics and reading constructs. Thus, these results suggested that SC/AA relations for mathematics and reading constructs did not significantly vary across gender for preadolescents. Consistent with a developmental difference between preadolescents and adolescents, the influence of gender on the tests of the I/E model was more profound with adolescents (see, e.g., Marsh et al., 1988).

Gender differences relevant to a multidimensional and hierarchical pattern of SC/AA relations were examined within a CFA framework by Byrne and Shavelson (1987) for a Canadian adolescent sample; the relevant within-network results for this study were reviewed above. Although this study was conducted with a different age group than is the focus of the present study, given its unique status, the results will be reviewed. Consistent with a multidimensional SC, Mathematics and English grades were more highly correlated with corresponding than non-corresponding SCs for girls. For boys, however, this pattern was true of only Mathematics grades; grades in English were equally correlated with Mathematics and English SC facets. Thus, results supported the content-specificity of multidimensional SC/AA relations of verbal and mathematics constructs for adolescent girls, but of only mathematics constructs for adolescent boys.

The hierarchical structure of SC/AA relations for mathematics achievement was also supported for both girls and boys (Byrne and Shavelson, 1987). That is, grades in mathematics were most correlated with Mathematics SC, less correlated with Academic SC, and least or uncorrelated with General SC. Support for the hierarchical structure of SC for verbal constructs was less consistent. Unexpectedly, grades in English were more highly correlated with Academic SC than with English SC, although they were least correlated with General SC; examined from the reverse perspective, however, the hierarchical structure of SC was supported. Thus, although the hierarchical structure of SC was strongly supported for mathematics constructs for both girls and boys, it was not supported for verbal constructs. Nevertheless, consistent with the pattern of the results reviewed earlier with regard to the within-network hierarchical structure of SC, lack of support for an hierarchical pattern of relations occurred with regard to the first-, as compared to the second-order level

constructs; support for an hierarchical pattern from both the first- and the second-order level of the hierarchy in relation to the third-level (i.e., general SC) was consistent.

Current research has not adequately addressed gender differences in the multidimensional, hierarchical pattern of relations between SC and AA constructs with preadolescents. Indeed, to the author's knowledge, CFA research testing for gender differences in the multidimensional and hierarchical pattern of relations between SC and AA constructs has not been conducted with preadolescents. Accordingly, there is a need for research examining this aspect of the construct validity of the SC nomological network.

Purpose of the Present Study

As the review of the literature revealed, SC instrument and construct validation research over the course of the past two decades has resulted in an increased understanding of the SC nomological network. However, these research efforts are limited for several reasons. The purpose of the present study was to address several limitations in current SC research. In particular, research examining the within- and between-network components of the SC nomological network has almost exclusively limited the study of the academic portion of SC to mathematics and reading constructs. The only research study to examine an expanded set of academic SCs was conducted with Australian boys, used only the ASDQ, did not include nonacademic SC facets, and for preadolescents, did not examine between-network SC/AA relations (Marsh, 1990d). Thus, there is a need for research examining the construct validity of the Marsh/Shavelson model of SC for preadolescents within a CFA framework in which multiple independent measures of nonacademic SC facets, and of an expanded set of subject-specific academic SC and achievement indices are considered.

In addition, much of the research examining gender differences in SC has focused on the substantive issue of differences in levels of specific SC facets. The important methodological and construct validation aspect of whether the measurement and structure of SC is equivalent across gender has rarely been examined. Moreover, this issue has never been addressed in relation to a broad range of academic SC facets. The between-network gender invariance of relations among a broad range of subject-specific SC and AA

constructs has similarly never been examined. Accordingly, research examining the measurement and structure of SC, and of SC/AA relations within and across gender using CFA procedures was needed.

The present study addressed these limitations in current SC research by examining the construct validity of an academic expansion of the Marsh/Shavelson model separately for preadolescent girls and boys, and then testing for the equivalence of SC models across gender. All analyses were conducted using CFA procedures within the framework of covariance structure modeling. Multiple nonacademic SC facets were measured using the SDQ-I (Marsh, 1992b) and the SPPC (Harter, 1985). An expanded set of academic SCs was measured using adaptations of the ASDQ-I (Marsh, 1990d) and the SPPC. Subject-specific AA constructs were obtained from children's grades and self-reported grades.

Tests of the construct validity of the expanded Marsh/Shavelson model were conducted for Grades 5 to 6 students within and across gender in relation to both the first- and second-order level of the SC construct. In addition, the construct validity of the between-network component of the SC nomological network was examined within and across gender in relation to subject-specific AA constructs at both the first- and second-order level of the SC construct. More specifically, the purposes of the study were (1) to test for the multidimensional structure of nonacademic SCs and an expanded set of academic SCs separately for girls and boys, (2) to test for a second-order hierarchical structure of this expanded multidimensional SC model separately for girls and boys, (3) to test for invariant measurement and structure across gender relative to a multidimensional and hierarchical SC, (4) to test for the multidimensional and hierarchical pattern of relations between SC and AA constructs separately for girls and boys, and (5) to test for the gender invariance of SC/AA relations across gender. The hypotheses and models tested will be reviewed in more detail in the method section.

Although not a focus of the present study, measures of general SC and of the importance of doing well in school were obtained. The appropriate measurement of general SC is the subject of theoretical controversy. At least eight operational definitions of general

SC have been proposed in the literature (for reviews, see, e.g., Marsh, 1986a, 1990a; Marsh & Hattie, 1996). For example, researchers have proposed that this construct is best represented as an hierarchical construct derived through analyses of lower-order SCs (e.g., Marsh, 1987c). Others have proposed that relations between more specific SC facets and general SC are mitigated by the importance which individuals attach to the more specific facets (e.g., Harter, 1985; Hoge & McCarthy, 1984). Given the significant scope of the present thesis, an examination of the validity of each of these theoretical perspectives was precluded. However, inclusion of measures tapping general SC and importance constructs allowed for future analyses examining this aspect of SC validation. Accordingly, although these data were collected, they did not form part of the present analyses.

Importance of the Present Study

Researchers have generally assumed that the measurement and structure of SC is invariant across gender. Tests of this assumption are an important methodological aspect of the construct validation process; unless such an assumption is warranted, research examining gender differences in SC may be seriously flawed. As previously noted, Byrne and Shavelson (1987) attributed much of the inconsistency in research examining gender differences in SC to the violation of this assumption. Furthermore, between-network investigations of gender differences in the SC nomological network can proceed more meaningfully when methodological issues related to measurement and structural issues have been addressed. With the expansion of the academic portion of the Marsh/Shavelson model of SC (Marsh, 1990d), a comprehensive investigation of the within- and between-network multidimensional and hierarchical aspects of the construct validity of this model within and across gender is of significant theoretical and methodological importance.

METHOD

Sample

Participants were 474 Grades 5 and 6 students drawn from elementary schools under the jurisdiction of three school boards in Ottawa and surrounding areas. The schools were located primarily in middle-class neighborhoods in suburban communities. Two hundred and thirty-eight of the participants were girls and 236 were boys. For both the sample of girls and the sample of boys, the mean age of the participants was 10.8 years and the age range was 9 to 13 years. The mean age of the Grades 5 and 6 samples of girls and boys is presented in Table 1. As Table 1 reveals, the number and mean age of the participants were remarkably evenly distributed across gender and, when accounting for the year difference implied, grade level.

Table 1

Sample Mean Age Within and Across Grade and Gender

	Girls		Boys		Total Sample	
	<u>n</u>	<u>M</u>	<u>n</u>	<u>M</u>	<u>N</u>	<u>M</u>
Grade 5	117	10.26	121	10.32	238	10.29
Grade 6	121	11.22	115	11.32	236	11.27
Grades 5 & 6	238	10.75	236	10.81	474	10.78

Although two school boards were initially targeted for inclusion in the present study, one of the school boards declined to participate due to large-scale standardized testing. Accordingly, two smaller school boards in the surrounding Ottawa area were invited, and consented to participate. The resulting potential student pool consisted of approximately 540 participants; the larger school board imposed a 300 subject limit, and there were

approximately 120 potential Grades 5 and 6 students from each of the two smaller school boards. Thus, the participation rate from the three consenting school boards was quite high, being approximately 87.8%.

Procedure

Subsequent to obtaining consent from the school boards, schools, and teachers, participants were recruited on a class-by-class basis. That is, the researcher made an oral presentation to the students in which they were provided with information about the study and asked to indicate, on a class list containing all of their names, whether they wished to participate (see Appendix A for the standardized student oral presentation). The students who wished to participate were given a hand-out which they were asked to provide to their parent(s). The hand-out contained a permission letter and two identical parental consent forms, one to be returned to the teacher and one to be kept by the parent(s); the former described the study and asked parents whether they wished their child to participate (see Appendix A for the parental permission letter and form). Only students who provided consent and for whom parental consent was obtained participated in the study.

Data collection took place beginning approximately two weeks after the first or second term report card in order to ensure that all students were cognizant of their most recent academic performance. Data were collected by the researcher or a trained research assistant. Non-participating students engaged in an alternate activity during administration of the instruments. All instruments were group administered to intact classes and were read aloud to the children in order to minimize difficulties associated with reading the items, although none were expected. Immediately prior to the beginning of the first administration session, the students were again provided with a brief oral description of the study, and the confidential and voluntary nature of their participation was reiterated (see Appendix A for the standardized student oral presentation given at the beginning of the first administration session).

Due to time constraints imposed by, and at the request of school officials, the SDQ-I and ASDQ-I instruments were integrated and administered as one instrument with one set of

instructions; the SDQ-I and ASDQ-I are hereafter collectively referred to as the (A)SDQ-I or (A)SDQ instruments. The instruments were administered over two separate sessions; in all cases, the sessions were on the same day and were separated by the recess or lunch break. During the first administration session, participants completed either the (A)SDQ-I or the SPPC instrument. During the second session, participants completed a second SC instrument, followed by a self-report measure of grades and two importance rating scales. The order of administration of the SC instruments was alternated across classes. The first administration session required 25-30 minutes and the second required 45-50 minutes.

Instrumentation

A battery of six self-report measuring instruments was used in the present study. These instruments may be categorized into three types, namely, self-report measures of SC, of grades, and of academic importance. The children completed three multidimensional preadolescent SC instruments, two of which, as previously noted, were administered as one instrument. These SC instruments tapped general SC, general-academic SC, as well as specific nonacademic and academic SC components. The children also completed three academic rating scales. One scale summarized their most recent AA in each of the academic areas tapped by the SC instruments and two scales required students to rate the importance of doing well in each of these same academic areas. In addition, the children's grades in the specific academic subject areas were obtained from their most recent school records.

Due to time restrictions imposed by school officials on the data collection process, the range of academic subject areas originally targeted for inclusion in the present study was reduced. That is, in the interest of reducing administration time, school officials requested that items designed to tap four academic content areas in French Language (i.e., Reading, Writing, Speaking, and Listening) be dropped from the study; accordingly, these academic areas were not measured. In addition, three academic content areas (i.e., Music, Drama, and Geometry) were measured for participants in only two of the three school boards, and thus, were dropped from analyses. Accordingly, a total of seven academic content areas originally targeted for inclusion in the present study were dropped from the hypothesized SC models.

The variation across school boards in the specific academic areas measured was related to differences among the participating school boards in the subject areas represented in their report cards. That is, the level of academic SC differentiation had been hypothesized to parallel the level of differentiation in the children's report cards, and thus, the academic areas included in the measures for each of the school boards were selected on this basis. As a result, not all of the academic areas were tapped across all three school boards. However, in order to permit the inclusion of participants from all three school boards, only the specific academic scales common to the three school boards were utilized in the analyses. The specific academic areas tapped in each of the school boards and the common academic areas included for analyses in the present study are summarized in Table 2 (see Appendix B for academic subject area descriptions for each of the school boards). The academic subject areas common across all three school boards were Social Studies, Science, two specific mathematics areas (i.e., Arithmetic and Measurement), four specific Language Arts areas (i.e., Listening, Speaking, Reading, and Writing), Art, Gym, and Religion. Note that Music (considered a subarea of Art) and Geometry (a Mathematics subarea) were measured in only two of the three school boards, and Drama (another Art subarea) was measured in only one school board. A total of 11 specific academic areas were measured across all three school boards, and thus, were included in analyses.

Self-concept Measures

Measures of SC were obtained using three multidimensional instruments. These included: (a) the Self-Description Questionnaire-I (SDQ-I; Marsh, 1992b), (b) the Academic Self-Description Questionnaire (ASDQ-I; Marsh, 1990d), and (c) the Self-Perception Profile for Children (SPPC; Harter, 1985). As previously noted, at the request of school officials, the (A)SDQ-I instruments were integrated and administered as one instrument with one set of instructions; the similarity of their instructions and response format rendered their integration feasible. We turn now to a description of each of these instruments and a review of research bearing on their psychometric properties.

Table 2

Summary of the Specific Academic Areas Measured in Each of the Three School Boards and the Specific Academic Constructs Common Across the School Boards

Common Academic Areas	School Board		
	A	B	C
Arithmetic	Arithmetic	Computation	Computation
Measurement	Measurement	Problem Solving	Measurement
	Geometry		Geometry
Science	Environmental Studies	Science	Natural Science
Social Studies	Social Studies	Social Studies	Social Studies
Listening in LA	Listening in LA	Listening in LA	Listening in LA
Speaking in LA	Oral Expression/ Speaking in LA	Oral Expression/ Speaking in LA	Speaking in LA
Writing in LA	Writing in LA	Writing in LA	Writing in LA
Reading in LA	Reading in LA	Reading in LA	Reading in LA
Gym	Physical Education	Physical Education	Physical Education
Religion	Religion and Family Life	Religion and Family Life	Religion or Moral Education
Art	Visual Arts	Art	Arts
	Music	Music	
	Drama		

Note. LA = Language Arts.

The Self-Description Questionnaire-I

The SDQ-I is a 76-item instrument designed to measure preadolescent general and general-academic SCs, as well as specific academic and nonacademic SCs (Marsh, 1992b; see Appendix C for the standardized instructions, response format, and items). The SDQ-I contains eight scales which measure: (1) general SC (i.e., the General-Self scale), (2) general-academic SC (i.e., the All School scale), (3) four specific nonacademic SCs (i.e., the Relations with Peers, Relations with Parents, Physical Ability, and Physical Appearance scales), and (4) two specific academic SCs (i.e., Reading and Mathematics scales). Only the four nonacademic and the general SC scales were used in the present study. Each of these scales contain eight positively worded self-descriptive statements (e.g., “I am a nice looking person”); an additional 12 items are negatively worded but are not included in scoring the SDQ-I. The rationale for this exclusion was based on research which suggested that preadolescent children do not respond appropriately to negatively worded items (e.g., Benson & Hocevar, 1985; Marsh, 1986b, 1992b; Marsh & MacDonald Holmes, 1990; Marsh, Barnes, et al., 1984). Children are required to respond on a 5-point Likert-type scale ranging from false to true.

At the request of school officials, only six of the eight positively worded self-descriptive statements were utilized in order to reduce administration time. Item selection was based on statistical data presented in the instrument manual available at the time of the study (Marsh, 1988; but see Marsh, 1992b, for the most recent SDQ-I manual). Specifically, the manual provided, based on the SDQ-I normative sample ($N = 3,562$) of Australian children in Grades 2 to 6, corrected item-scale correlations representing the correlation between each SDQ-I item and the remaining items in the matching scale. The six items from each scale having the highest item to scale correlations were selected for the present study (see Appendix C for the list of included and excluded items).

Research has provided strong support for the validity of the SDQ-I as a measure of preadolescent multidimensional SC. The factor structure of this instrument has been replicated across a variety of preadolescent Australian samples in Grades 2 to 6 using both

exploratory and confirmatory factor analytic procedures (e.g., Marsh, Barnes, et al., 1984; Marsh & Hocevar, 1985; for reviews, see Marsh, 1990a, 1992b). CFA and tests of factorial invariance have supported both the SDQ-I scales and their factorial invariance across preadolescent Australian and British samples (Marsh & Smith, 1987). Exploratory factor analyses with preadolescent Nigerian (Watkins & Akande, 1992), Nepalese (Watkins et al., 1991), Filipino (Watkins & Gutierrez, 1989), and Zimbabwean (Watkins & Mpofu, 1994) samples reported to be fluent in English have provided support for the specific SDQ-I scales, although support for the more general scales was weaker. In addition, exploratory and confirmatory factor analyses supported the SDQ-I scales for Singaporean preadolescents (Tanzer et al., 1992; Tanzer & Sim, 1991), and the general and academic scales of the SDQ-I were supported for a preadolescent Canadian sample (Byrne & Worth Gavin, 1996).

Support for the concurrent, convergent and/or discriminant validity of corresponding SDQ-I scales has been demonstrated in MTMM studies based on the Campbell-Fiske criteria and CFA using the inferred ratings of peers, teachers, and/or parents (e.g., Marsh, Barnes & Hocevar, 1985; Marsh & Craven, 1991; Marsh, Smith & Barnes, 1983, 1984; Marsh, Smith, et al., 1983), as well as other preadolescent multidimensional SC instruments, most notably the PCSC (Byrne & Schneider, 1988; Marsh, 1990b; Marsh & Gouvenet, 1989). Correlational analyses of corresponding scales of the SDQ-I and the MSCS have also provided at least moderate support for the convergent and discriminant validity of the social and physical scales of this instrument, but only weak support for the SDQ-I Mathematics scale (Delugach, Bracken, Bracken, & Schicke, 1992).

Internal consistency reliability coefficients, based on the SDQ-I normative sample ($N = 3,562$), were available for all but the General-Self Scale, the latter of which were based on a subsample ($n = 739$; Marsh, 1992b). Reliability coefficients for all scales were reported to be in the 80s and 90s. In addition, test-retest reliability coefficients for a sample of Grades 5 and 6 students over a six-month period ranged from .54 to .74 (mean $r = .62$; Marsh, 1992b). Thus, a plethora of construct validity research has, with few exceptions, supported the psychometric soundness of the SDQ-I as a measure of multidimensional preadolescent SC.

The Academic Self-Description Questionnaire-I

The ASDQ-I is a 78-item measure of multiple dimensions of preadolescent academic SC facets (Marsh, 1990d). As with the SDQ-I, this instrument was designed on the basis of the Shavelson model. As originally utilized, the ASDQ-I includes 13 scales, 12 of which measure specific academic SCs, and one of which measures general-academic SC (i.e., the General-School scale). Each scale contains five positively worded items and one negatively worded item; the items in each scale are identical except for the specific academic area tapped. Given the potential response bias associated with negatively worded items with preadolescent children, only the five positively worded items were included in the present analyses. The ASDQ-I was used to measure general-academic SC, and specific academic SCs consistent with the grading classification systems adopted by the educational boards, as outlined in Table 2. The ASDQ-I response format requires children to respond on a 6-point Likert-type response scale ranging from true to false (Marsh, 1990d). However, in order to permit joint administration of the (A)SDQ-I instruments, the 5-point SDQ-I response format was utilized given the more extensive validation work that has been conducted with this instrument (see Appendix C for the ASDQ-I items).

Although instrument validation research specific to the ASDQ-I is not extensive, the design of the ASDQ-I closely approximates that of the SDQ-I. In fact, the ASDQ-I items parallel, or were derived from those of the SDQ-I. Accordingly, the plethora of research supporting the validity of the SDQ-I has bearing on the psychometric properties of the ASDQ-I. Consistent with this, exploratory and confirmatory factor analyses of the responses of a sample of Grades 5 and 6 Australian boys have supported the factor structure of the 13 ASDQ-I scales; their internal consistency has similarly been supported (i.e., reliability coefficients were reported to range from .88 to .94, *Mdn* = .91; Marsh, 1990d).

The Self-Perception Profile for Children

The SPPC is a 36-item multidimensional instrument designed for use with children in Grades 3 to 6. As previously noted, the SPPC represents a revision and expansion of a previous version of this instrument, the PCSC (Harter, 1982). The SPPC was designed to

measure preadolescent self-perceptions of competence in five specific areas, in addition to their general sense of self-worth (Harter, 1985; see Appendix C for the SPPC standardized instructions, response format, and items). The specific areas tapped by this instrument include self-perceptions of Scholastic Competence, Social Acceptance, Athletic Competence, Physical Appearance, and Behavioral Conduct. Each scale contains six items and each item contains one pair of opposing statements, one indicating high competence, the other indicating low competence (e.g., “Some kids do very well at all kinds of sports BUT other kids don’t feel that they are very good when it comes to sports”). Items are reversed such that half of the items within each scale have the high competence statement on the left, the other half of which have it on the right. Children are required to decide which of the two statements within each question is more typical of themselves, and then rate the statement as “really” or “sort of” true of themselves; thus, a 4-point structured-alternative format is used. The rationale behind the choice in response format was a hypothesized reduction in the influence of social desirability response sets, although empirical evidence of the effectiveness of this approach was not provided.

All of the SPPC scales except the Behavioral Conduct scale were used in the present study. The Social Acceptance, Athletic Competence, and Physical Appearance scales of the SPPC were used to represent Peer, Physical Ability, and Physical Appearance SCs, respectively. The Parent SC facet was not measured with the SPPC; thus, this SC facet was measured with only one instrument. As with the ASDQ-I, the Scholastic Competence scale items were adapted to measure general-academic SC, as well as each of the specific academic SC facets outlined in Table 2.

At the request of school officials, only a subset of the SPPC Scholastic scale items were utilized in order to reduce administration time. That is, only four of the six SPPC Scholastic scale items were administered to the students. Item selection was based on statistical data available in the instrument manual (Harter, 1985). Specifically, the SPPC manual provided individual item data based on the age group relevant to the present sample, namely, Grades 5 and 6 students. On the basis of the factor loadings of each item on its

respective scale, the two items having the lowest target factor loadings were selected for exclusion (see Appendix C for the list of included and excluded items).

Research relevant to the validity and reliability of the SPPC was reported in the instrument manual (Harter, 1985). Norms were based on four independent samples of Grades 3 to 8 children drawn from schools in Colorado. Internal consistency reliabilities were reported to range between .71 and .86 for four samples of children in Grades 3 to 8. Factor analyses of the SPPC has supported the separation of its scales for Grades 5 to 8 samples. In addition, the development of the SPPC was based on numerous analyses of the initial version of this instrument (i.e., the PCSC) with geographically diverse samples of elementary school children (Harter, 1982). Test-retest correlations covering intervals of one month to one year were reported to range from .40 to .65 (Harter, 1990).

Independent research efforts using preadolescent Australian (Marsh & Gouvernet, 1989; Marsh & MacDonald Holmes, 1990) and Canadian (Byrne & Schneider, 1988) samples have provided evidence for the validity of the academic, physical, social, and general scales of the PCSC. In particular, exploratory and confirmatory factor analyses have provided support for the factorial validity of the PCSC scales. Correlational analyses have provided support for concurrent, but not convergent validity, although the latter was interpreted to reflect the choice of external criterion (Byrne & Schneider, 1988). In addition, Byrne and Shavelson (1996) conducted CFAs, separately for Grades 3, 7, and 11 Canadian children, of the general SC scale and a multidimensional expansion of the social SC scale of both the SPPC and the SDQ instruments. The social scales were expanded to measure specific (i.e., classmates, teachers, siblings, and parents for the Grade 3 sample) and more general (i.e., general-social, social-school, and social-family) social SC facets. Results provided support for the multidimensional adaptation of these instruments for measuring a broader range of social SC dimensions in each of the three grades.

MTMM analyses of the PCSC and the corresponding SDQ-I scales have also provided support for convergent validity, and modest to strong support for discriminant validity (Marsh & Gouvernet, 1989; Marsh & MacDonald Holmes, 1990). However, as

previously noted, Marsh (1990b; Marsh & MacDonald Holmes, 1990) found evidence of a moderate method effect with the PCSC, as compared to little indication of a method effect with the SDQ-I; these differing results were attributed to the complexity of the PCSC response format. Other researchers have expressed similar concern with the response format of the PCSC and SPPC instruments (e.g., Keith & Bracken, 1996), and Marsh (1990b; Marsh & MacDonald Holmes, 1990) noted that a number of children experienced difficulty responding appropriately to this response format. However, analyses of the SDQ-I were based on item-pairs whereas analyses of the PCSC were based on single items. Although the method effect associated with the PCSC was likely related to its response format, the differing basis for the analyses likely also contributed to the magnitude of the method effect associated with this instrument. As outlined in greater detail in the data analyses section, single-item analyses contributes to a higher likelihood of the occurrence of idiosyncratic and trivial sample-specific artifacts. In contrast, the use of item-pairs increases the reliability and validity of indicator variables across samples (for a discussion of this issue, see, for example, Marsh, 1992b, 1993b; Marsh & O'Neill, 1984).

Although psychometric support for the reliability and validity of Harter's instruments was, and remains weaker and less extensive than that for the SDQ-I, moderate to strong support for this instrument as a measure of multidimensional preadolescent SC has been provided across a diversity of samples. As noted by Byrne (1996b), these instruments are among the most widely used measures of SC, and additionally, on the basis of empirical research, have been recommended for use with preadolescent samples (e.g., Marsh & Gouvenet, 1989; Marsh & MacDonald Holmes, 1990). Harter's instruments, and particularly the SDQ instruments, represented the most thoroughly researched and empirically validated instruments amenable to the investigation of the Marsh/Shavelson model (for a review of SC instruments across a wide range of age groups, see Byrne, 1996b).

Academic Achievement Measures

SC researchers have relied upon a variety of types of AA indicators. These indicators have included: (a) teacher ratings of academic ability, (b) standardized tests, (c) children's

school grades, and/or (d) self-reports of school grades (e.g., Byrne & Worth Gavin, 1996; Marsh, Smith, & Barnes, 1985; Shavelson & Bolus, 1982). Given the requirement of statistical identification in the estimation of CFA models, two indicators of AA were required for each latent AA construct. A desire on the part of school officials to not impose any time requirements on the part of the teachers precluded the utilization of teacher ratings of academic ability. The lack of availability of standardized tests in each of the academic areas considered, in addition to practical (i.e., the time required to administer tests yielding scores in each of the academic areas) and ethical considerations (i.e., the time constraints imposed by school officials on the data collection process, and the need to minimize the impositions on teaching time), precluded the utilization of standardized tests as a measure of AA. Accordingly, the two AA indicators obtained for the present study were (1) the children's most recent grades in each of the common academic areas, obtained from their most recent school records; and (2) their self-reports of general school achievement, and of their grades in each of the academic areas outlined in Table 2. Thus, report card grades represented one indicator variable, and self-reported grades represented the other indicator variable; indicator variables for only the 11 common academic areas were included in the present analyses.

Grades

At the request of school officials, the children's grades were utilized as one indicator of AA. In most instances, the grading classification system (as represented in report cards) for the participating school boards provided for only one grade in each of the targeted academic areas. However, for some of the academic areas, the report cards for one or two school boards allowed for a greater level of differentiation in grading than was available in the other school board(s). For example, one school board permitted teachers to provide two grades in Religion, as represented by sub-categories under this academic area. In order to maintain consistency of indicator variables across the three school boards, a composite AA indicator variable was calculated within each school board when grades were more differentiated than the academic areas targeted for measurement. Specifically, as was

necessary, the mean of the children's grades were computed when multiple grades were available for a single common academic area. Thus, in the example above, for each participating child within the school board, the mean of the two grades for Religion was calculated to form a single indicator variable representing the grade indicator variable for Religion.

The use of grades as a measure of AA has a long history in between-network investigations of SC. In fact, theory and empirical research has suggested that children's academic SCs are derived more from school performance indicators than from standardized ability tests (e.g., Hansford & Hattie, 1982; Marsh, 1990a; Wylie, 1979). The rationale for this argument is that students, although having frequent feedback on, and access to their school performance, typically do not have access to standardized test results. Thus, logically, children's grades would be expected to have a greater impact on the formation of their academic SCs than other indicators of academic performance to which they typically do not have access. Accordingly, in addition to the practical advantages of using grades, use of this type of AA indicator offered theoretical advantages.

Self-report of Grades

For the second indicator of AA, a self-report measure of the children's grades was developed. This measure, the Children's Report of School Grades, required that each participant indicate their most recent grade in each of the specific academic areas tapped, in addition to their general school performance. Although all three school boards utilized five grading categories in their report cards, the associated labels (i.e., whether letters or numbers) and descriptive statements (i.e., very good, good, etc.) varied across school boards. In order to be consistent with the report cards the children obtained, the grading category labels and descriptive statements were adapted within each school board in order to correspond to those utilized by the school board (see Appendix C for the instructions and response format of this measure). Given that this self-report measure was developed for the present study, data relevant to reliability and validity are not available. However, given (1) the simple and straightforward nature of this instrument; and (2) its high face validity,

particularly with regard to its correspondence with the report card format of the specific school boards from which the students were drawn, measurement problems associated with the use of this instrument were not anticipated.

Theoretically, children's self-reports of their AA are of interest in examining relations between SC and AA constructs. Although empirical research has suggested that preadolescent children have a tendency to overestimate their abilities, this tendency has also been found to decline across the elementary school years (e.g., Frey & Ruble, 1987; Harter, 1990; Stipek & Daniels, 1988; Stipek & MacIver, 1989). It was also expected that the specific request that children report their most recent grades, rather than their perceived abilities, would minimize this tendency. Moreover, self-report measures of grades have been utilized as indicators of AA (e.g., Byrne & Worth Gavin, 1996; Watkins et al., 1991), and are justified from a theoretical perspective. In particular, within the I/E Model developed to explain the low correlations observed between Reading and Mathematics SCs (despite higher correlations among their corresponding achievement measures), it is posited that children's self-perceptions rather than objective indicators of AA contribute to the formation of their academic SCs. Accordingly, a self-report measure of grades allowed for the children's recollections of their AA to be reflected in each of the latent AA constructs. Thus, the utilization of self-reports of grades as one component of the AA construct is an important and theoretically relevant aspect of this construct.

Academic Importance Ratings

Although the measures of the importance construct did not form part of the analyses for the present study, for the sake of completeness, information regarding these instruments will be briefly reviewed. Specifically, participants completed two measures tapping the importance of doing well in each of the specific academic subject areas outlined in Table 2, as well as the overall importance of doing well in school. One instrument, the Children's Ratings of the Importance of School Subjects, was developed for inclusion in the present study. It contains one item asking the children to rate, on a 5-point scale, the importance of doing well in each of the targeted specific school subjects, and in school overall (see

Appendix C). The second instrument was adapted from the Harter Importance Rating Scale (Harter, 1985), an instrument designed to tap the importance of the domains measured by the SPPC; the original instrument contained a total of 10 items (two for each of the five specific domains examined with the SPPC). The two Scholastic Importance scale items were adapted to measure this construct in each of the specific academic areas (as outlined in Table 2), in addition to the importance of doing well in school in general (see Appendix C for the items and response format of this instrument).

Data Analyses and Research Hypotheses

Analyses were conducted in several stages in order to address the hypotheses to be tested. The first stage involved preliminary analyses, which included data preparation, determining descriptive statistics, and tests of instrument validity. The second stage involved the main analyses, and thus, the testing of the hypotheses specified in the present study. The EQS statistical analysis program (Bentler, 1992) was used in all analyses and, in particular, to formulate the factor analytic model necessary to test the specified hypotheses. The main analyses were conducted using CFA procedures within the framework of covariance structure modeling. We turn now to a review of the preliminary analyses conducted.

Preliminary Analyses

Preliminary analyses initially involved data preparation, which included dealing with missing data, reverse scoring appropriate items, creating composite AA indicator variables for which multiple grades were available for a targeted academic area, and creating SC item pairs. With regard to missing data, variables having more than 5% missing data, and cases having more than 5% missing data at either of two levels were excluded from analyses. Specifically, participants having more than 5% missing data across all items, or within the (A)SDQ, the SPPC, and the AA or importance construct were identified and excluded from further analyses. Regression analyses were utilized, separately for girls and boys, to impute values for the remaining missing data; in all cases, items from the corresponding scale of the same instrument were utilized in the regression analyses.

Where appropriate, items were reverse scored in order to render their scoring direction consistent with all other variables measuring the same construct set. The SC items were scored such that a low score represented a low SC and a high score represented a high SC. This required that a number of SPPC items be reversed scored; the five (A)SDQ-I items were all positively worded and thus, did not require reverse scoring. Grades and self-reports of grades were scored consistent with the system used by the school boards; in contrast to the SC items, a high number represented a lower grade (or self-reported grade) and a low number represented a higher grade (or self-reported grade). However, to facilitate interpretation, results for the main analyses will be presented consistent with the meaning of the constructs. That is, relations in which higher grades correspond to higher SCs will be represented as positive associations, and relations in which higher grades correspond to lower SCs will be represented as negative associations.

Data preparation included the creation of composite AA indicator variables, as was previously indicated, and SC item-pairs. Consistent with the utilization of the (A)SDQ instruments (e.g., Marsh, 1992b, 1993b), and as argued in the SDQ instrument manuals (Marsh, 1992b, 1992c, 1992d), whenever possible, analyses were conducted using item-pairs. The rationale for the use of item-pairs involved several issues (Marsh, 1993b; Marsh, Barnes, et al., 1984; Marsh & O'Neill, 1984). Specifically, the creation of item-pairs offered the dual advantage of reducing the complexity of the model(s) to be tested and increasing the subject-to-variable ratio. Theoretically, item-pairs reduce the occurrence of sample-specific artifacts due to idiosyncrasies associated with particular items, thereby increasing the stability of the indicators; the net effect should be more reliable and valid indicators with smaller unique variances. Finally, from a practical and economical perspective, the use of item-pairs substantially reduces computational requirements. Given the complexity of the models under study (i.e., the most complex models tested involved 199 single items), and the fact that the use of single items would exceed the limitations of currently available software programs (i.e., the EQS program limits the number of variables and factors to a total of 150), practical considerations precluded the utilization of single items.

The use of item-pairs, as has been noted by Marsh (e.g., Marsh, 1993b; Marsh, Barnes, et al., 1984) and others (e.g., Byrne & Shavelson, 1986), however, requires the assumption of homogeneity with respect to the items measuring each scale. Accordingly, in recognition of this assumption, items were grouped on the basis of the similarity of their content. For example, Items 3 and 5 of the Physical Ability SC scale were selected to form an item pair because both items referred to children's general physical ability in sports and as athletes; in contrast, the other items from this scale referred to discrete physical activities or characteristics. In addition, given that the SPPC contained items that required reverse-scoring, whenever possible, these items were balanced across item-pairs. With one exception, all SC indicator variables represented item-pairs; due to an uneven number of items in the ASDQ-I academic scales (i.e., five items per scale), two ASDQ-I indicator variables represented paired items, and one was a single item indicator.

The primary disadvantage with respect to the use of item-pairs, as opposed to single item analysis, is the loss of information relevant to specific items (e.g., Byrne & Shavelson, 1987; Marsh, 1993b). A resulting disadvantage of this approach is that further refinement of an instrument is precluded. Single item analyses, therefore, are required during the instrument development and validation phase of research, when information specific to particular items is of prime concern. However, when the focus of research is the validation of a construct, and instruments having adequate psychometric properties are utilized, the advantages of item-pair analysis outweigh the loss of information specific to single items, particularly with complex models. Despite the practical impossibility of single item analysis, item-pair analysis was considered a more appropriate choice than was single item analysis given the construct validation focus of the present study, the complexity of the models examined, and the instrument validation research underlying the SPPC and (A)SDQ-I instruments.

The final phase of data preparation involved the identification of multivariate outliers. Multivariate outliers were identified utilizing Mardia's measure of multivariate kurtosis (Bentler, 1992; for a review of options available for identifying multivariate outliers

in structural equation modeling, see West, Finch, & Curran, 1996). Specifically, the EQS program provides for the identification of the five cases demonstrating the largest contribution to multivariate kurtosis, as based on Mardia's measure of multivariate kurtosis (Bentler, 1992; Byrne, 1994). Cases whose values are large relative to those of the remaining cases identified are indicative of outliers; these cases are excluded and analyses rerun until outliers are no longer identified. This process was used, separately for girls and boys, for the identification and elimination of outliers.

Once data preparation was completed, the descriptive statistics were computed, for girls and boys separately, for each of the SC scale scores, as well as the SC and AA indicator variable scores. The descriptive statistics obtained included the mean, standard deviation, skewness, and kurtosis values. Skewness and kurtosis values were examined in order to determine whether the data were normally distributed, that is, whether the skewness and kurtosis values fell within the -1.00 to 1.00 recommended range for normality (Muthén & Kaplan, 1985).

The determination of whether the indicator variable scores were normally distributed had important implications for the indices used in the evaluation of the hypotheses. Specifically, when the data, or parts of the data cannot be considered normally distributed, the violation of the assumption of multivariate normality underlying the method of model estimation, and the associated fit indices in CFA dictates the use of their robust (i.e., corrected for distributional non-normality) versions. More specifically, the Satorra-Bentler chi-square serves as a correction to the chi-square statistic when data are highly kurtotic (Satorra & Bentler, 1988, 1994). In addition, this corrected chi-square value can be used in calculating certain fit indices, and thus, they may similarly be considered to be robust indices (see below for further elaboration). On the basis of these inspections, robust estimates and fit indices were utilized when the EQS program allowed; robust estimates and fit indices are not available in multigroup analyses, the form of analysis required when testing for invariance across samples. Thus, with the exception of the tests of invariance, robust estimates and fit indices were utilized.

The final phase of preliminary analyses involved an examination of the SC instruments. Specifically, coefficient alphas were computed for each of the SC scale instruments, as well as across each of the instruments. In addition, CFAs were conducted separately for girls and boys in order to ascertain that the instruments were measuring the SC dimensions represented in the main hypotheses. These analyses were initially conducted for the SC instruments both separately and in combination. As a result of these analyses, the SPPC was excluded in tests of the main hypotheses (see further elaboration in the results section). We turn now to a more detailed review of the hypotheses specified and the analyses conducted in testing the hypotheses.

Research Hypotheses

In this section, the hypotheses, and their theoretical and empirical underpinnings, are presented along with the general framework within which the analyses were conducted. As previously noted, all analyses were conducted separately for girls and boys. A review of these analyses and the criteria used in evaluating the hypotheses are presented subsequent to the review of the hypotheses, to which we now turn.

Hypothesis 1: For both girls and boys, preadolescent self-concept is a multidimensional construct

Based on the review of the literature and for purposes of the present study, multidimensionality was comprised of general-academic SC, the four specific nonacademic SCs, and the 11 specific academic SC dimensions common to the three school boards. In particular, the specific academic SC dimensions included Arithmetic, Measurement, Science, Social Studies, Religion, Art, Gym, as well as Listening, Reading, Writing, and Speaking in Language Arts. The nonacademic components included Physical Ability, Physical Appearance, Parent, and Peer SC facets. The CFA model of multidimensional SC hypothesized a priori, for both preadolescent girls and boys, that (a) SC was best represented as a 16-factor model comprising each of the aforementioned SC facets; (b) each indicator variable would have a non-zero loading on the factor it is designed to measure and a zero loading on the non-target factors; (c) the factors would be correlated, with the exception that

the correlations among the Mathematics and Language Arts SC facets may approach zero; and (d) the measurement error terms would be uncorrelated.

Hypothesis 1 was tested by comparing the goodness-of-fit of the hypothesized 16-factor model of SC with two less differentiated multidimensional SC models. In particular, the fit of the hypothesized 16-factor model of SC (see Figure 4 for a pictorial representation of the core academic portion of this model) was contrasted with the fit of both a 15- and a 13-factor model of SC (see Figures 5 and 6, respectively). The 15-factor model of SC differed from the 16-factor model in that the two specific mathematics SC dimensions (i.e., Measurement and Arithmetic) were represented as one Mathematics factor, rather than two separate and more specific SC factors. The 13-factor model differed from the hypothesized model in that the four Language Arts SC facets (i.e., Reading, Writing, Speaking, and Listening) were represented as one, rather than four SC facets. Hypotheses will be supported if the 16-factor model provides a better fit to the data than either the 15- or 13-factor models for both girls and boys.

Hypothesis 2: For both girls and boys, preadolescent self-concept is an hierarchically structured construct

All second-order (i.e., hierarchical) CFAs were conducted for each gender using the best-fitting models determined by the first-order analyses; by convention, these final models are termed baseline models. A total of four second-order SC models were hypothesized and tested for their goodness-of-fit to the sample data, each of which assumed that Hypothesis 1 had been supported. Thus, in each of these models, general-academic SC and each of the specific academic and nonacademic SC components delineated in Hypothesis 1 appeared at the base, or first level of the SC hierarchy. The four second-order hierarchical models of preadolescent SC differed only with regard to the second level of the SC hierarchy. Specifically, the relations specified between the first- and second-order SC facets, and the nature or number of second-order SC facets changed with each successive model. These hierarchical models are represented pictorially in Figures 7 to 10. The four hierarchical models specified were hypothesized utilizing the Marsh/Shavelson Model of SC (Marsh &

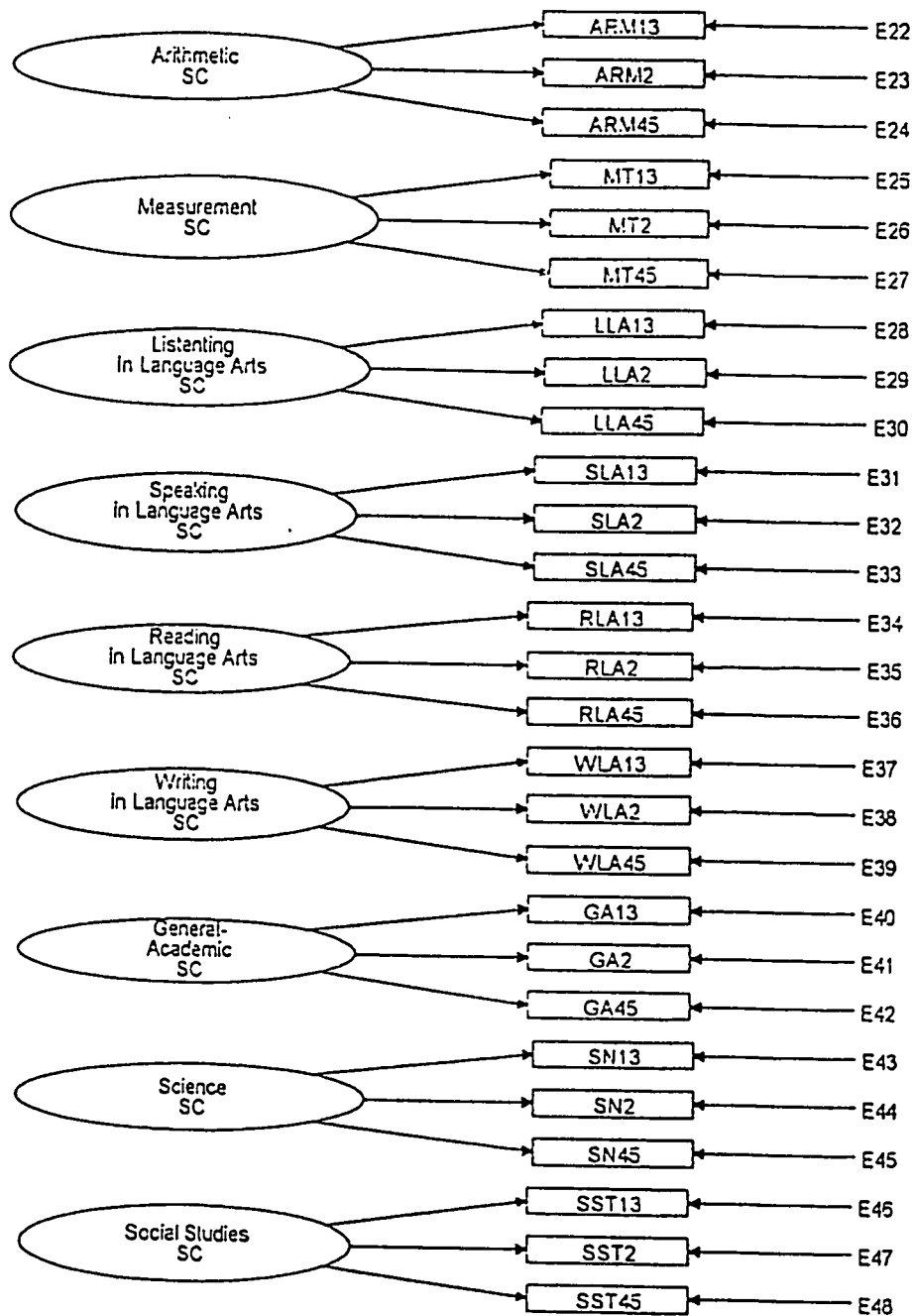


Figure 4. A pictorial representation of the core academic portion of the first-order 16-factor self-concept (SC) model.

Note. Boxes represent indicator variables for the corresponding first-order SC factor, with the letters representing this same factor, and the numbers representing the ASDQ-I item number(s) from which they were derived. Thus, for example, SN13 represents the indicator variable derived from the ASDQ-I Science scale Item numbers 1 and 3.

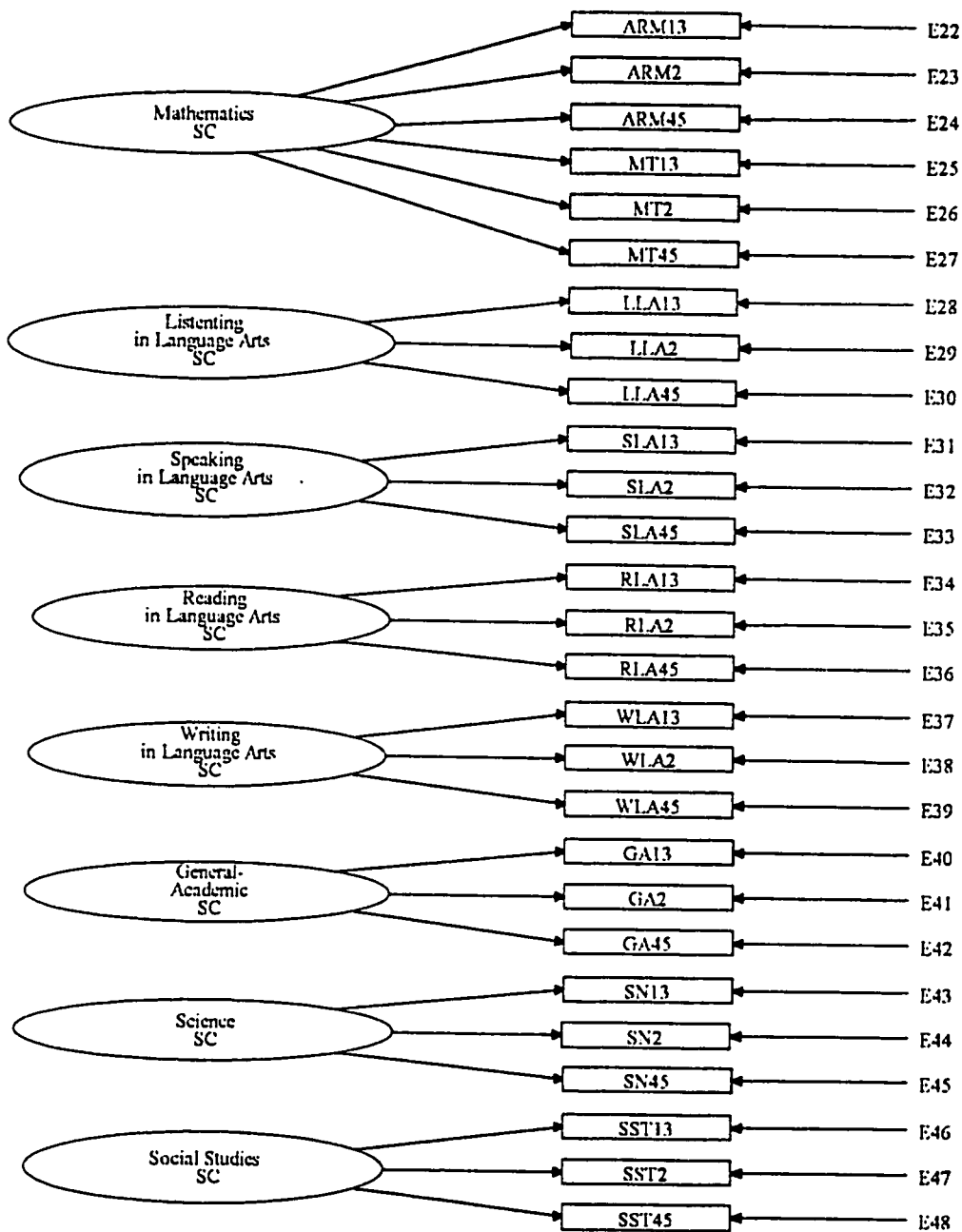


Figure 5. A pictorial representation of the core academic portion of the first-order 15-factor self-concept (SC) model.

Note. Boxes represent indicator variables for the corresponding first-order SC factor, with the letters representing this same factor, and the numbers representing the ASDQ-I item number(s) from which they were derived. Thus, for example, SN13 represents the indicator variable derived from the ASDQ-I Science scale Item numbers 1 and 3. However, the indicator variables for Mathematics SC are derived from the ASDQ-I Arithmetic and Measurement scales, which are abbreviated ARM and MT, respectively.

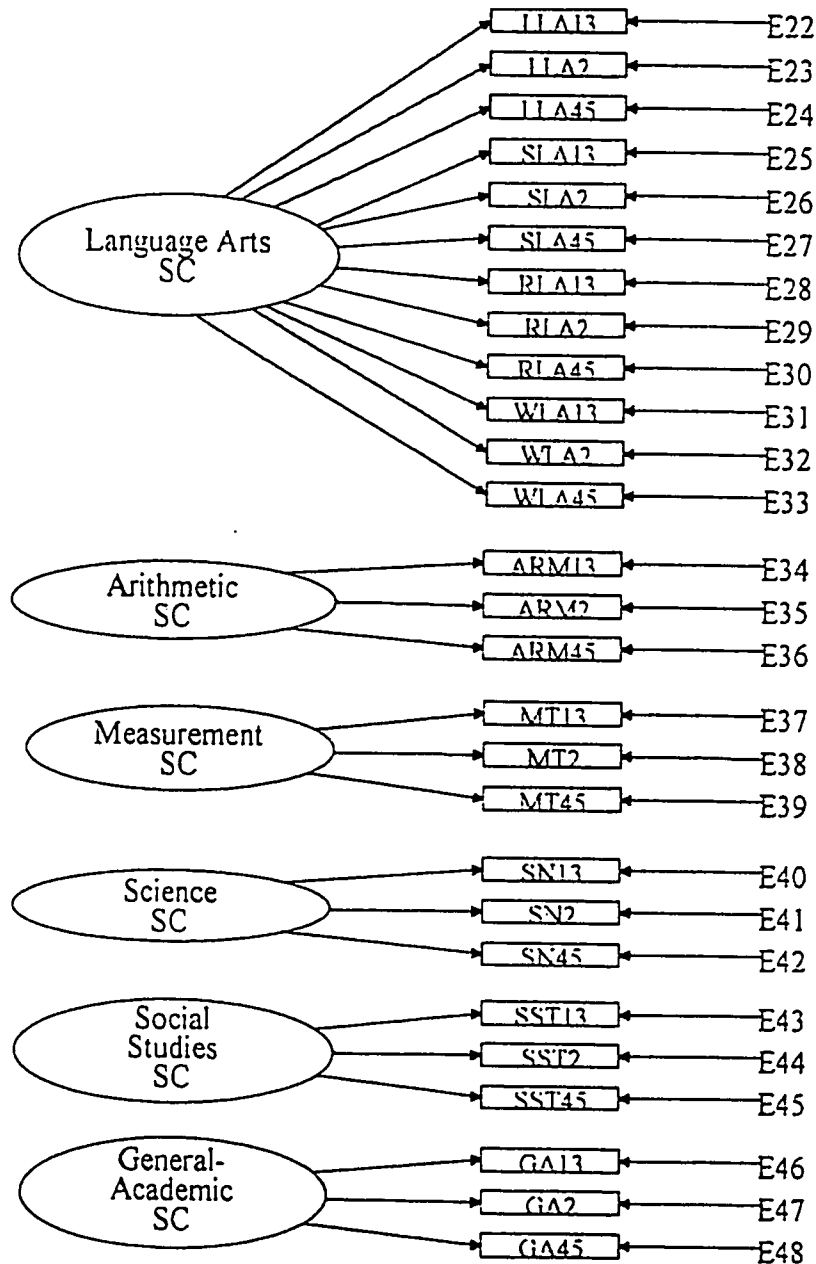


Figure 6. A pictorial representation of the core academic portion of the first-order 13-factor self-concept (SC) model.

Note. Boxes represent indicator variables for the corresponding first-order SC factor, with the letters representing this same factor, and the numbers representing the ASDQ-I item number(s) from which they were derived. Thus, for example, ARM13 represents the indicator variable derived from the ASDQ-I Arithmetic scale Item numbers 1 and 3. However, the indicator variables for Language Arts SC are derived from the ASDQ-I Listening, Speaking, Reading, and Writing in Language Arts SC scales, which are abbreviated L/S/R/WLA, respectively.

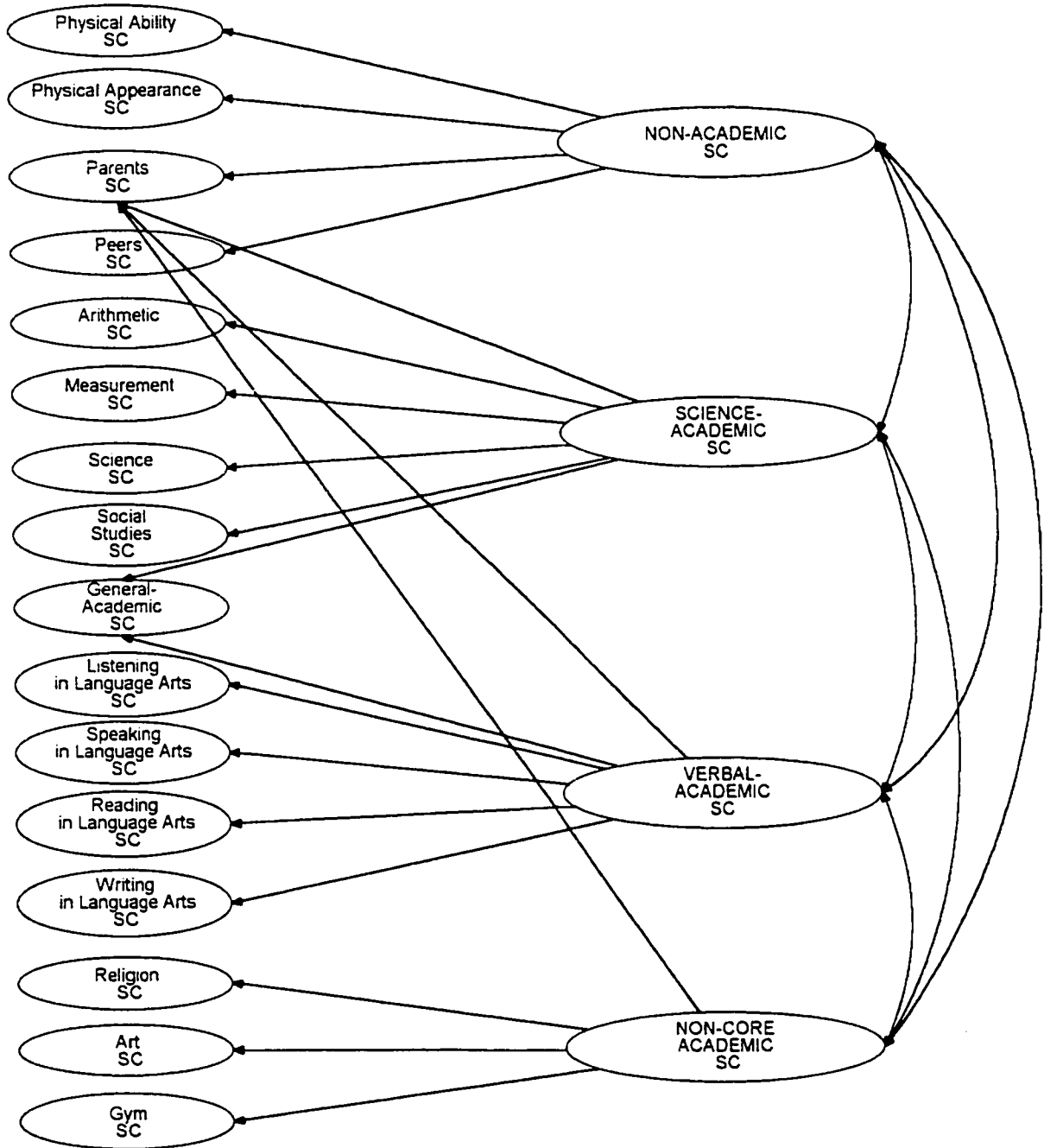


Figure 7. A pictorial representation of the structural portion of hierarchical self-concept (SC) Model H1.

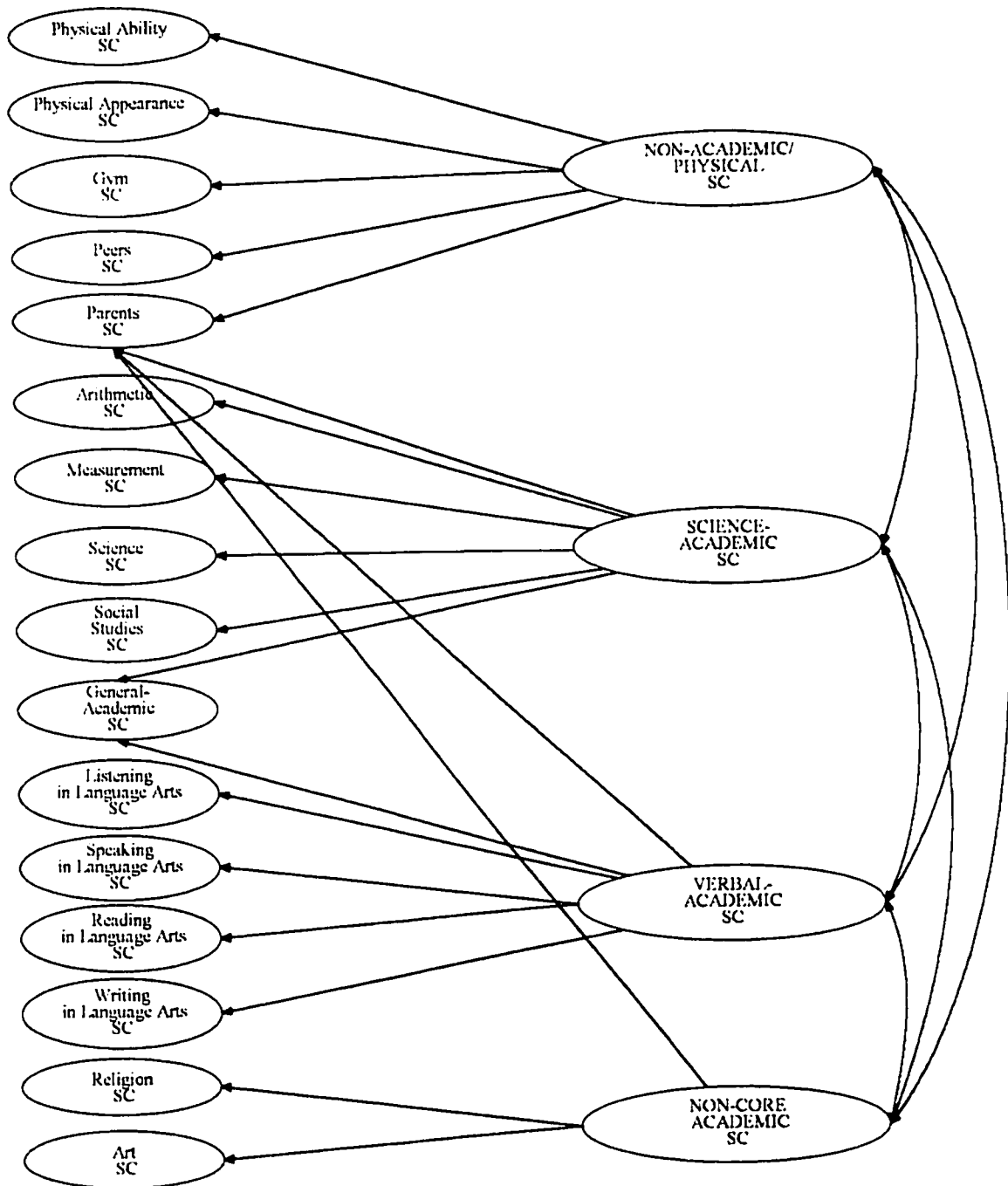


Figure 8. A pictorial representation of the structural portion of hierarchical self-concept (SC) Model H2.

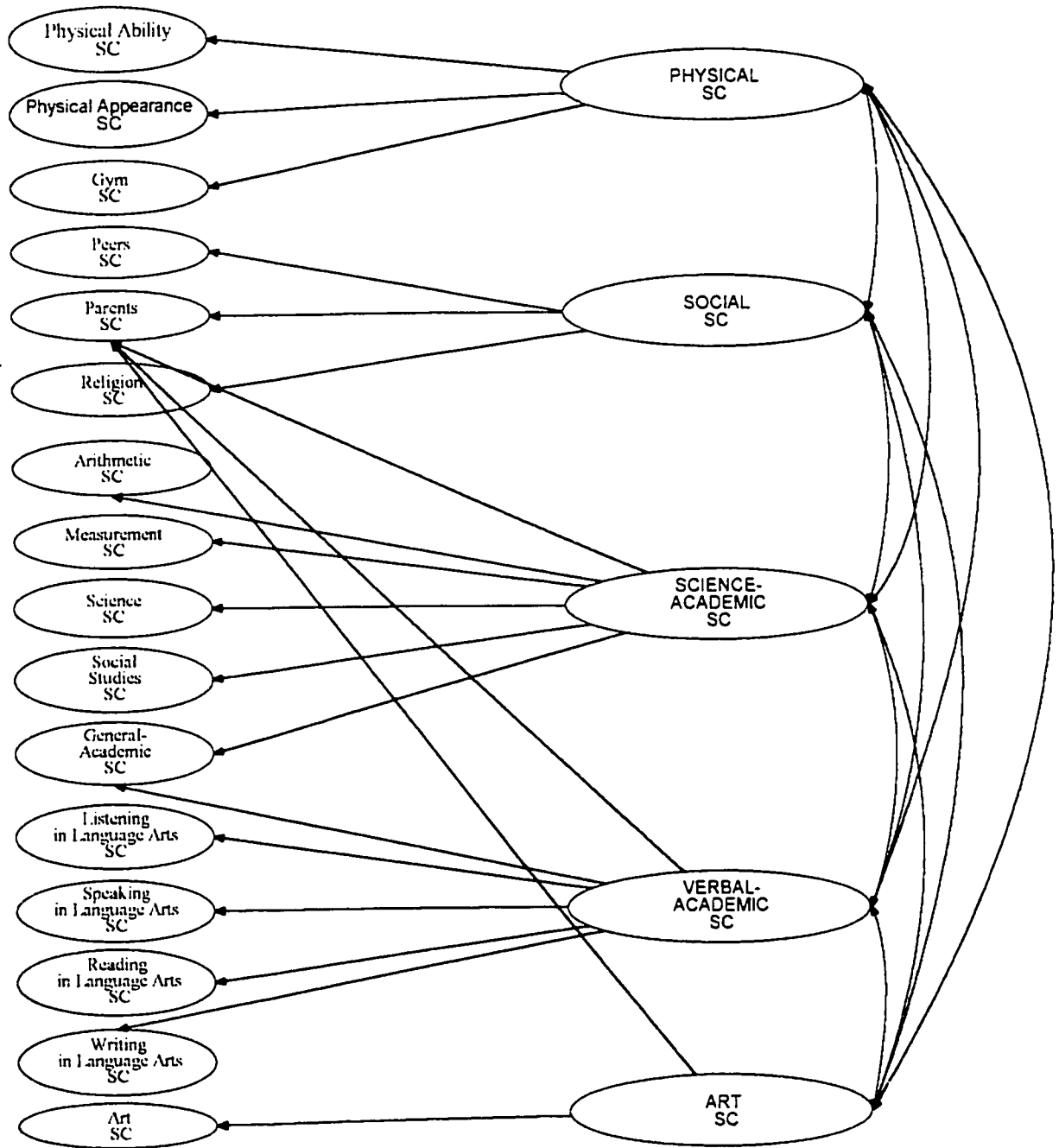


Figure 9. A pictorial representation of the structural portion of hierarchical self-concept (SC) Model H3.

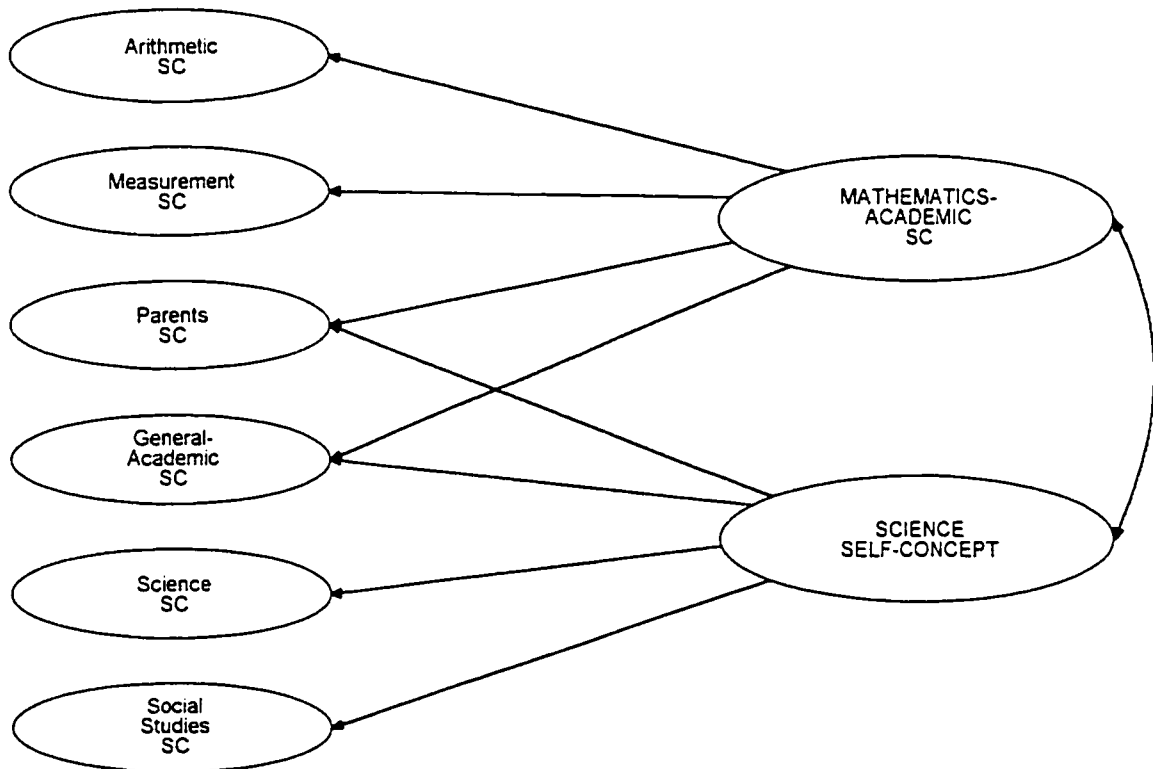


Figure 10. A pictorial representation of the structural portion of the hierarchical self-concept (SC) Model H4 expansion.

Shavelson, 1985) as their basis, and expanded on the basis of theory and the study (Marsh, 1990d) of an expanded set of academic SC facets. We turn now to an elaboration of each of these models, and the theoretical and empirical basis underlying their specification.

Model H1. The first hierarchical model of SC postulated (see Figure 7) was, with two exceptions, an exact duplication of the Marsh/Shavelson model of SC. First, the suggestion that the second-order Mathematics-Academic SC facet is more appropriately labelled a science facet was incorporated into this model (Marsh, 1990d). Second, incorporating research demonstrating that this model does not adequately account for relations among first-order non-core academic SC facets (Marsh, 1990d), an additional second-order factor was hypothesized to represent relations among the first-order non-core academic SC facets measured (i.e., Religion, Art, and Gym). Thus, the adequacy of just one additional second-order SC facet in explaining relations among an expanded set of academic SC facets was tested in this model. Accordingly, four second-order factors were hypothesized to represent the relations among the first-order constructs; these were second-order Nonacademic, Science-Academic, Verbal-Academic, and Non-Core Academic SCs.

Consistent with the Marsh/Shavelson model and SDQ research, the second-order Nonacademic, Science-Academic, and Verbal-Academic SC facets were hypothesized to define the same first-order SC facets identified in the Marsh/Shavelson Model, albeit with greater specificity of academic content areas. In particular, the second-order Nonacademic SC construct was postulated to define the two first-order physical (i.e., ability and appearance) and the two first-order social (i.e., Parent and Peer) SC facets. The second-order Science-Academic SC facet was posited to define the two first-order mathematics (i.e., Arithmetic and Measurement) and the two first-order science (i.e., Science and Social Studies) SC constructs. The second-order Verbal-Academic SC facet was hypothesized to define the four Language Arts (i.e., Listening, Speaking, Reading, and Writing) SC factors. As in the Marsh/Shavelson model, and consistent with SDQ research, the two core Verbal- and Science-Academic second-order SC facets also defined the first-order General-Academic SC facet. The second-order Non-Core Academic SC construct, which was not

posited in the Marsh/Shavelson Model, was postulated to define first-order SCs in the non-core academic areas, namely, Religion, Art, and Gym. Finally, consistent with SDQ research showing relations between the Parent SC facet and both nonacademic and academic SC facets, Parent SC was postulated to be defined by all second-order academic SC facets in addition to nonacademic SC.

Model H2. Model H2 (see Figure 8) was postulated to empirically examine the suggestion that Gym SC might more appropriately be represented under the same second-order SC facet defining the two other physical SC facets (i.e., ability and appearance; Marsh, 1990d). Thus, Model H2 differed from Model H1 only in that the first-order Gym SC facet was defined by the second-order Nonacademic SC facet, rather than the Non-Core Academic SC facet. To more appropriately reflect its changing configuration, this second-order SC facet was relabelled Nonacademic/Physical SC.

Model H3. The third hierarchical model (see Figure 9) represented a more differentiated model with respect to the nonacademic portion of the SC hierarchy. Consistent with the previous hierarchical models, this model examined the adequacy of the Marsh/Shavelson model in accounting for relations among the first-order SCs with just one additional second-order academic SC facet. However, an additional second-order nonacademic SC facet was hypothesized. Thus, in contrast to the four second-order SC facets hypothesized in hierarchical Models H1 and H2, Model H3 specified five second-order SC facets. In Model H3, an alternate second-order representation of the first-order non-core academic SC facets, and an associated reconceptualization and increased second-order differentiation of the nonacademic SC facets was postulated. Specifically, the second-order Non-Core Academic SC facet was relabelled Art SC and was defined by the first-order Art and Parent SC facets, the inclusion of the latter SC being consistent with previous research demonstrating relations between the Parent and academic SC facets. The remaining non-core academic SCs (i.e., Gym and Religion) were represented under conceptually similar but more differentiated second-order nonacademic SCs. Specifically, Gym SC was represented, along with the first-order Physical Ability and Physical Appearance SC facets, under a

second-order Physical SC facet. Religion SC (as reflected in the three school boards' labelling of this subject area; see Table 2) was represented under a second-order Social SC facet that was defined by the two first-order social (i.e., Peer and Parent) SC facets.

Model H4. The final model (see Figure 10) represented an increased level of SC differentiation at the second-order level of the core academic portion of the Marsh/Shavelson model. Given the increased specificity with which science and mathematics SCs were measured in the present study, this model tested whether these facets were more appropriately represented separately at the second level of the SC hierarchy. Thus, the two specific first-order mathematics SCs (i.e., Arithmetic and Measurement), and the two first-order science SCs (i.e., Science and Social Studies) were each postulated to form, along with General-Academic SC, their own respective SC dimensions at the second level of the SC hierarchy (i.e., Mathematics- and Science-Academic SCs). This model was incorporated into the best-fitting of the three previous hierarchical models. Thus, the exact formulation of Model H4 was dependent upon which of the previous hierarchical models provided the best theoretical and statistical representation of the SC construct for each gender.

The a priori hypotheses with regard to the hierarchical structure of SC, for both preadolescent girls and boys, were as follows. Model H2 was expected to provide a better fit to the data than Model H1. This hypothesis was grounded in research which demonstrated the need to represent the first-order Gym SC facet separate from the non-core Art and Religion SC facets at the second level of the SC hierarchy (i.e., under the rubrics of Physical Education and Art SC, respectively; Marsh, 1990d). Given the lack of empirical research examining the hierarchical structure of preadolescent SC with first-order SCs comprising both nonacademic and an expanded set of academic components, hypotheses regarding the differential goodness-of-fit of the three remaining models were not specified. Hypothesis 2 was tested by comparing the relative fit of Model H1 to H4 separately for girls and boys. Hypotheses will be supported if Model H2 provides a better fit to the data than does Model H1 for both girls and boys.

Hypothesis 3: The within-network structure of preadolescent self-concept is invariant across gender

Consistent with CFA research with Australian preadolescents (Marsh, 1987b; Marsh, Smith, & Barnes, 1985), measurement and structural SC relations were expected to be equivalent across gender. The CFA model hypothesized a priori that corresponding estimated factor loadings, regression paths linking first- and second-order factors, as well as second-order factor covariances would be invariant across the girl and boy samples.

Hypothesis 3 was tested by first establishing a best-fitting first-order multidimensional and second-order hierarchical SC structure separately for girls and boys under Hypothesis 1 and 2, respectively. Within the CFA framework, tests of the invariance of the fit of a model across samples require that the fit of the model to each of the samples be estimated simultaneously. The invariance of parameters are supported if the parameters constrained equal are not found to be significantly different from each other. Thus, the hypothesis of invariant preadolescent SC measurement and structure will be supported if the parameters specified as invariant are not identified as significantly different from each other. As previously noted, currently available software, however, does not permit the derivation of robust estimates and fit indices in multigroup analyses. Thus, tests of the invariance of parameters, and estimates of the adequacy of the fit of models of invariance can currently only be based on non-robust parameter estimates and fit indices. Accordingly, this represents a limitation in current methodology, and may reduce one's confidence in the results obtained when data are highly non-normally distributed, or for parameters that are non-normally distributed.

Hypothesis 4: For both girls and boys, relations among preadolescent latent self-concept and academic achievement constructs will be consistent with the multidimensional and hierarchical structure posited in the Marsh/Shavelson model

With regard to the hypotheses of multidimensional and hierarchical SC structure, SC research has provided strong support for the multidimensionality of SC. Between-network SC research has similarly provided support for a stronger pattern of relations between AA

and generic measures of academic SC, as compared to relations between AA and general SC. Support for a stronger pattern of relations between specific AA and academic SC facets, as compared to that between specific AA measures and more generic academic SC facets has, however, been inconsistent. None of these research efforts contrasted the latent correlations between AA constructs and first-order SC facets against the correlations between AA constructs and second-order SC facets estimated within the context of CFA, as was done in the present study. Such a comparison represents a test of the hierarchical characteristic of the Marsh/Shavelson model that is more representative of the hypotheses relevant to this aspect of the model. Accordingly, it is possible that the inconsistency in empirical research bearing on an hierarchically structured SC is due to the nature of the data upon which this criterion was assessed. As a result, the hypotheses were specified consistent with the predictions derived from the Marsh/Shavelson model.

The hypotheses with regard to the pattern of correlations among latent SC and AA constructs were as follows. Specifically, consistent with the multidimensionality of SC, latent SC and AA constructs, although intercorrelated to varying extents, were hypothesized to be measurable as distinct constructs. In addition, latent AA constructs in particular academic areas were hypothesized to be more highly correlated with latent academic SC constructs in matching content areas than with those in other academic areas. Support for the differentiability of SC and AA constructs was assessed by examining the magnitude of the correlations among these constructs. Support for the content-specificity of multidimensional SC/AA relations was assessed by contrasting the magnitude of the latent correlations between corresponding subject-specific SC and AA constructs against the magnitude of the latent correlations among non-corresponding SC and AA constructs. Hypotheses will be supported if, for both the sample of girls and of boys, the latent correlations are consistent with the differentiability of the latent factors and the specified pattern of relations.

The hierarchical structure of SC similarly implies a systematic pattern of correlations among SC and AA constructs. Specifically, latent AA constructs were hypothesized to be more highly correlated with corresponding latent first-order SC constructs than with

corresponding second-order latent SCs. Stated in reverse, second-order academic SC facets were hypothesized to be less highly correlated with corresponding latent AA constructs than were corresponding first-order SC constructs. This hypothesis was examined by comparing the magnitude of the factor correlations among the specified constructs. Support for an hierarchical pattern of SC/AA relations will be demonstrated if the latent correlations are consistent with the specified pattern for both girls and boys. Accordingly, the fit of the models did not constitute the data upon which this hypothesis was evaluated. Rather, the feasibility (i.e., that the values of the latent correlations did not imply a lack of distinction among constructs) and the pattern of correlations represented the information upon which the hypothesis of multidimensional and hierarchical SC/AA relations was assessed.

Within the CFA framework, the simultaneous estimation of latent correlations between one set of constructs (i.e., AA), and both the first- and second-order level of a second set of constructs (i.e., SC) is not technically permissible. Rather, the estimation of latent correlations between the AA constructs and the first-order SC constructs must be conducted separately from the estimation of the latent correlations between the AA constructs and the second-order SC constructs.

Hypothesis 5: Preadolescent self-concept/academic achievement relations, as described under Hypothesis 4, will be equivalent across gender

The only research to examine this aspect of the SC construct within a CFA framework was conducted with adolescents (Byrne & Shavelson, 1987). Byrne and Shavelson (1987) found gender differences in both the within- and between-network multidimensional and hierarchical pattern of SC/AA relations. However, research examining the validity of the I/E model has suggested that the relations between reading and mathematics SC and AA indices are not significantly altered by gender for preadolescents (Marsh, Smith, & Barnes, 1985). Thus, it seemed likely that Byrne and Shavelson's results were specific to adolescents and related to a developmental phenomenon. Accordingly, SC/AA relations were hypothesized to be equivalent across gender for preadolescents.

On the basis of results from the test of the invariance of the measurement and structure of SC outlined under Hypothesis 3, the CFA model hypothesized and tested for the gender invariance of (a) the measurement and structural SC parameters found to be invariant under Hypothesis 3, and (b) the corresponding SC/AA factor covariances. Hypotheses will be supported if the parameters constrained to be equal across girls and boys are not identified as significantly different from each other.

Tests of Hypotheses: Model Evaluation and Modification

In CFA, two central issues confront the researcher. The first of these is the estimation method to be used. The second issue is the criteria to be used in testing the hypotheses under study (for reviews of this methodology, see Long, 1983a, 1983b; West et al., 1996). That is, subsequent to the postulation of an a priori model, the hypothesized model is fit to the sample data, and the model parameters are estimated so as to minimize the discrepancy between the model and the sample data; the larger this discrepancy, the poorer the fit of the model to the data. Several methods of estimation are available to the researchers in fitting the model to the data (e.g., maximum likelihood, generalized least squares). Of these, the most widely used and recommended approach is the maximum likelihood method (Hoyle & Panter, 1995; Sugawara & MacCallum, 1993). The maximum likelihood method has been demonstrated to generally perform most adequately and consistently across a wide variety of analytic conditions (e.g., Hu, Bentler, & Kano, 1992). Accordingly, maximum likelihood estimation was selected as the estimation method in the present study.

With regard to the second issue, that of selecting the criteria to be used in testing the hypotheses under study, an overall indication of the discrepancy between the model and the data is given by the chi-square likelihood ratio test (Hu & Bentler, 1995). Specifically, the chi-square statistic is a test of the probability with which a researcher can have confidence in the tenability of the null hypothesis (i.e., that the model perfectly represents the sample data). Under this approach, if the chi-square value of a model exceeds the critical value (at the designated alpha level and the model's degrees of freedom), the null hypothesis is not supported, and thus, the model cannot be considered to adequately represent the data.

Accordingly, unlike traditional statistical methods, a nonsignificant (i.e., smaller) chi-square value supports the adequacy of the hypothesized model in representing the sample data, and is thus, the desired outcome.

Although the availability of this statistic initially led researchers to conclude that the subjectivity inherent in exploratory factor analysis had been largely overcome with the development of confirmatory modes of analyses, subsequent research demonstrated that the evaluation of model fit was much more complex than the simple determination of the significance (or nonsignificance) of a model's chi-square value (Hu & Bentler, 1995). That is, numerous problems associated with the performance of the chi-square statistic under varying real-world conditions have since been identified. Most critical among these problems is the sensitivity of the chi-square statistic to sample size; research has demonstrated that the likelihood of rejection of a given model under the chi-square statistic increases as sample sizes increases (e.g., Boomsma, 1982; Bearden, Sharma, & Teel, 1982). This statistic has also been demonstrated to be sensitive to such factors as the complexity of the model, the degree of model misspecification, and violations of the assumptions underlying the method of estimation (e.g., multivariate normality; Bearden et al., 1982; Hu et al., 1992; La Du & Tanaka, 1989; for a review of these issues, see Hu & Bentler, 1995).

Given that structural equation modeling is grounded in asymptotic (i.e., large-sample) theory, the sensitivity of the chi-square statistic to variations in sample size mitigates its use as a dichotomous criterion in evaluating the fit of the model(s) under examination (e.g., Hu & Bentler, 1995). In addition, given that models can never completely represent the data, some degree of misfit is considered inevitable (Browne & Cudeck, 1993; MacCallum, 1995; MacCallum, Roznowski, & Necowitz, 1992). As a result, it is now generally accepted that the chi-square statistic is best regarded as an index of fit, rather than as a test statistic in the classical hypothesis-testing sense (e.g., Hu & Bentler, 1995; Joreskog & Sorbom, 1993; Long, 1983a, 1983b; Satorra & Saris, 1985). The central question in CFA, thus, becomes not whether the model fits the data, but rather, evaluating the extent to which the hypothesized model adequately represents the sample data.

In an attempt to address this issue and overcome the limitations associated with the chi-square statistic, a plethora of alternative fit indices have been developed to assess the degree of fit (i.e., congruence) or misfit (i.e., discrepancy) between the hypothesized model and the data (e.g., Bentler & Bonnett, 1980; Steiger & Lind, 1980; for reviews, see Browne & Cudeck, 1993; Hu & Bentler, 1995; Tanaka, 1993). These alternative indices of fit, however, have similarly been plagued with problems related to their sensitivity to such issues as sample size, model complexity, and violations of distributional assumptions (e.g., Bentler, 1990; Gerbing & Anderson, 1993; Mulaik, James, Van Alstine, Bennett, Lind, & Stidwell, 1989).

Given the problems associated with both the chi-square statistic and the alternative fit indices, sole reliance on any one criterion in evaluating model fit is considered to be statistically and theoretically untenable (Browne & Cudeck, 1993; Hoyle, 1995; Hu & Bentler, 1995; Tanaka, 1993). In addition, any information suggested by these “objective” indicators must be evaluated for its substantive meaningfulness (MacCallum, 1986). In particular, researchers have emphasized the need to consider both statistical (i.e., the chi-square statistic) and practical (i.e., alternative fit indices) criteria within a substantive context when evaluating the fit of a model, and most importantly, when considering whether and how to modify a model so as to achieve a better fit (e.g., MacCallum, 1986; MacCallum et al., 1992; Silvia & MacCallum, 1988). In addition, there has been a shift in an emphasis on the “fit” of a single model towards the importance of specifying multiple alternative, competing models and the evaluation of their comparative fit (e.g., Hoyle, 1995; Hoyle & Panter, 1995; MacCallum et al., 1992; MacCallum, Wegener, Uchino, & Fabrigar, 1993). Finally, researchers have emphasized the importance of considering the parsimony of each of the competing models in multiple model comparisons (e.g., Williams & Holahan, 1994). That is, should two models provide an equally good fit to the sample data from a substantive, statistical, and practical perspective, the more parsimonious model is preferred.

Given the sensitivity of both the chi-square statistic and alternative fit indices to sample size, the adequacy of sample size is an important consideration. In addition,

problems related to non-convergence and improper solutions have been found when samples are small (e.g., Boomsma, 1985). Although a methodology for determining the required sample size for a given level of power when testing models having a specific level of complexity is not available, analyses based on simulated (i.e., Monte Carlo) and real data have led to the specification of certain guidelines. In particular, a minimum sample size of 100 is considered necessary, and a preferred sample size of 200 has been recommended, with researchers noting that, in highly complex models, these requirements may need to be increased (e.g., Bearden et al., 1982; Boomsma, 1982, 1985). In the present study, the upper bound minimum recommended sample size of 200 was attained for each gender. We turn now to a review of the criteria used in evaluating the fit of the hypothesized models.

Model Evaluation

As has been emphasized in the literature, multiple multidimensional and hierarchical models of SC were specified and their goodness-of-fit compared. Thus, for Hypothesis 1, a series of decreasingly differentiated multidimensional models of preadolescent SC were specified and their differential goodness-of-fit to the sample data were contrasted for both girls and boys. Likewise, for Hypothesis 2, a series of hierarchical SC models were tested and contrasted for their goodness-of-fit to the observed data. In addition, multiple criteria were considered in evaluating and contrasting model fit under both Hypotheses 1 and 2. That is, substantive, statistical, and practical criteria were considered and, on the basis of the descriptive statistics reviewed later, robust versions of the estimates and criteria were utilized. Specifically, the criteria utilized in evaluating and contrasting the fit of the models under Hypotheses 1 and 2 included: (a) the substantive meaningfulness and parsimony of the models (e.g., MacCallum, 1986); (b) a scaled version of the chi-square statistic, the Satorra-Bentler chi-square statistic ($S-\beta \chi^2$; Satorra & Bentler, 1988, 1994), which corrects the maximum likelihood chi-square statistic for non-normal data; (c) the corrected (i.e., robust) version of the Comparative Fit Index (*CFI; Bentler, 1990); (d) the robust version of the Expected Cross-Validation Index (*ECVI; Browne & Cudeck, 1989); and (e) the Root Mean Square Error of Approximation (RMSEA; Steiger & Lind, 1980).

The Satorra-Bentler chi-square is a corrected version of the chi-square statistic developed for use with samples in which the distributional assumptions underlying maximum likelihood estimation are violated. The correction incorporates information relevant to the model tested, the estimation method, and the sample multivariate kurtosis values (Hu et al., 1992; West et al., 1996). Studies have demonstrated the superior performance of this statistic over that of the asymptotically distribution-free estimation methods under a variety of conditions and in all but excessively large samples, where performance was similar (e.g., Hu et al., 1992). In addition, the EQS program provides robust model parameter estimates with this statistic.

When models are nested within one another, the difference in chi-square values between the models can be used to evaluate the statistical significance of the variation in the fit of competing models. Models are considered to be nested when one model is a more restricted version of the other, and they contain the same number of factors (Hoyle, 1995). The difference in chi-square ($\Delta\chi^2$) values between nested models is itself chi-square-distributed, with the degrees of freedom equal to the difference in their degrees of freedom (Δdf). Thus, the chi-square/degrees of freedom differential ($\Delta\chi^2/df$) can itself be tested for significance, with a significant differential indicating a significant improvement (in the case where parameters are added) or decrement (in the case where parameters are removed) in model fit. When comparing nested models, this differential was used in contrasting the fit of both the first- and the second-order SC models and, in particular, in determining which of the specified models best represented the sample data.

The Comparative Fit Index (CFI; Bentler, 1990) is a revised version of an earlier fit index, the Bentler-Bonett normed fit index (Bentler & Bonett, 1980). The CFI is derived from the comparison of the hypothesized model with the null model (i.e., a model in which all indicator variables are uncorrelated). Its value ranges from zero to 1.00, with higher values indicative of a superior fit; values greater than .90 are considered to indicate that the model represents an acceptable fit to the data (e.g., Byrne, 1994; Hoyle, 1995). The value of the CFI can be interpreted to reflect the proportion of variation in the data that is explained

by the model (Hu & Bentler, 1995; Mulaik et al., 1989). A corrected version of this fit index (*CFI), in which the Satorra-Bentler chi-square statistic is substituted, was used in the present study.

The Expected Cross-Validation Index (ECVI; Browne & Cudeck, 1989) is a measure designed to assess, in a single sample, the likelihood that the estimated model will be replicated (i.e., will cross-validate) across samples of a similar size drawn from the same population. Specifically, it measures the discrepancy between the covariance matrix expected in another same-size sample and the current sample's fitted covariance matrix. As with the CFI, the robust version of this index (*ECVI) is calculated using the corrected version of the chi-square statistic (the $S-\beta \chi^2$). The ECVI (*ECVI) is a relative rather than a stand-alone index in that it assumes a comparison with ECVI values associated with other models. The model having the lowest value is the model with the highest expectation of replication; thus, the ECVI (*ECVI) is used to help in the selection of a model among a set of models. This index may be conceptualized as a measure of the comparative predictive validity of multiple models (Browne & Cudeck, 1989).

The final criterion used in determining the adequacy of the fit of the hypothesized models to the sample data was the Root Mean Square Error of Approximation (RMSEA; Steiger & Lind, 1980). The RMSEA is derived from the residual matrix generated by the discrepancy between the sample and model covariance matrix; it represents an overall measure of lack of fit per degree of freedom (Hu & Bentler, 1995; MacCallum, 1995). Given that more complex models, merely by virtue of the loss of degrees of freedom associated with their greater number of parameters, will generally yield a better fit to the data than less complex models, inclusion of an index that takes degrees of freedom into account and reflects the misfit of a model has been recommended in the methodological literature. In particular, the RMSEA has been recommended as an appropriate choice (e.g., Browne & Cudeck, 1993; MacCallum, 1995; Sugawara & MacCallum, 1993). RMSEA values below .10 are considered to represent a reasonable fit, and values below .05 to represent a very

good fit, with a value of .08 having been recommended as representing an appropriate cut-off point (Browne & Cudeck, 1993; Sugawara & MacCallum, 1993).

As with the Satorra-Bentler chi-square, the change in the values of the *CFI (Δ *CFI) and the RMSEA among nested models was also used in contrasting the fit of the models to the data. However, unlike the chi-square statistic, which can be evaluated for its significance, a clear standard as to what constitutes a statistically or practically important change has not been established for these indices. The change in the values of these indices are used, as they have been by others (e.g., Byrne & Shavelson, 1996), as an additional source of information in nested model comparisons. We turn now to the issues and information utilized in considering modifications to the hypothesized models.

Model Modification

Given that even true models can never completely represent sample data (Browne & Cudeck, 1993; MacCallum, 1995; MacCallum et al., 1992), initially postulated theoretical models will always include some degree of misfit. Under Hypotheses 1 and 2, consistent with model evaluation and modification procedures, detection of model misfit was based, at a general level, on the overall fit of the model (as outlined above), and at a more specific level, on the assessment of the parameter estimates and parameter misspecification (Byrne, 1994; Hoyle, 1995). Assessment of the appropriateness of the parameter estimates entailed an inspection of their values to ensure that they fell within an admissible range. Assessment of parameter misspecification involved a review of the results of the multivariate Lagrange-Multiplier Test (hereafter referred to as the LM-Test).

The LM-Test provides for a multivariate test of fixed parameters in the model which, if estimated in a subsequent test of the model, would improve its fit to the sample data; the parameters are listed such that the first parameter will provide the greatest enhancement in the fit of the model, with each successive parameter providing a lesser degree of improved fit. Each LM-Test parameter has an associated chi-square value which represents the expected drop in the model chi-square value should the specified parameter be included in the model. In tests of the hypothesized first-order models under Hypothesis 1, the presence

of misspecifications related to the multidimensionality of the SC construct were examined; thus, misspecifications representing correlated measurement errors and cross-loadings were estimated. In tests of the hypothesized second-order models under Hypothesis 2, misspecifications in the structure of SC were examined; accordingly, misspecifications due to the presence of correlated disturbance terms, and alternative relations among the first- and second-order factors were examined.

Decision-making with regard to whether a parameter(s) listed by the LM-Test was added to the models under Hypotheses 1 and 2 was based on a number of criteria. The first and most important criterion was the theoretical and substantive meaningfulness of the identified parameter(s) (Byrne, 1994, 1998; MacCallum, 1986; MacCallum et al., 1992). Only misspecifications that were consistent with previous empirical research, or had a solid basis in theory were considered for inclusion in post hoc models. Relatedly, the availability of alternative a priori models to address the nature of the model misfit was preferred over the utilization of post hoc procedures (e.g., Hoyle, 1995; MacCallum et al., 1993); the latter represents a movement from a hypothesis testing mode to an exploratory mode.

The relative values of the chi-square statistic associated with the LM-Test parameters was also considered in assessing model misfit. Specifically, a parameter(s) with an LM chi-square value of a magnitude that was distinct from that of the remaining chi-square values was considered more suggestive of model misfit than parameters with relatively non-distinct chi-square values (i.e., values that decreased in relatively small and consistent decrements). In addition, the inclusion of additional parameters had to be justified statistically. That is, the improvement in the fit of a model with the addition of one or several parameters had to be statistically significant, as reflected in a statistically significant difference in the chi-square value (relative to the change in the degrees of freedom) from one model to the next.

The decision to include post hoc parameters was also made within the context of the overall adequacy of the fit of the model. As the goodness-of-fit of the models increased, the inclusion of additional parameters required increased justification. Research has demonstrated that error in model modification (i.e., inaccuracy and poor likelihood of

replication) increases with the number of post hoc parameters, and that sources of misfit associated with smaller discrepancies in model fit are less stable across samples (e.g., MacCallum, 1986). Logically, the magnitude of model misfit tends to decrease as the fit of a model increases. Thus, the requirement that the addition of post hoc parameters be justified substantively and statistically was applied more stringently as the goodness-of-fit of the model increased. This minimized the possibility of overfitting the model to the data, and the risk of capitalizing on chance factors that were specific to the sample data.

A final aspect considered in evaluating the hypothesized models was the statistical significance of the parameter estimates. The EQS program provides a multivariate test of the significance of model parameters based on a .05 significance level; this test is called the multivariate Wald-Test (hereafter referred to as the Wald-Test). Given that second-order factors are specified to represent the relations among first-order factors, the Wald-Test parameters representing first-order factor covariances were not considered for exclusion in first-order models. This aspect of the SC construct was addressed in tests of the significance of the parameters in second-order models. In addition, when identified by the Wald-Test, measurement error and disturbance terms were retained as estimated parameters given the inappropriateness of their exclusion in model estimation (e.g., Bentler & Wu, 1995).

Model modification for the tests of invariance specified under Hypotheses 3 and 5 were based on the LM-Test. Specifically, in tests related to the invariance of a model across groups, the LM-Test identifies the significance level associated with each of the parameters constrained to be equal across groups; parameters identified as significant (i.e., not equal across groups) are not considered to be invariant. Accordingly, under Hypotheses 3 and 5, parameters identified as significant at the .05 alpha level were allowed to vary (i.e., were freely estimated) across groups in subsequent model estimation procedures. Using the baseline model of SC measurement and structure established under Hypotheses 1 and 2, Hypothesis 4 was evaluated by first estimating the latent SC/AA relations, and then examining the feasibility of the parameter estimates, and their consistency with the postulated multidimensional and hierarchical pattern of relations. Thus, evaluation of

Hypothesis 4 was not based on fit indices. Rather, the pattern of latent correlations among the SC and AA facets formed the data upon which this hypothesis was evaluated in that they were examined for their congruence with the patterns hypothesized.

RESULTS

For purposes of clarity, the results will be provided in two sections. The first section will focus on the preliminary analyses that were conducted in preparation for the main analyses. In particular, information relevant to data preparation, descriptive statistics, and instrument validity will be reviewed. In the second section, the results from the main analyses will be presented, with those relevant to the within-network component of SC presented first and followed by between-network analyses of SC/AA relations.

Preliminary Analyses

Data Preparation

The first stage of data preparation involved the elimination of cases through a visual inspection of the data in order to identify those cases for which all, or a large portion of the data for the grades or an instrument were not available. A total of 36 participants were excluded on the basis of these two criteria. Of these, grades were not available for 6 girls and 10 boys, and data were not available across all instruments for 20 participants. Of the 20 students for whom data were not complete, 3 girls and 5 boys had completed only one of the two testing sessions, and 4 girls and 8 boys either left class prior to the completion of testing, opted not to complete the instruments, or rendered their responses unscorable by endorsing both sides of the SPPC. Thus, a total of 13 girls and 23 boys were excluded from subsequent analyses; a review of the cases with missing data did not reveal a pattern, and thus, it was concluded that the data were missing completely at random (Muthén, Kaplan, & Hollis, 1987). The resulting sample consisted of 438 participants, of which 225 were girls and 213 were boys.

As described in the method section, where appropriate, composite AA indicator variables were computed. The data were then examined in order to eliminate those variables and cases having more than 5% missing data. With regard to the variables, AA measures in Measurement, Social Studies, and Religion were excluded from further analyses; these subject areas were, in some classes, not covered until later in the school year, and thus, the children had not received a grade when the instruments were administered. With regard to

the elimination of cases, 2 girls and 6 boys were identified and eliminated. The resulting sample size was 223 for girls and 207 for boys. Subsequent to reverse-scoring appropriate SC items, regression analyses were conducted for girls and boys separately in order to impute values for the remaining cases with missing data.

The final aspects of data preparation involved computing SC item-pairs, and the subsequent identification of multivariate outliers in relation to the main hypotheses of the present study. A total of 5 subjects, 3 girls and 2 boys, were identified as multivariate outliers and eliminated from subsequent analyses. Thus, the final samples utilized in testing the hypotheses included 220 girls and 205 boys, which represents 92.4 and 86.9% of the original sample, respectively, prior to the elimination of cases. If the total potential pool of participants across the three school boards are considered, a more conservative estimate of the subject response rate is 78.7%.

The mean age of the final sample, by grade and gender, is summarized in Table 3. As a review of this table indicates, the composition of the samples is again quite similar across groups. Although the greatest difference in the number of participants is reflected in the subgroup of Grade 6 boys, which is moderately underrepresented as compared to other groups, the mean ages are remarkably consistent when accounting for grade level.

Table 3

Final Sample Mean Age Within and Across Grade and Gender

Grade	Girls		Boys		Total Sample	
	<u>n</u>	<u>M</u>	<u>n</u>	<u>M</u>	<u>N</u>	<u>M</u>
Grade 5	108	10.26	107	10.27	215	10.27
Grade 6	112	11.22	98	11.30	210	11.26
Grades 5 & 6	220	10.75	205	10.76	425	10.75

Descriptive Statistics

Descriptive statistics were computed separately for girls and boys for each of the SC instruments scale scores, and for each of the SC and AA indicator variable scores. These statistics are presented in Tables 4 and 5 for the (A)SDQ-I and SPPC instrument scale scores, respectively. The descriptive statistics for the ASDQ-I, the SDQ-I, the SPPC academic, and the SPPC nonacademic indicator variable scores are presented in Tables 6 to 9, respectively. Those for the AA indicator variable scores are presented in Table 10. SC items were paired as represented in Tables 6 to 9. For ease of identification, the (A)SDQ-I items were represented with the numbers 1 through 6, and the SPPC items were represented with numbers greater than 6. The indicator variables were designated with their content area, both item numbers from which they were derived, and, if applicable, an “r” when they had been reverse-scored. Thus, for example, indicator variable “WLA8r9” represented the pairing of Item numbers 8 and 9 from the Writing in Language Arts scale, with Item 8 having been reverse-scored. For the AA constructs, grades represented one indicator variable and self-reported grades represented the other indicator variable.

The skewness and kurtosis values for the majority of the indicator variables and scales fell within acceptable ranges for normality. However, for the PRT25 indicator variable of the SDQ-I Parent SC scale, skewness and kurtosis values fall outside the range for the data to be considered normal, and are relatively high, particularly for the sample of boys. For example, kurtosis values for the SDQ-I indicator variable PRT25 are 5.047 and 6.671 for the girl and boy samples, respectively. Interestingly, the mean value for this indicator variable for both gender groups approaches the highest possible value of five, and is the highest of any of the (A)SDQ-I indicator variables for each gender group. In addition, the standard deviation value associated with this indicator variable is the lowest of any of the (A)SDQ-I indicator variables for each group. Thus, participants appear to have endorsed a uniformly high Parent SC for this indicator variable.

For the sample of boys, descriptive statistics for the Gym SC scales and indicator variables are relatively highly non-normal, particularly for the (A)SDQ-I instrument, as are

Table 4

ASDQ-I and SDQ-I Scale Score Mean, Standard Deviation, Skewness, and Kurtosis Values for the Sample of Girls (N=220) and of Boys (N=205)

SC scale	Girls				Boys			
	<u>M</u>	<u>SD</u>	Skewness	Kurtosis	<u>M</u>	<u>SD</u>	Skewness	Kurtosis
WLA	3.955	0.826	-0.438	-0.428	3.732	0.936	-0.616	-0.051
RLA	4.053	0.832	-0.646	0.069	3.834	0.911	-0.603	-0.288
LLA	3.975	0.755	-0.148	-1.016	3.755	0.885	-0.478	-0.273
SLA	3.876	0.843	-0.201	-0.890	3.723	0.879	-0.364	-0.415
MT	3.702	0.992	-0.505	-0.308	4.138	0.913	-1.152	0.851
ARM	3.963	0.958	-0.594	-0.538	4.296	0.845	-1.142	0.517
GA	4.096	0.717	-0.561	-0.111	4.068	0.871	-1.042	0.809
ART	4.284	0.721	-0.733	-0.486	4.013	0.967	-1.062	0.738
GYM	4.137	0.818	-1.032	1.015	4.467	0.762	-1.902	4.153
REL	4.137	0.811	-0.922	0.691	4.024	0.878	-0.860	0.274
SST	3.695	0.930	-0.286	-0.709	3.663	0.937	-0.508	-0.091
SN	3.801	0.869	-0.359	-0.565	3.824	0.957	-0.630	-0.216
PEER	3.929	0.749	-0.581	0.281	3.908	0.820	-0.912	0.928
PRT	4.144	0.826	-1.192	1.353	4.244	0.797	-1.409	2.136
AB	3.758	0.766	-0.489	-0.078	4.243	0.738	-1.425	2.244
AP	3.487	0.822	-0.317	-0.299	3.674	0.915	-0.600	-0.094

Note. ASDQ-I = Academic Self-Description Questionnaire-I (Marsh, 1990d); SDQ-I = Self-Description Questionnaire-I (Marsh, 1992b); SC = self-concept; W/R/L/SLA = Writing/Reading/Listening/Speaking in Language Arts, respectively; MT = Measurement; ARM = Arithmetic; GA = General-Academic; REL = Religion; SST = Social Studies; SN = Science; PRT = Parent; AB = Physical Ability; AP = Physical Appearance.

Table 5

SPPC Scale Score Mean, Standard Deviation, Skewness, and Kurtosis Values for the Sample of Girls (N=220) and of Boys (N=205)

SC scale	Girls				Boys			
	<u>M</u>	<u>SD</u>	Skewness	Kurtosis	<u>M</u>	<u>SD</u>	Skewness	Kurtosis
WLA	3.115	0.680	-0.408	-0.644	3.035	0.702	-0.532	-0.131
RLA	3.187	0.665	-0.577	-0.268	3.110	0.663	-0.513	0.078
LLA	3.194	0.616	-0.535	0.027	3.076	0.635	-0.525	0.091
SLA	3.075	0.668	-0.245	-0.722	2.955	0.666	-0.359	0.003
MT	2.994	0.775	-0.458	-0.511	3.349	0.689	-0.942	0.216
ARM	3.155	0.744	-0.613	-0.265	3.403	0.639	-0.885	0.073
GA	3.274	0.654	-0.621	-0.363	3.321	0.659	-0.935	0.434
ART	3.343	0.582	-0.595	-0.415	3.207	0.759	-0.948	0.364
GYM	3.265	0.674	-0.830	0.272	3.562	0.581	-1.561	2.556
REL	3.259	0.636	-0.676	0.128	3.285	0.655	-0.965	0.865
SST	2.890	0.736	-0.138	-0.790	2.962	0.741	-0.338	-0.538
SN	3.074	0.667	-0.449	-0.366	3.141	0.690	-0.492	-0.588
PEER	3.051	0.665	-0.706	-0.096	3.125	0.647	-0.819	0.256
AB	2.862	0.707	-0.251	-0.580	3.244	0.660	-0.774	-0.038
AP	2.720	0.666	-0.138	-0.462	3.144	0.641	-0.610	0.000

Note. SPPC = Self-Perception Profile for Children (SPPC; Harter, 1985); SC = self-concept; W/R/L/SLA = Writing/Reading/Listening/Speaking in Language Arts, respectively; MT = Measurement; ARM = Arithmetic; GA = General-Academic; REL = Religion; SST = Social Studies; SN = Science; AB = Physical Ability; AP = Physical Appearance.

Table 6

ASDQ-I Indicator Variable Score Mean, Standard Deviation, Skewness, and Kurtosis Values for the Sample of Girls (N=220) and of Boys (N=205)

SC indicator variable	Girls				Boys			
	M	SD	Skewness	Kurtosis	M	SD	Skewness	Kurtosis
WLA13	3.916	0.890	-0.434	-0.568	3.644	1.011	-0.406	-0.494
WLA2	3.959	1.053	-0.795	0.092	3.732	1.081	-0.672	0.025
WLA45	3.993	0.870	-0.535	-0.214	3.820	1.005	-0.831	0.390
RLA13	4.061	0.886	-0.599	-0.375	3.832	0.971	-0.623	-0.178
RLA2	3.991	1.007	-0.686	-0.190	3.766	1.054	-0.482	-0.463
RLA45	4.075	0.877	-0.745	0.150	3.870	0.980	-0.681	-0.316
LLA13	3.970	0.821	-0.364	-0.497	3.680	0.909	-0.410	-0.304
LLA2	4.000	0.927	-0.626	0.143	3.754	1.059	-0.727	0.042
LLA45	3.968	0.815	-0.076	-1.260	3.832	0.984	-0.601	-0.067
SLA13	3.881	0.916	-0.440	-0.593	3.738	0.894	-0.296	-0.702
SLA2	3.891	0.992	-0.346	-0.816	3.644	1.060	-0.521	-0.035
SLA45	3.863	0.897	-0.256	-0.821	3.749	0.963	-0.501	-0.198
MT13	3.779	1.033	-0.688	0.017	4.193	0.938	-1.228	0.995
MT2	3.677	1.098	-0.417	-0.522	4.137	1.048	-1.102	0.496
MT45	3.639	1.032	-0.444	-0.356	4.083	0.975	-1.040	0.509
ARM13	3.989	0.984	-0.687	-0.259	4.295	0.884	-1.258	1.013
ARM2	3.935	1.092	-0.612	-0.566	4.288	0.924	-1.170	0.739
ARM45	3.952	0.994	-0.529	-0.743	4.300	0.891	-1.199	0.603
GA13	4.109	0.817	-0.791	0.223	4.064	0.936	-0.971	0.510
GA2	4.273	0.859	-0.950	0.432	4.128	1.015	-1.225	1.246
GA45	3.995	0.820	-0.394	-0.454	4.041	0.932	-1.076	0.965
ART13	4.273	0.772	-0.814	-0.192	4.018	1.013	-0.998	0.431
ART2	4.254	0.877	-0.982	0.286	3.976	1.157	-0.989	0.211
ART45	4.310	0.787	-0.823	-0.607	4.026	1.004	-1.017	0.602
GYM13	4.015	0.927	-0.879	0.527	4.391	0.806	-1.564	2.647
GYM2	4.159	0.969	-1.116	0.967	4.502	0.861	-1.915	3.799
GYM45	4.248	0.865	-1.083	0.814	4.524	0.803	-2.084	4.665
REL13	4.166	0.839	-0.886	0.577	4.018	0.940	-0.808	0.165
REL2	4.040	1.004	-0.844	0.102	3.904	1.086	-0.829	0.101
REL45	4.158	0.865	-0.866	0.162	4.090	0.951	-1.064	0.642
SST13	3.793	0.957	-0.394	-0.733	3.730	0.985	-0.619	-0.012
SST2	3.630	1.086	-0.518	-0.279	3.611	1.129	-0.460	-0.495
SST45	3.630	1.000	-0.158	-0.905	3.621	1.007	-0.431	-0.283
SN13	3.918	0.912	-0.630	-0.055	3.905	1.006	-0.791	0.110
SN2	3.800	1.075	-0.594	-0.217	3.833	1.057	-0.668	-0.065
SN45	3.683	0.928	-0.260	-0.597	3.740	1.056	-0.593	-0.406

Note. ASDQ-I = Academic Self-Description Questionnaire-I (Marsh, 1990d); SC = self-concept; W/R/L/SLA =

Writing/Reading/Listening/Speaking in Language Arts, respectively; MT = Measurement; ARM = Arithmetic; GA =

General-Academic; REL = Religion; SST = Social Studies; SN = Science. The numbers represent the single- or dual-items from which the indicator variables were drawn.

Table 7

SDQ-I Indicator Variable Score Mean, Standard Deviation, Skewness, and Kurtosis Values for the Sample of Girls (N=220) and of Boys (N=205)

SC indicator variable	Girls				Boys			
	<u>M</u>	<u>SD</u>	Skewness	Kurtosis	<u>M</u>	<u>SD</u>	Skewness	Kurtosis
PEER12	3.973	0.854	-0.636	0.117	3.940	0.886	-0.795	0.549
PEER36	4.027	0.822	-0.624	0.028	4.007	0.906	-0.982	0.903
PEER45	3.788	0.906	-0.442	-0.327	3.776	1.019	-0.742	0.084
PRT14	3.897	1.031	-0.739	-0.273	4.118	0.984	-1.131	0.571
PRT25	4.518	0.716	-1.991	5.047	4.535	0.744	-2.353	6.671
PRT36	4.018	1.065	-1.049	0.456	4.078	1.003	-1.136	0.910
AB14	3.672	0.974	-0.450	-0.333	3.920	1.056	-0.893	0.088
AB26	3.689	0.844	-0.288	-0.645	4.346	0.773	-1.498	2.132
AB35	3.914	1.003	-0.796	0.110	4.464	0.856	-1.757	2.552
AP15	3.464	0.890	-0.287	-0.059	3.732	1.040	-0.782	0.116
AP23	3.781	0.969	-0.630	-0.009	3.991	0.898	-0.859	0.563
AP46	3.215	0.993	-0.185	-0.374	3.299	1.057	-0.281	-0.572

Note. SDQ-I = Self-Description Questionnaire-I (Marsh, 1992b); SC = self-concept; PRT = Parent; AB = Physical Ability; AP = Physical Appearance. The numbers correspond to the two SDQ-I item numbers from which the indicator variables were derived.

Table 8

SPPC Academic Self-concept (SC) Indicator Variable Score Mean, Standard Deviation, Skewness, and Kurtosis Values for the Sample of Girls (N=220) and of Boys (N=205)

SC indicator variable	Girls				Boys			
	<u>M</u>	<u>SD</u>	Skewness	Kurtosis	<u>M</u>	<u>SD</u>	Skewness	Kurtosis
WLA7r10r	3.128	0.747	-0.604	-0.282	3.055	0.742	-0.682	0.197
WLA8r9	3.102	0.724	-0.280	-0.984	3.015	0.771	-0.460	-0.394
RLA7r10r	3.232	0.747	-0.663	-0.357	3.147	0.714	-0.643	0.008
RLA8r9	3.142	0.708	-0.506	-0.392	3.073	0.727	-0.499	-0.097
LLA7r10r	3.218	0.672	-0.665	0.129	3.072	0.681	-0.565	0.235
LLA8r9	3.170	0.674	-0.347	-0.599	3.080	0.721	-0.404	-0.301
SLA7r10r	3.105	0.712	-0.441	-0.410	3.002	0.734	-0.564	-0.001
SLA8r9	3.046	0.729	-0.194	-0.815	2.909	0.737	-0.277	-0.355
MT7r10r	2.968	0.824	-0.533	-0.398	3.346	0.742	-1.065	0.428
MT8r9	3.020	0.806	-0.490	-0.444	3.351	0.730	-0.960	0.189
ARM7r10r	3.243	0.767	-0.894	0.378	3.474	0.696	-1.239	0.846
ARM8r9	3.068	0.818	-0.510	-0.563	3.332	0.705	-0.666	-0.638
GA7r10r	3.362	0.687	-0.916	0.147	3.395	0.703	-1.125	0.918
GA8r9	3.186	0.735	-0.656	-0.294	3.247	0.743	-0.846	0.083
ART7r10r	3.407	0.642	-0.885	0.082	3.232	0.802	-1.045	0.640
ART8r9	3.278	0.663	-0.446	-0.914	3.182	0.834	-0.773	-0.336
GYM7r10r	3.268	0.745	-0.847	0.085	3.607	0.629	-1.788	3.006
GYM8r9	3.261	0.720	-0.710	-0.316	3.517	0.646	-1.288	1.202
REL7r10r	3.285	0.679	-0.754	0.286	3.327	0.709	-1.003	0.662
REL8r9	3.233	0.709	-0.710	-0.115	3.243	0.726	-0.728	-0.212
SST7r10r	2.804	0.819	-0.089	-0.845	2.866	0.804	-0.283	-0.684
SST8r9	2.977	0.734	-0.255	-0.643	3.057	0.782	-0.528	-0.356
SN7r10r	3.093	0.721	-0.626	0.266	3.190	0.767	-0.775	0.027
SN8r9	3.055	0.724	-0.547	-0.302	3.093	0.755	-0.483	-0.536

Note. SPPC = Self-Perception Profile for Children (SPPC; Harter, 1985); W/R/L/SLA =

Writing/Reading/Listening/Speaking in Language Arts, respectively; MT = Measurement;

ARM = Arithmetic; GA = General-Academic; REL = Religion; SST = Social Studies; SN =

Science. The numbers correspond to the two item numbers from which the indicator variables were derived, with an "r" representing a reverse-scored item.

Table 9

SPPC Nonacademic Self-concept (SC) Indicator Variable Score Mean, Standard Deviation, Skewness, and Kurtosis Values for the Sample of Girls (N=220) and of Boys (N=205)

SC indicator variable	Girls				Boys			
	<u>M</u>	<u>SD</u>	Skewness	Kurtosis	<u>M</u>	<u>SD</u>	Skewness	Kurtosis
PEER8r9	3.156	0.845	-0.828	-0.231	3.205	0.804	-0.832	-0.078
PEER10r7	3.067	0.742	-0.653	-0.149	3.176	0.704	-0.782	0.330
PEER12r11	2.929	0.818	-0.492	-0.373	2.994	0.822	-0.478	-0.632
AB7r11	3.059	0.800	-0.519	-0.541	3.454	0.696	-1.110	0.130
AB9r12	2.850	0.780	-0.193	-0.611	3.217	0.775	-0.673	-0.543
AB10r8	2.678	0.888	-0.161	-1.037	3.061	0.881	-0.626	-0.551
AP7r10	2.810	0.815	-0.334	-0.546	3.181	0.759	-0.633	-0.372
AP8r9	2.672	0.895	-0.198	-0.781	3.185	0.834	-0.825	-0.161
AP12r11	2.677	0.819	-0.091	-0.695	3.067	0.774	-0.521	-0.538

Note. SPPC = Self-Perception Profile for Children (Harter, 1985); AB = Physical Ability; AP = Physical Appearance. The numbers correspond to the two item numbers from which the indicator variables were derived, with an “r” representing a reverse-scored item.

Table 10

Academic Achievement (AA) Indicator Variable Score Mean, Standard Deviation, Skewness, and Kurtosis Values for the Sample of Girls (N=220) and of Boys (N=205)

AA indicator variable	Girls				Boys			
	<u>M</u>	<u>SD</u>	Skewness	Kurtosis	<u>M</u>	<u>SD</u>	Skewness	Kurtosis
WLA-GR	2.122	0.796	0.179	-0.768	2.431	0.836	0.059	-0.472
WLA-RGR	1.861	0.813	0.773	0.459	2.185	0.795	0.128	-0.574
RLA-GR	2.065	0.786	0.217	-0.725	2.220	0.784	-0.075	-0.808
RLA-RGR	1.836	0.777	0.588	-0.259	2.070	0.777	0.065	-0.902
LLA-GR	2.053	0.844	0.145	-1.073	2.239	0.865	-0.033	-0.711
LLA-RGR	1.756	0.767	0.629	-0.452	2.088	0.853	0.309	-0.436
SLA-GR	2.061	0.829	0.167	-0.944	2.230	0.753	-0.272	-0.900
SLA-RGR	1.918	0.795	0.477	-0.139	2.238	0.833	0.199	-0.547
ARM-GR	2.018	0.891	0.550	-0.277	2.048	0.908	0.347	-0.899
ARM-RGR	1.777	0.794	0.752	-0.065	1.722	0.832	0.870	-0.147
SN-GR	2.050	0.797	0.542	0.040	2.228	0.801	0.278	-0.272
SN-RGR	1.920	0.856	0.513	-0.646	2.063	0.841	0.529	0.285
ART-GR	1.945	0.634	0.156	-0.584	2.222	0.596	-0.172	-0.334
ART-RGR	1.562	0.641	0.922	0.761	1.854	0.766	0.586	-0.113
GYM-GR	2.178	0.661	-0.045	-0.608	2.054	0.715	0.212	-0.809
GYM-RGR	1.789	0.806	0.776	-0.023	1.578	0.727	1.001	0.167

Note. GR = grade; RGR = self-report of grade; W/R/L/SLA = Writing/Reading/Listening/
Speaking in Language Arts, respectively; ARM = Arithmetic; SN = Science.

those for the (A)SDQ-I Physical Ability scale and indicator variables. Slight non-normality of the data is also evident for these scales for the sample of girls, and for the (A)SDQ-I Measurement, Arithmetic, and to a lesser extent, General-Academic scales for the sample of boys. Thus, although the descriptive statistics suggest that the majority of the variables are normally distributed, there is clear evidence of non-normality for the Parent SC scale for both girls and boys, as well as for the (A)SDQ-I Physical Ability and Gym SC scales, and to a lesser extent, for the SPPC Gym SC scale for the sample of boys. The distributional non-normality of the Physical Ability and Gym SC scales and indicator variables for the sample of boys is more evident for the (A)SDQ-I than for the SPPC instruments. In addition, examination of the descriptive statistics for the (A)SDQ-I Gym and Physical Ability scales and indicator variables for the sample of boys reveals that the mean, standard deviation, skewness, and kurtosis values are more consistent across the Gym than across the Physical Ability SC indicator variables. Given the presence of non-normality for parts of the data for both girls and boys, all analyses of covariance structures were, whenever possible, based on the Satorra-Bentler chi-square test statistic and robust estimates. As noted earlier, this test statistic and the robust estimates provide for a correction to their usual values when the assumption of multivariate normality does not hold across variables. We turn now to a review of analyses relevant to instrument validity.

Instrument Validity: The Self-Description Questionnaires and
the Self-Perception Profile for Children

Cronbach Alpha Coefficients: Self-concept Instruments

The Cronbach alpha coefficients for each of the (A)SDQ-I and SPPC instrument scale scores are presented separately for girls and boys in Table 11. The coefficients provide support for the internal consistency of each of the (A)SDQ-I SC scales for both the sample of girls (i.e., coefficients range from .74 to .94, $\underline{M} = .88$) and the sample of boys (i.e., coefficients range from .76 to .93, $\underline{M} = .89$); only the Physical Ability scale coefficients fall below .80 (i.e., $\alpha = .76$ for girls and .74 for boys). The coefficients for the SPPC scales also fall within acceptable ranges for both girls (i.e., $\alpha = .71$ to .89, $\underline{M} = .82$) and boys (i.e., $\alpha =$

Table 11

Cronbach Alpha Coefficients for Girls (N=220) and for Boys (N=205) for Each of the (A)SDQ-I and SPPC Scale Scores, and the Mean Across Scales for Each Instrument

SC scale	Cronbach alpha coefficients			
	(A)SDQ-I		SPPC	
	Girls	Boys	Girls	Boys
WLA	0.882	0.895	0.830	0.839
RLA	0.902	0.903	0.806	0.817
LLA	0.876	0.898	0.807	0.781
SLA	0.898	0.908	0.837	0.782
MT	0.940	0.922	0.893	0.858
ARM	0.944	0.932	0.865	0.798
GA	0.828	0.901	0.815	0.795
ART	0.874	0.917	0.742	0.837
GYM	0.882	0.919	0.817	0.795
REL	0.900	0.874	0.809	0.798
SST	0.917	0.900	0.887	0.854
SN	0.893	0.906	0.825	0.782
PEER	0.841	0.846	0.771	0.777
PRT	0.853	0.848	N/A	N/A
AB	0.742	0.761	0.821	0.790
AP	0.833	0.905	0.705	0.744
Mean	0.875	0.890	0.815	0.803

Note. (A)SDQ-I = Academic Self-Description Questionnaire-I (Marsh, 1990d) and Self-Description Questionnaire-I (Marsh, 1992b); SPPC = Self-Perception Profile for Children (Harter, 1985); SC = self-concept; W/R/L/SLA = Writing/Reading/Listening/Speaking in Language Arts, respectively; MT = Measurement; ARM = Arithmetic; GA = General-Academic; REL = Religion; SST = Social Studies; SN = Science; PRT = Parent; AB = Physical Ability; AP = Physical Appearance; N/A = not applicable.

.74 to .86, $M = .80$). However, for both samples, the coefficients are less favorable for the SPPC than for the (A)SDQ-I instruments, particularly for boys. For example, coefficients fall below .80 for 10 of the 15 SPPC scales for boys, whereas this is true of only 3 scales for girls. Thus, the internal consistency coefficients for the (A)SDQ-I instruments are somewhat higher than for the SPPC for both girls and boys, with the lowest scale consistencies obtained with the SPPC instrument for the sample of boys.

Confirmatory Factor Analyses: Self-concept Instruments

In order to examine the validity of the multiple scales of the (A)SDQ-I and SPPC instruments, CFAs were conducted, separately for girls and boys, on: (a) the (A)SDQ-I, (b) the SPPC, and (c) both the (A)SDQ-I and SPPC instruments combined. Analyses of the (A)SDQ-I instruments were based on each of the 16 SC factors these instruments were designed to measure in the present study. Analyses of the SPPC were based on a 15-factor model; this instrument was not designed with a Parent scale, and thus, one less factor is specified with this instrument than with the (A)SDQ-I.

CFA of the 16 scales of the (A)SDQ-I instruments provides an excellent fit to the data for both the sample of girls ($S-\beta \chi^2_{(960)} = 1249.56$, $*CFI = .962$, $RMSEA = .053$), and the sample of boys ($S-\beta \chi^2_{(960)} = 1193.29$, $*CFI = .954$, $RMSEA = .056$). For the sample of girls, the LM-Test results do not reveal any distinctly ill-fitting parameters, whereas the results for the sample of boys reveal three parameters with chi-square values that are, at most, minimally distinct from those of the remaining parameters. Given the excellent fit of the 16-factor (A)SDQ-I model, and the lack of strong distinction among the chi-square values associated with the LM-Test parameters for both girls and boys, results are considered to provide support for a 16-factor representation of this instrument.

CFA of the 15 SPPC scales, as with the results for the (A)SDQ-I instrument, provides an excellent fit to the data for both girls ($S-\beta \chi^2_{(390)} = 583.83$, $*CFI = .949$, $RMSEA = .063$; $*ECVI = 4.23$) and boys ($S-\beta \chi^2_{(390)} = 520.02$, $*CFI = .957$, $RMSEA = .057$; $*ECVI = 4.23$). The LM-Test results for girls do not reveal any distinctly ill-fitting parameters, although, for boys, four parameters exhibit chi-square values that are, again, at

most, minimally distinct from those of the remaining parameters. Given the excellent fit of the 15-factor SPPC model to the data and the lack of any distinct model misfit for both girls and boys, the hypothesized 15-factor model is considered to represent an accurate depiction of the SPPC instrument. In conclusion, single-instrument CFAs of both the SPPC and the (A)SDQ-I instruments provide support for the validity of these instruments as measures of preadolescent multidimensional SC for both girls and boys.

CFAs of a 16-factor model of SC with indicator variables from both the (A)SDQ-I and the SPPC instruments, however, do not provide an adequate fit to the data for either the sample of girls ($S-\beta \chi^2_{(3039)} = 5071.32$, *CFI = .854; *ECVI = 25.73), or the sample of boys ($S-\beta \chi^2_{(3039)} = 4945.93$, *CFI = .813; *ECVI = 27.01)¹. For both samples, the robust CFI values fall below the .90 cut-off point for acceptable fit, particularly for boys. The difference in the adequacy of the fit of the single-, as compared to the dual-instrument SC models suggests the possibility that a method effect is occurring for both samples. Although the chi-square values associated with the LM-Test parameters begin at a moderately high value (LM $\chi^2 = 62.70$ for girls, and LM $\chi^2 = 53.63$ for boys), their relative values are not suggestive of any distinctly ill-fitting parameters.

However, examination of the specific parameters identified does reveal a common thread for both samples. In particular, the LM-Test parameters primarily represent correlated measurement error among the SPPC indicator variables, particularly the “8r9” academic indicator variable pairs. For example, for the sample of girls, of the first 25 variables identified by the LM-Test, 23 involve correlated errors among the SPPC indicator variables. A similar pattern is evident among the next set of 25 LM-Test parameters. Likewise, for the sample of boys, correlated errors among the SPPC indicator variables were also identified by the LM-Test as the source of model misfit. In fact, the first 110 parameters represent correlated error among the SPPC indicator variables, the majority of which are among SPPC academic indicator variables.

¹ Computational limitations precluded the estimation of the RMSEA index for combined SC instrument analyses.

The inadequate fit and the pattern of LM-Test results of the dual-instrument 16-factor SC model, coupled with the adequate fit of the single-instrument SC models for both girls and boys suggest the presence of a method effect specific to the SPPC instrument. In addition, both the fit indices and the LM-Test results are consistent in suggesting the possibility that this method effect is stronger for boys than for girls. That is, the fit indices are more favorable for the sample of girls than for the sample of boys, and the LM-Test results for boys are more exclusively suggestive of a method effect than are the results for girls. These findings are also consistent with the generally lower internal consistencies obtained for the SPPC than for the (A)SDQ-I instruments, as well as with the lower internal consistencies obtained for boys than for girls with the SPPC instrument scales. Thus, inclusion of the SPPC in tests of the hypotheses of the present study would potentially introduce a confound that would render the interpretability of the findings in relation to the SC construct, and in particular, gender differences in SC dubious. That is, findings related to the SPPC method effect would be confounded with those related to the SC construct itself, particularly in relation to gender differences. Accordingly, exclusion of the SPPC from investigations of the SC construct (i.e., tests of hypotheses) was deemed necessary; thus, all further analyses were conducted with only the (A)SDQ-I instruments.

Self-concept Measurement and Structure: Within-network Analyses

As specified under Hypothesis 1 in the method section, three first-order SC models were tested separately for girls and boys for their goodness-of-fit to the sample data. As previously noted, these SC models differ with regard to the extent to which the specific mathematics (i.e., Arithmetic and Measurement) or the specific Language Arts (i.e., Reading, Writing, Listening, and Speaking) SCs are represented as single or multiple dimensions. In particular, the three first-order SC models (see Figures 4 to 6 in the method section) include (a) the hypothesized 16-factor model of SC in which both the specific mathematics and the specific Language Arts SCs are postulated to form separate dimensions; (b) a 15-factor model in which the two mathematics SC dimensions are postulated to form one, rather than two SC constructs; and (c) a 13-factor SC model in which the four Language

Arts dimensions are postulated to form one, rather than four SC constructs. Note that, due to the exclusion of the SPPC instrument from analyses, the 16-factor model is identical to that tested in relation to the validity of the (A)SDQ-I instrument. A summary of the indices of fit for the 16-, 15-, and 13-factor first-order models of SC is presented in Table 12 for girls, and in Table 13 for boys.

Preadolescent Girls: Testing for the Multidimensional Structure of Self-concept

Recall that the hypothesized 16-factor model of SC (designated Model 16G for girls) provides an excellent fit to the sample data for girls ($S-\beta \chi^2_{(960)} = 1249.56$, $*CFI = .962$, $RMSEA = .053$; $*ECVI = 7.68$), and that no single parameter, or group of parameters was identified by the LM-Test as having a degree of misfit that is distinct from that of the remaining parameters. Accordingly, this model is considered to represent the final best-fitting 16-factor model of SC for preadolescent girls. Model 16G explains 96.2% of the variance in the data for the sample of girls. The Wald-Test results for Model 16G identifies 23 nonsignificant parameters, all of which represent factor covariances. Given that second-order factors are postulated to represent the relations among the first-order factors, this aspect of the SC construct will be addressed when hypotheses relevant to the hierarchical structure are examined. We now turn to a review of results related to the 15-factor model with the sample of girls.

The 15-factor model of SC (Model 15G for girls) yields a substantially poorer fit to the data ($S-\beta \chi^2_{(975)} = 1516.40$, $*CFI = .928$, $RMSEA = .065$; $*ECVI = 8.76$) than does Model 16G ($S-\beta \chi^2_{(960)} = 1249.56$, $*CFI = .962$, $RMSEA = .053$; $*ECVI = 7.68$). Furthermore, the LM-Test results reveal correlated measurement errors with associated χ^2 values that are substantial, and distinct from the remaining parameters, among the three indicator variables for the Measurement factor ($LM \chi^2 = 103.00$ to 156.44). It is noteworthy that the correlated errors identified are among indicator variables representing one of the collapsed factors. Nevertheless, the presence of measurement error covariances within a single measuring instrument is neither uncommon, nor unexpected in the measurement of psychological phenomena (see, e.g., Byrne, 1988, 1994; Tanaka & Huba, 1984). Indeed,

Table 12
Summary of the Goodness-of-fit Indices for the Sample of Girls: First-order 16-, 15-, and 13-factor Self-concept (SC) Models

Model	S- $\beta\chi^2$	df	$\Delta S-\beta\chi^2$ (Δdf) from matching a priori model	*CFI	RMSEA	*ECVI
16G: 16-factor SC model	1249.56*	960		.962	.053	7.68
15G: 15-factor SC model	1516.40*	975		.928	.065	8.76
15Ga: 15-factor SC model with correlated measurement error of MT13, MT2; MT13, MT45; and MT2, MT45	1284.58*	972	231.83 (3)**	.959	.054	7.73
13G: 13-factor SC model	1454.95*	1002		.940	.061	8.23
13Ga: 13-factor SC model with correlated measurement error of SLA13 and SLA2	1411.51*	1001	43.44 (1)**	.946	.059	8.04

Note. S- $\beta\chi^2$ = Satorra-Bentler chi-square; *CFI = robust Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; *ECVI = robust Expected Cross-Validation Index; MT = Measurement; SLA = Speaking in Language Arts.

* $p \leq .00000$. ** $p \leq .001$.

Table 13
 Summary of the Goodness-of-fit Indices for the Sample of Boys: First-order 16-, 15-, and 13-factor Self-concept (SC) Models

Model	S- $\beta\chi^2$	df	$\Delta S-\beta\chi^2$ (Δdf) from matching a priori model	*CFI	RMSEA	*ECVI
16B: 16-factor SC model	1193.29*	960		.954	.056	7.97
15B: 15-factor SC model	1430.50*	975		.910	.067	8.98
15Ba: 15-factor SC model with correlated measurement error of MT13, MT2;						
MT13, MT45; MT2, MT45	1212.94*	972	217.56 (3)**	.952	.056	7.95
13B: 13-factor SC model	1348.39*	1002		.931	.062	8.32
13Ba: 13-factor SC model with correlated measurement error of SLA13 and SLA2	1305.91*	1001	42.48 (1)**	.940	.060	8.12

Note. S- $\beta\chi^2$ = Satorra-Bentler chi-square; *CFI = robust Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; *ECVI = robust Expected Cross-Validation Index; MT = Measurement; SLA = Speaking in Language Arts.
 p=.00000. ** p \leq .001.

error covariances between items from the same scale of an instrument are particularly common given their shared content. Given the theoretical feasibility of these correlated measurement errors, these parameters were incorporated in a subsequent 15-factor post hoc model, Model 15Ga.

Model 15Ga ($S-\beta \chi^2_{(972)} = 1284.58$, $*CFI = .959$, $RMSEA = .054$; $*ECVI = 7.73$) provides a statistically significant improvement in model fit as compared to Model 15G ($\Delta S-\beta \chi^2_{(3)} = 231.83$, $p \leq .001$, $\Delta *CFI = .031$). The fit of Model 15Ga, however, is poorer than that of Model 16G ($S-\beta \chi^2_{(960)} = 1249.56$, $*CFI = .962$, $RMSEA = .053$; $*ECVI = 7.68$). Given the poorer fit of the 15-factor model relative to the 16-factor model, and the reduction in parsimony inherent in the addition of correlated measurement errors, the 15-factor model will not be evaluated further and is rejected in favor of a 16-factor representation of the data for girls. We turn now to the results for the 13-factor SC model with the sample of girls.

The 13-factor model of SC (Model 13G for girls) provides a relatively good fit to the data for the sample of girls ($S-\beta \chi^2_{(1002)} = 1454.95$, $*CFI = .940$, $RMSEA = .061$; $*ECVI = 8.23$). However, the fit of Model 13G is poorer than that of Model 16G ($S-\beta \chi^2_{(960)} = 1249.56$, $*CFI = .962$, $RMSEA = .053$; $*ECVI = 7.68$). The LM-Test results reveal a correlated error, with a χ^2 value distinct from that of the remaining LM-Test parameters, between indicator variable SLA13 and SLA2 ($LM \chi^2 = 55.22$); again, these parameters represent one of the collapsed factors. The presence of correlated measurement error between these indicator variables is theoretically feasible given their shared content area, and thus, this parameter was included in post hoc Model 13Ga. However, Model 13Ga ($S-\beta \chi^2_{(1001)} = 1411.51$, $*CFI = .946$, $RMSEA = .059$; $*ECVI = 8.04$) nevertheless yields a poorer fit than is provided by Model 16G ($S-\beta \chi^2_{(960)} = 1249.56$, $*CFI = .962$, $RMSEA = .053$; $*ECVI = 7.68$). Furthermore, the LM-Test results for Model 13Ga identifies correlated errors, with χ^2 values that are moderately distinct from the remaining parameters, among two sets of Language Arts indicator variables (i.e., LLA13 with LLA2, and RLA13 with RLA2; $LM \chi^2 = 32.75$ and 26.55 , respectively). Given these results, the 13-factor model will not be

evaluated further and is rejected in favor of the 16-factor model. Thus, Model 16G is considered to best represent SC structure for preadolescent girls.

The magnitude of the correlations among the mathematics and Language Arts SC facets in Model 16G, however, are worthy of note (see Appendix D for the latent factor correlations among the first-order SC constructs obtained under Model 16G). That is, Arithmetic SC is moderately correlated with the four specific Language Arts SC facets ($r_s = .34$ to $.46$, mean $r = .42$), as is Measurement SC ($r_s = .47$ to $.51$, mean $r = .48$). We turn now to a review of the results for the sample of boys.

Preadolescent Boys: Testing for the Multidimensional Structure of Self-concept

Recall that the hypothesized 16-factor model of SC (Model 16B for boys), as with the same model for girls, provides an excellent fit to the data for the sample of boys ($S\text{-}\beta \chi^2_{(960)} = 1193.29$, $*CFI = .954$, $RMSEA = .056$; $*ECVI = 7.97$), and that the chi-square values associated with the first three parameters identified by the LM-Test are perhaps slightly distinct from the remaining parameters. Given the high degree of fit of the model to the data, and the lack of clear distinction among the LM-Test parameters, Model 16B, which explains 95.4% of the variance in the data, is considered to represent the final best-fitting 16-factor model of SC for preadolescent boys. The Wald-Test results from Model 16B identifies 51 parameters as nonsignificant to the model. These parameters represent factor covariances and, as with the sample of girls, will be evaluated when hypotheses relevant to the hierarchical structure are examined. We now turn to a review of results related to the 15-factor model with the sample of boys.

The 15-factor model of SC (designated Model 15B with the sample of boys), consistent with the pattern of results for girls, yields a substantially poorer fit to the data for the sample of boys ($S\text{-}\beta \chi^2_{(975)} = 1430.50$, $*CFI = .910$, $RMSEA = .067$; $*ECVI = 8.98$) than does Model 16B ($S\text{-}\beta \chi^2_{(960)} = 1193.29$, $*CFI = .954$, $RMSEA = .056$; $*ECVI = 7.97$). Also consistent with the results for girls, the LM-Test for Model 15B identifies correlated measurement errors with substantial χ^2 values among the three indicator variables for the Measurement factor ($LM \chi^2 = 105.06$ to 112.66). Inclusion of these correlated measurement

errors in a subsequent model, Model 15Ba ($S-\beta \chi^2_{(972)} = 1212.94$, *CFI = .952, RMSEA = .056; *ECVI = 7.95), results in a statistically significant improvement in model fit over Model 15B ($\Delta S-\beta \chi^2_{(3)} = 217.56$, $p \leq .001$, Δ *CFI = .042). However, the fit of Model 15Ba is not different from that of Model 16B ($S-\beta \chi^2_{(960)} = 1193.29$, *CFI = .954, RMSEA = .056; *ECVI = 7.97). Given that Model 15Ba incorporates correlated errors, is an a posteriori model, and does not provide a better fit than does Model 16B, Model 16B is considered a superior model. Accordingly, as with the sample of girls, the 15-factor model of SC is rejected in favor of the 16-factor representation of SC with the sample of boys.

The 13-factor model of SC (designated Model 13B for boys), as with the results for the sample of girls, yields a poorer fit to the data for boys ($S-\beta \chi^2_{(1002)} = 1348.39$, *CFI = .931, RMSEA = .062; *ECVI = 8.32) than does Model 16B ($S-\beta \chi^2_{(960)} = 1193.29$, *CFI = .954, RMSEA = .056; *ECVI = 7.97). The LM-Test results identify the same parameter identified with this model for girls, namely, a correlated measurement error, with a χ^2 value distinct from those associated with the other parameters, between indicator variable SLA13 and SLA2 ($LM \chi^2 = 51.01$). The inclusion of this correlated measurement error in a subsequent post hoc model, designated Model 13Ba, improves the fit of the model to the data ($S-\beta \chi^2_{(1001)} = 1305.91$, *CFI = .940, RMSEA = .060; *ECVI = 8.12) relative to Model 13B ($\Delta S-\beta \chi^2_{(1)} = 42.48$, $p \leq .001$, Δ *CFI = .009). However, the fit of Model 13Ba remains poorer than that of Model 16B ($S-\beta \chi^2_{(960)} = 1193.29$, *CFI = .954, RMSEA = .056; *ECVI = 7.97). Accordingly, as with the corresponding models for girls, Model 13Ba is rejected in favor of Model 16B for the sample of boys.

Consistent with the latent factor correlations observed with Model 16G for girls, the Mathematics and Language Arts SC facets are moderately correlated in Model 16B (see Appendix D for the latent factor correlations among the first-order SCs obtained for Model 16B). That is, Arithmetic SC is moderately correlated with the four specific Language Arts SC facets ($r_s = .42$ to $.47$, mean $r = .46$), as is Measurement SC ($r_s = .38$ to $.43$, mean $r = .41$).

Testing for the Multidimensional Structure of Preadolescent Self-concept for Girls and Boys: A Summary

For both the sample of girls and the sample of boys, the 13- and 15-factor models of SC were rejected in favor of the hypothesized 16-factor model of preadolescent SC structure. Statistically, the 16-factor model of SC allows for a greater fit of the model to the sample data than does either the 13- or the 15-factor models for both girls and boys. Likewise, the practical fit indices suggest that the fit of the 16-factor model is superior to that of the 13- and 15-factor models. The identification of correlated errors among the indicator variables representing the collapsed factors with the 13- and 15-factor models for both girls and boys gives further credence to the theoretical appropriateness of the differentiation hypothesized in the 16-factor model. In addition, representing the variance in the data in terms of latent constructs is theoretically superior to doing so with correlated errors. Thus, covariance structure analyses are consistent with the statistical, practical and theoretical superiority of a 16-factor representation of the within-network structure of SC for both preadolescent girls and boys.

The final 16-factor models, namely Model 16G and 16B, both provide an excellent fit to the data and represent the baseline 16-factor multidimensional models of preadolescent SC upon which subsequent analyses were based for girls and boys, respectively. A pictorial representation of the core academic portion of these first-order baseline models was previously presented in Figure 4. We turn now to a review of analyses related to Hypothesis 2, which argued that SC is better described within the framework of an hierarchical model.

Testing for the Hierarchical Structure of Self-concept: An Overview

As specified under Hypothesis 2, CFAs examining a higher-order (i.e., hierarchical) SC structure were conducted separately for each gender using the best-fitting baseline models determined by the first-order analyses. As noted previously, a total of four hierarchical models, representing an expansion and adaptation of the Marsh/Shavelson model of SC (Marsh & Shavelson, 1985) are hypothesized; the four models were represented pictorially in Figures 7 to 10 (see method section). A summary of the goodness-of-fit

statistics related to each of the hierarchical models is presented in Table 14 for the sample of girls, and in Table 15 for the sample of boys.

Preadolescent Girls: Testing for the Hierarchical Structure of Self-concept

Recall that the first hierarchical model of SC postulated (Model H1; see Figure 7) is, with two exceptions, an exact duplication of the Marsh/Shavelson model of SC. Specifically, the second-order Mathematics-Academic SC facet was relabeled a Science-Academic SC facet, and an additional second-order factor was hypothesized to represent relations among the non-core first-order academic SC facets measured, namely, Religion, Art, and Gym (i.e., Physical Education). Thus, in Model H1, four second-order factors are hypothesized to represent the relations among the first-order constructs; specifically, second-order Nonacademic, Science-Academic, Verbal-Academic, and Non-Core Academic SCs.

Covariance structure analysis of Model H1 for the sample of girls (designated Model H1G) indicates that this model provides a well-fitting representation of the relations among the first-order SC facets ($S-\beta \chi^2_{(1054)} = 1551.68$, *CFI = .934, RMSEA = .062; *ECVI = 8.20). The LM-Test identifies two parameters exhibiting chi-square values distinct from those associated with the remaining parameters. These parameters represent correlated disturbance terms between the Gym and Physical Ability SC facets (LM $\chi^2 = 82.70$), and between the Arithmetic and Measurement SC facets (LM $\chi^2 = 63.71$). In the analysis of higher-order covariance structures, disturbance terms represent residual error associated with the prediction of the first-order factors by the higher-order factors. Thus, the correlation between the disturbance terms for the Gym and Physical Ability SC facets indicate that the residual error associated with the prediction of the Gym SC facet is correlated with that of the Physical Ability SC facet. Similarly, the second identified parameter represents correlated residual error between the Arithmetic and Measurement SC facets.

The alternate hierarchical models hypothesized in the present study offer the potential to address the areas of model misfit represented in the correlated residual errors identified under Model H1G. Specifically, Model H2 (see Figure 8) allows for the hierarchical representation of the Physical Ability and Gym SC facets under the rubric of the

Table 14
Hierarchical Models: Summary of the Goodness-of-fit Indices for the Sample of Girls (N = 220)

Model	S- $\beta\chi^2$	df	$\Delta S-\beta\chi^2$ (Δdf) from corresponding a priori model	*CFI	RMSEA	*ECVI
H1G	1551.68*	1054		.934	.062	8.20
H1G ² : Model H1G without paths from NA2, SN2, and VB2 to PRT; and from VB2 to GA	1555.52*	1058	3.84 (4)	.934	.062	8.18
H2G	1502.57*	1054		.941	.059	7.98
H2G ² : Model H2G without paths from NaP2, SN2, and VB2 to PRT; and from VB2 to GA	1506.89*	1058	4.32 (4)	.941	.059	7.96
H3G	1482.16*	1050		.943	.059	7.92

Table 14 continued.

Model	S- $\beta\chi^2$	df	$\Delta S-\beta\chi^2$ (Δdf) from corresponding a priori model	*CFI	RMSEA	*ECVI
H4G	1416.80*	1048		.951	.056	7.64
H4G ² : Model H4G without paths from NaP2, MTH2, SN2, and VB2 to PRT, and from SN2 to GA	1423.09*	1053	6.29 (5)	.951	.056	7.62

Note. S- $\beta\chi^2$ = Satorra-Bentler chi-square, *CFI = robust Comparative Fit Index, RMSEA = Root Mean Square Error of Approximation; *ECVI = robust Expected Cross-Validation Index; NA2 = Nonacademic; SN2 = Science-Academic; VB2 = Verbal-Academic; PRT = Parents; GA = General-Academic; NaP2 = Nonacademic/Physical, MTH2 = Math-Academic.

* $p \leq .00000$.

Table 15
Hierarchical Models: Summary of the Goodness-of-fit Indices for the Sample of Boys (N = 205)

Model	S- $\beta\chi^2$	df	$\Delta S-\beta\chi^2$ (Δdf) from corresponding a priori model	*CFI	RMSEA	*ECVI
H1B	1431.12*	1054		.925	.063	8.21
H1B ² : Model H1B without paths from NA2 and SN2 to PRT	1431.77*	1056	0.65 (2)	.926	.063	8.20
H2B	1309.34*	1054		.949	.060	7.61
H2B ² : Model H2B without paths from SN2 and NC2 to PRT	1309.96*	1056	0.62 (2)	.950	.060	7.60
H2B ³ : Model H2B without paths from SN2, NC2, and NaP2 to PRT	1315.30*	1057	5.96 (3)	.949	.060	7.61
H3B	1300.87*	1050		.950	.060	7.61

Table 15 continued.

Model	S- $\beta\chi^2$	df	$\Delta S-\beta\chi^2$ (Δdf) from corresponding a priori model	*CFI	RMSEA	*ECVI
H4B	1264.13*	1048		.957	.058	7.45
H4B ² : Model H4B without paths from SN2, MTH2, and NC2 to PRT; and from SN2 to GA	1264.98**	1052	0.85 (4)	.958	.058	7.42
H4B ³ : Model H4B without paths from SN2, MTH2, NC2 and NaP2 to PRT; and from SN2 to GA	1270.34*	1053	6.21 (5)	.957	.058	7.43

Note. S- $\beta\chi^2$ = Satorra-Bentler chi-square; *CFI = robust Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; *ECVI = robust Expected Cross-Validation Index; NA2 = Nonacademic; SN2 = Science-Academic; PRT = Parents; NC2 = Non-core Academic; NaP2 = Nonacademic/Physical; MTH2 = Mathematics-Academic; GA = General-Academic.

* $p \leq .00000$ ** $p \leq .00001$.

same second-order SC factor (i.e., Nonacademic/Physical SC). The separate second-order representation of these factors under Model H1 may not adequately account for their association, and thus, leads to a correlation between their disturbance terms. If this is so, Model H2 should provide a better fit to the data than does Model H1. Conversely, Model H4 (see Figure 10) offers the potential to address the correlated residual error between the two specific mathematics SC facets identified under Model H1 by allowing for a second-order separation between the first-order science and mathematics SC facets. If the correlation among the disturbance terms for the two specific first-order mathematics SC constructs found for Model H1G is due to the lack of second-order differentiation of these constructs from the first-order order science SC facets, Model H4G should provide a better fit to the data than does Model H1G. Thus, the two areas of misfit identified in Model H1G may potentially be addressed by the alternate theoretical models hypothesized in the present study.

Given the particular areas of misfit identified in Model H1G, and the potential for the alternate hypothesized theoretical models (i.e., models H2 and H4) to correct these areas of misfit with a priori specified model modifications, the two correlated disturbance terms identified by the LM-Test for Model H1G were not incorporated into this model. Instead, the fit of this model will be contrasted with the fit of the alternate hypothesized models. Accordingly, we turn to the Wald-Test results for Model H1G to determine the presence of nonsignificant model parameters.

The Wald-Test results for Model H1G reveal four nonsignificant model parameters. These parameters represent the path linking the first-order General-Academic SC factor with the second-order Verbal-Academic SC facet, and the paths linking the first-order Parent SC factor with the second-order: (a) Nonacademic, (b) Science-Academic, and (c) Verbal-Academic SC facets. It is important to note that deletion of these paths from Model H1G, in model H1G², in no way deteriorates the previously noted good fit to the data ($S-\beta \chi^2_{(1058)} = 1555.52$, *CFI = .934, RMSEA = .062; *ECVI = 8.18). That is, as expected, the difference in fit between Model H1G and Model H1G² is not significant ($\Delta S-\beta \chi^2_{(4)} = 3.84$, $\Delta^*CFI =$

.000), thereby supporting the deletion of the four nonsignificant paths from Model H1G. Accordingly, the final model, Model H1G², which explains 93.4% of the variance in the data, is considered to best represent the fit of Model H1 to the sample data for girls when statistical, practical, and theoretical criteria are considered.

In testing for the validity of the second hypothesized hierarchical model for girls (Model H2G), results reveal that this model also provides a good fit to the data ($S-\beta \chi^2_{(1054)} = 1502.57$, $*CFI = .941$, $RMSEA = .059$; $*ECVI = 7.98$). As expected, Model H2G yields an improvement in the values of the model chi-square and practical fit indices as compared to Model H1G ($\Delta S-\beta \chi^2_{(0)} = 49.11$, $\Delta *CFI = .007$). The LM-Test identifies the same two distinctly ill-fitting parameters that were identified with Model H1G, however, the magnitude of their associated chi-square values are consistent with expectations. That is, the parameter representing the correlation among the disturbance terms for the first-order mathematics SC constructs (i.e., Arithmetic and Measurement; $LM \chi^2 = 62.17$) exhibits the larger LM chi-square value in Model H2G. Thus, as expected, the correlated error among these SC facets remains as an area of misfit, and the associated chi-square value is of a similar value to that obtained with Model H1G ($LM \chi^2 = 63.71$ vs. 62.17 for the H1G and H2G models, respectively). Also consistent with expectations, the correlated error among the first-order Gym and Physical Ability SC facets drops to the second, rather than the first parameter identified, and the size of its associated chi-square value is substantially smaller for Model H2G than for Model H1G ($LM \chi^2 = 48.60$ vs. 82.70 , respectively). Thus, although the correlated error for the Physical Ability and Gym SC facets remains with Model H2G, the size of the systematic error in the prediction of these facets is substantially diminished. In addition, the drop in the chi-square value associated with this parameter renders its distinction from the remaining parameters less marked.

The continued presence of the correlated disturbance associated with the Gym and Physical Ability SC facets, despite its diminishing distinction, suggests the possibility that Model H2G does not fully account for the misfit in this area of the model. However, the a priori hypothesized Model H3 (see Figure 9) offers the potential to address the remainder of

this misfit in Model H2G with an increased second-order differentiation. Specifically, Model H3 differentiates the first-order nonacademic physical (i.e., Gym, Physical Ability, and Physical Appearance) and social (i.e., Peer and Parent) SC facets at the second-order level. If the continued misfit related to the first-order Gym and Physical Ability SC facets is due to their lack of differentiation from the nonacademic social SC facets, the second-order differentiation offered with Model H3 should yield a model with a better fit to the data than either the H1 or H2 models, which do not allow for this differentiation.

Given that, as with Model H1G, the areas of misfit identified in Model H2G can potentially be accounted for by the a priori modifications specified in subsequent models (i.e., models H3 and H4), the correlated disturbance terms identified for Model H2G were not incorporated into this model. Accordingly, the Wald-Test results for Model H2G were examined for the presence of nonsignificant parameters, of which five are identified. One of these parameters is the disturbance term representing the error in the prediction of the first-order Physical Ability SC facet; thus, the representation of the Physical Ability and Gym SC facets under the same second-order SC factor results in a nonsignificant error in the prediction of one of these facets. Given that the deletion of a parameter representing error in the prediction of a latent variable is considered inappropriate (e.g., Bentler, 1990), this parameter is retained as a freely estimated parameter.

The remaining parameters identified as nonsignificant in Model H2G represent paths linking the first- and second-order SC factors. These paths are analogous to those identified for Model H1G. Specifically, the parameters represent the path linking the first-order General-Academic SC facet with the second-order Verbal-Academic SC facet, as well as the paths linking the first-order Parent SC facet with three of the second-order SC factors (i.e., Nonacademic/Physical, Verbal-Academic, and Science-Academic). As expected, the deletion of these four parameters does not significantly alter the final H2G model, Model H2G² ($\Delta S-\beta \chi^2_{(4)} = 4.32, \Delta *CFI = .000$), which explains 94.1% of the variance in the data ($S-\beta \chi^2_{(1058)} = 1506.89, *CFI = .941, RMSEA = .059; *ECVI = 7.96$). In addition, Model

H2G² yields an improvement in the value of the chi-square and practical fit indices, as compared to Model H1G² ($\Delta S-\beta \chi^2_{(0)} = 48.63$, $\Delta *CFI = .007$).

We turn now to a review of the results for the next hierarchical model, Model H3 (see Figure 9), in which the first-order nonacademic social and physical SC dimensions are represented separately at the second-order level; recall that this model also reconceptualized the non-core academic portion of the SC hierarchy in that first-order Art and Parent SCs are defined by a second-order Art SC facet, and that Religion SC is represented under the second-order Social SC factor. Overall fit indices obtained for the third hierarchical model for girls (designated Model H3G) suggest that this model also provides a good fit to the data ($S-\beta \chi^2_{(1050)} = 1482.16$, $*CFI = .943$, $RMSEA = .059$; $*ECVI = 7.92$). However, the paths linking the first-order facets with the second-order Art SC facet are negative, and the correlations among the second-order Social and Art SC facets are not only very high, but also yield negative signs. Thus, the appropriate and separate estimation of the second-order Social and Art SCs is problematic. Given the estimation problems with this model, it is not assessed further.

Given the results of the second-order CFAs of the three hierarchical models reviewed thus far, Model H2G² is considered to best represent SC structure for preadolescent girls. That is, comparison of the fit indices for the H1G and H2G models suggests that the overall fit of Model H2G² (and H2G) is superior to that of Model H1G² (and H1G). The size of the chi-square values associated with the correlated error of prediction between Physical Ability and Gym SC facets for these models indicates that the correlation between their disturbance terms is smaller for the H2G models than for the H1G models. Taken together, the final H2G² model is considered to better represent the structure of SC for girls than does the final H1G² model. Accordingly, the hierarchical representation of SC specified in Model H2 was utilized to test the increased hierarchical differentiation hypothesized in Model H4 (i.e., between science and mathematics SC facets; see Figure 10) for girls. The nonsignificant parameters identified in analyses of Model H2G were, however, retained as estimated

parameters in analyses of Model H4 in order to allow for an independent evaluation of this model.

Estimation of Model H4 (designated Model H4G for girls) indicates that this expanded model provides an excellent fit to the sample data for girls ($S-\beta \chi^2_{(1048)} = 1416.80$, *CFI = .951, RMSEA = .056; *ECVI = 7.64). The fit of Model H4G is better than that of the model (Model H2G) upon which it is based ($S-\beta \chi^2_{(1054)} = 1502.57$, *CFI = .941, RMSEA = .059; *ECVI = 7.98). Not surprisingly, the LM-Test, as with previous models, indicates that the fit of this model can be most improved with the addition of a parameter which represents correlated disturbance terms between the Physical Ability and Gym SC facets ($LM \chi^2 = 47.30$). As expected, the value associated with this parameter is similar to that obtained with Model H2G (i.e., $LM \chi^2 = 48.60$), and smaller than that of Model H1G (i.e., $LM \chi^2 = 82.70$). However, unlike Model H2G but also as expected, the correlated residual error detected in previous models between the two first-order mathematics SCs is not identified with Model H4G. Thus, the separation of the specific first-order mathematics and science SC facets at the second-order level eliminates the presence of correlated residual error in the prediction of the two specific first-order mathematics SC facets. Taken together, results indicate that Model H4G provides an excellent fit to the data, addresses the misfit among the specific mathematics SCs in Model H2G, and, as with Model H2G, results in a reduction in the misfit of the first-order Gym and Physical Ability SC components identified in Model H1G.

On the basis of a number of criteria, the correlated disturbance term between the Gym and Physical Ability SC facets representing an area of potential model misfit in Model H4G was not incorporated into a subsequent H4G model. Specifically, these considerations included: (a) the already excellent fit of Model H4G to the sample data for girls; (b) the reduction in misfit achieved with the change in the second-order representation of Gym SC under Model H4G as compared to Model H1G, and thus, the reduction in the gain in model fit with the inclusion of this parameter; (c) the lack of substantive meaningfulness of the parameter; and (d) the loss of model parsimony and interpretability associated with the

introduction of a correlated disturbance term. Accordingly, Model H4G is further evaluated only in terms of nonsignificant model parameters.

The Wald-Test for Model H4G identifies eight parameters as being nonsignificant to this model. Three of these parameters represent variances associated with the prediction of the first-order SC facets, namely, SCs in Physical Ability, and both Listening and Writing in Language Arts. As with Model H2G, parameters representing error in the prediction of a latent variable remained estimated despite their nonsignificance. The remaining parameters identified in Model H4G are, with one exception, analogous to the nonsignificant paths identified with the previous H1G and H2G models. Specifically, the analogous paths identified are those linking the first-order Parent SC construct with the second-order (a) Nonacademic/Physical SC facet, and (b) core academic SC facets (in this case, Mathematics-, Science- and Verbal-Academic SCs). The path linking first-order General-Academic SC with the second-order Science-Academic SC factor is also identified as nonsignificant; in Model H1G and H2G, the path linking General-Academic SC with the second-order Verbal-Academic SC facet is not significant.

As expected, the deletion of the five nonsignificant paths linking the first- and second-order SC facets does not significantly alter the fit of the final H4G model, Model H4G² ($\Delta S-\beta \chi^2_{(5)} = 6.29$, $\Delta *CFI = .000$). With the deletion of these paths, only the variance representing the error in the prediction of the Physical Ability SC facet is identified as nonsignificant by the Wald-Test. Model H4G² yields an excellent fit to the sample data for girls ($S-\beta \chi^2_{(1053)} = 1423.09$, $*CFI = .951$, $RMSEA = .056$; $*ECVI = 7.62$) and explains 95.1% of the variance in the data. Thus, Model H4G² is considered the final best-fitting H4G model. In addition, Model H4G² provides a better fit to the sample data for girls than does the final model upon which it is based, namely, Model H2G² ($S-\beta \chi^2_{(1058)} = 1506.89$, $*CFI = .941$, $RMSEA = .059$; $*ECVI = 7.96$). In sum, of the four hierarchical models hypothesized, Model H4G² is considered to best represent the structure of SC for preadolescent girls. We turn now to a review of results related to the tests of competing hierarchical structures for the sample of boys (see Table 15).

Preadolescent Boys: Testing for the Hierarchical Structure of Self-concept

Covariance structure analysis of Model H1 (see Figure 7), designated Model H1B for boys, indicates that this model provides an adequate fit to the sample data for boys ($S-\beta \chi^2_{(1054)} = 1431.12$, $*CFI = .925$, $RMSEA = 0.063$; $*ECVI = 8.21$). As with the results for girls, the distinct parameters identified by the LM-Test represent correlated disturbance terms between the Gym and Physical Ability SC facets ($LM \chi^2 = 93.89$), and between the Arithmetic and Measurement SC facets ($LM \chi^2 = 35.51$). Interestingly, the chi-square value associated with the former correlated disturbance term under Model H1B is slightly higher than that obtained for the corresponding model for girls (i.e., $LM \chi^2 = 82.70$). The value obtained for the second parameter, however, is substantially lower under Model H1B than that obtained for girls under Model H1G (i.e., $LM \chi^2 = 63.71$). As a result, for Model H1B, only the parameter representing the correlation among the disturbance terms for Gym and Physical Ability SC is distinct from the remaining parameters; the chi-square value associated with the second parameter, the correlation between the disturbance terms for Arithmetic and Measurement SCs, is only modestly distinct from the remaining chi-square values. Tests of the invariance of the SC construct across girls and boys, reviewed later, allowed an empirical examination of the significance of this differing pattern of results.

As with the sample of girls, the alternate hypothesized models offer the potential to correct the areas of misfit identified for Model H1B. Accordingly, Model H1B is further evaluated only to determine if there are unnecessary model parameters whose deletion does not significantly degrade the fit of this model. The Wald-Test results from Model H1B identifies two parameters as being nonsignificant, namely, the regression paths linking the first-order Parent SC factor with the second-order Nonacademic and Science-Academic SC facets. The deletion of these two paths in Model H1B² yields an adequate fit to the sample data for boys ($S-\beta \chi^2_{(1056)} = 1431.77$, $*CFI = .926$, $RMSEA = .063$; $*ECVI = 8.20$), with no additional parameters identified as nonsignificant by the Wald-Test. The change in fit from Model H1B to Model H1B² is not significant ($\Delta S-\beta \chi^2_{(2)} = 0.65$, $\Delta *CFI = .001$), and thus, supports the deletion of these paths. Accordingly, Model H1B², which explains 92.6% of the

variance in the data for the sample of boys, is considered the best Model H1 representation of preadolescent SC for boys.

Hierarchical CFA of Model H2 (see Figure 8) for the sample of boys (designated Model H2B) yields a model which provides an excellent fit to the data ($S-\beta \chi^2_{(1054)} = 1309.34$, $*CFI = .949$, $RMSEA = .060$; $*ECVI = 7.61$)², and a substantial increase in fit over that of Model H1B ($\Delta S-\beta \chi^2_{(0)} = 121.78$, $\Delta *CFI = .024$). Interestingly, the improvement in model fit with the change from model H1 to H2 is substantially larger for boys than for girls; that is, the change in both the model chi-square and fit indices for the sample of boys is double that obtained for the sample of girls. The LM-Test results for Model H2B reveal three modestly distinct ill-fitting parameters. As with the results for the H2 model for girls, these parameters suggest model misfit in relation to first-order SC facets represented under the rubrics of the second-order mathematics/science and nonacademic/physical portions of this model. That is, the LM-Test for Model H2B identifies the correlation between the disturbance terms for the first-order mathematics SC constructs (i.e., Arithmetic and Measurement; $LM \chi^2 = 34.31$). However, as with the pattern of results for Model H1, the LM chi-square value associated with this parameter for boys is, in comparison to that for girls under Model H2G (i.e., $LM \chi^2 = 62.17$), substantially smaller, and as a result, only modestly distinct from that of the remaining parameters.

With regard to the nonacademic/physical portion of Model H2B, model misfit is represented differently for boys than is true of the corresponding model for girls. Specifically, the first and third parameters identified by the LM-Test represent, respectively, a correlation between the disturbance terms for the Peer and Physical Appearance SCs ($LM \chi^2 = 38.19$), and the regression path linking the first-order Physical Appearance SC facet with the second-order Verbal-Academic SC factor ($LM \chi^2 = 29.44$). Thus, the misfit under the rubric of the second-order Nonacademic/Physical SC facet in Model H2 is represented

² The variance associated with the error in the prediction of Physical Ability SC approached zero, and thus, was constrained to its lower bound (i.e., equal to a value of 0.0) by the EQS program; this was true of all H2B models.

by the physical SCs (i.e., Physical Ability and Gym) for girls, whereas, for boys, it is represented in the Physical Appearance and Peer SC facets. The difference in the area of representation of misfit in the H2 models across gender is consistent with both (a) the modestly larger LM chi-square value for the correlated disturbance terms between the Gym and Physical Ability SCs facets obtained with the sample of boys than with the sample of girls under Model H1, and (b) the substantially larger gain in fit of Model H2 over Model H1 for boys than for girls. Tests of the gender invariance of SC measurement and structure, reviewed later, allow for an examination of gender differences relevant to this differing pattern of results.

As with the H2 models for the sample of girls, the areas of model misfit identified for Model H2B can potentially be addressed with the a priori specified H3 and H4 models. Accordingly, Model H2B is further evaluated only for the presence of nonsignificant model parameters. The Wald-Test results for Model H2B identifies two nonsignificant parameters. These parameters represent the regression paths linking the first-order Parent SC factor with both the second-order Science-Academic and Non-Core Academic SC facets. Accordingly, these parameters were not included as freely estimated parameters in Model H2B² ($S-\beta \chi^2_{(1056)} = 1309.96$, *CFI = .950, RMSEA = .060; *ECVI = 7.60), with the fit of the model not diminishing to a significant degree ($\Delta S-\beta \chi^2_{(2)} = 0.62$, $\Delta^*CFI = .001$). However, an additional regression path is identified as being nonsignificant, namely, the path linking the Parent SC facet with the second-order Nonacademic/Physical SC facet. The deletion of this third path results in a final H2B³ model which yields an excellent fit to the data ($S-\beta \chi^2_{(1057)} = 1315.30$, *CFI = .949, RMSEA = .060; *ECVI = 7.61). The change in fit from Model H2B to H2B³ is not significant ($\Delta S-\beta \chi^2_{(3)} = 5.96$, $\Delta^*CFI = .000$), despite the change in fit from Model H2B² to H2B³ being significant at the .05 alpha level ($\Delta S-\beta \chi^2_{(1)} = 5.34$, $\Delta^*CFI = .001$). The fit of Model H2B³ is significantly better than that of Model H1B² ($\Delta S-\beta \chi^2_{(1)} = 116.47$, $\Delta^*CFI = .023$). Thus, statistical and practical criteria suggest that Model H2B³ provides a better fit than does Model H1B². Accordingly, Model H1B² is rejected in favor of

Model H2B³. We turn now to a review of results related to Model H3 (see Figure 9) for the sample of boys.

Estimation of Model H3 (designated Model H3B for the sample of boys), as with the results for girls, is problematic. That is, the problematic association identified between the second-order Art and Social SCs with the sample of girls is present to a larger extent with the sample of boys. In particular, the regression paths linking the second-order Art SC facet with its first-order SCs are negative, and the EQS program constrained the parameter representing the association between the second-order Social and Art SC facets to a value of -1.00. Accordingly, this model is not evaluated further, and, as with the sample of girls, Model H2B is utilized in the expansion hypothesized under Model H4 (see Figure 10).

Estimation of Model H4 for the sample of boys (designated Model H4B) reveals that this model provides an excellent fit to the data ($S-\beta \chi^2_{(1048)} = 1264.13$, *CFI = .957, RMSEA = .058; *ECVI = 7.45)³. The fit of Model H4B is substantially better than that of the original Model H2B upon which it is based ($S-\beta \chi^2_{(1054)} = 1309.34$, *CFI = .949, RMSEA = .060; *ECVI = 7.61). As with the sample of girls, the correlated error among the disturbance terms for the two first-order mathematics SC facets is not identified. Thus, the separation of the specific first-order mathematics and science SCs at the second-order level eliminates the presence of a significant correlated error in the prediction of the two specific first-order mathematics SC facets. This pattern of results, and the consistency of the pattern across gender, indicates that the separation of the first-order mathematics SC facets from the first-order science SC facets at the second level of the SC hierarchy yields a better fitting model than does their joint second-order representation.

The LM-Test results bearing on the SC facets under the rubric of the second-order Nonacademic/Physical SC for Model H4B are consistent with those obtained for Model H2B. That is, not unexpectedly and as with Model H2B, the LM-Test for Model H4B reveals a modest degree of misfit for the Physical Appearance and Peer SC facets.

³ As with the H2B models, the error variance associated with the prediction of the Physical Ability SC facet for all H4B models was constrained at its lower bound by the EQS program.

Specifically, three parameters exhibiting, at most, a modest degree of misfit are identified. These parameters represent (a) a correlated disturbance term between the Peer and the Physical Appearance SC facets (LM $\chi^2 = 38.58$), and (b) the regression paths linking the second-order Verbal-Academic SC facet with the first-order Peer (LM $\chi^2 = 29.46$) and Physical Appearance SC facets (LM $\chi^2 = 20.11$). Given that the first-order SC facets represented in the first parameter are the same as those identified in the second and third parameters, it is likely that incorporation of the first parameter will address the misfit represented in all three parameters; this is substantiated with further analyses. However, as with this model for the sample of girls, and on the basis of the same criteria, this post hoc model parameter is not incorporated into Model H4B. Accordingly, Model H4B is further evaluated only for the presence of nonsignificant parameters.

A total of four parameters are identified as nonsignificant by the Wald-Test for Model H4B. These include the regression paths linking Parent SC to the second-order (a) Non-core Academic, (b) Mathematics-Academic, and (c) Science-Academic SC facets. In addition, the path linking the first-order General-Academic and second-order Science-Academic SC facets is not significant. The deletion of these parameters in Model H4B² yields a model with an excellent fit to the data (S- β $\chi^2_{(1052)} = 1264.98$, *CFI = .958, RMSEA = .058; *ECVI = 7.42), with the drop in the fit of Model H4B², as compared to Model H4B, being nonsignificant (Δ S- β $\chi^2_{(4)} = 0.85$, Δ *CFI = .001). However, as with Model H2B², the Wald-Test for Model H4B² identifies an additional nonsignificant parameter representing the regression paths linking the first-order Parent SC facet with the second-order Nonacademic/Physical SC factor. The deletion of this parameter in Model H4B³ again yields a model which provides an excellent fit to the data (S- β $\chi^2_{(1053)} = 1270.34$, *CFI = .957, RMSEA = .058; *ECVI = 7.43). Also consistent with the findings for the H2B models, the shift to Model H4B³ from Model H4B does not result in a significant degradation in model fit (Δ S- β $\chi^2_{(5)} = 6.21$, Δ *CFI = .000), despite the shift from Model H4B² to Model H4B³ being significant at the .05 alpha level (Δ S- β $\chi^2_{(1)} = 5.36$, Δ *CFI = .001). The fit of Model H4B³ is superior to that of both its corresponding H2B model, Model H2B³ (S- β $\chi^2_{(1057)} =$

1315.30, *CFI = .949, RMSEA = .060; *ECVI = 7.61), and the initial Model H2B upon which it is based ($S-\beta \chi^2_{(1054)} = 1309.34$, *CFI = .949, RMSEA = .060; *ECVI = 7.61).

Taken together, these results indicate that the final H4B model, Model H4B³, yields a better fitting representation of the relations among the first-order SC facets than do the H2B models. Accordingly, Model H2B³ is rejected in favor of Model H4B³. Model H4B³ is, therefore, the baseline model used for the sample of boys in subsequent analyses, namely, in testing for the gender invariance of the SC construct, and in examining SC/AA relations. A summary of goodness-of-fit statistics related to the competing hierarchical models for preadolescent boys was presented in Table 15.

Testing for the Hierarchical Structure of Preadolescent Self-concept for Girls and Boys:

A Summary

The multidimensional and hierarchical models considered to best represent preadolescent SC structure for both girls and boys are based on the hypothesized Model H4. The final second-order baseline models for girls and boys, Models H4G² and H4B³, respectively, along with their first- and second-order robust estimates, are presented in Figures 11 to 13 for girls, and in Figures 14 to 16 for boys. For ease of presentation, the estimates associated with the nonacademic and non-core academic SC facets are presented in Figures 11 and 14 for girls and boys, respectively. The estimates associated with the first-order portion of the core academic SC facets are presented in Figure 12 for girls, and in Figure 15 for boys. The estimates representing the relations between the first- and second-order SC facets, and among the second-order SC facets are presented in Figures 13 and 16 for the sample of girls and boys, respectively.

The paths included in the final baseline hierarchical SC models are, with the exception of one structural parameter, identical for girls and boys. At the first-order level of the SC construct, the baseline SC models for both girls and boys are comprised of the 16 hypothesized SC facets. Likewise, at the second-order level of the SC construct, SC is represented by the same five second-order SC facets, namely, Nonacademic/Physical, Mathematics-Academic, Science-Academic, Verbal-Academic, and Non-Core Academic

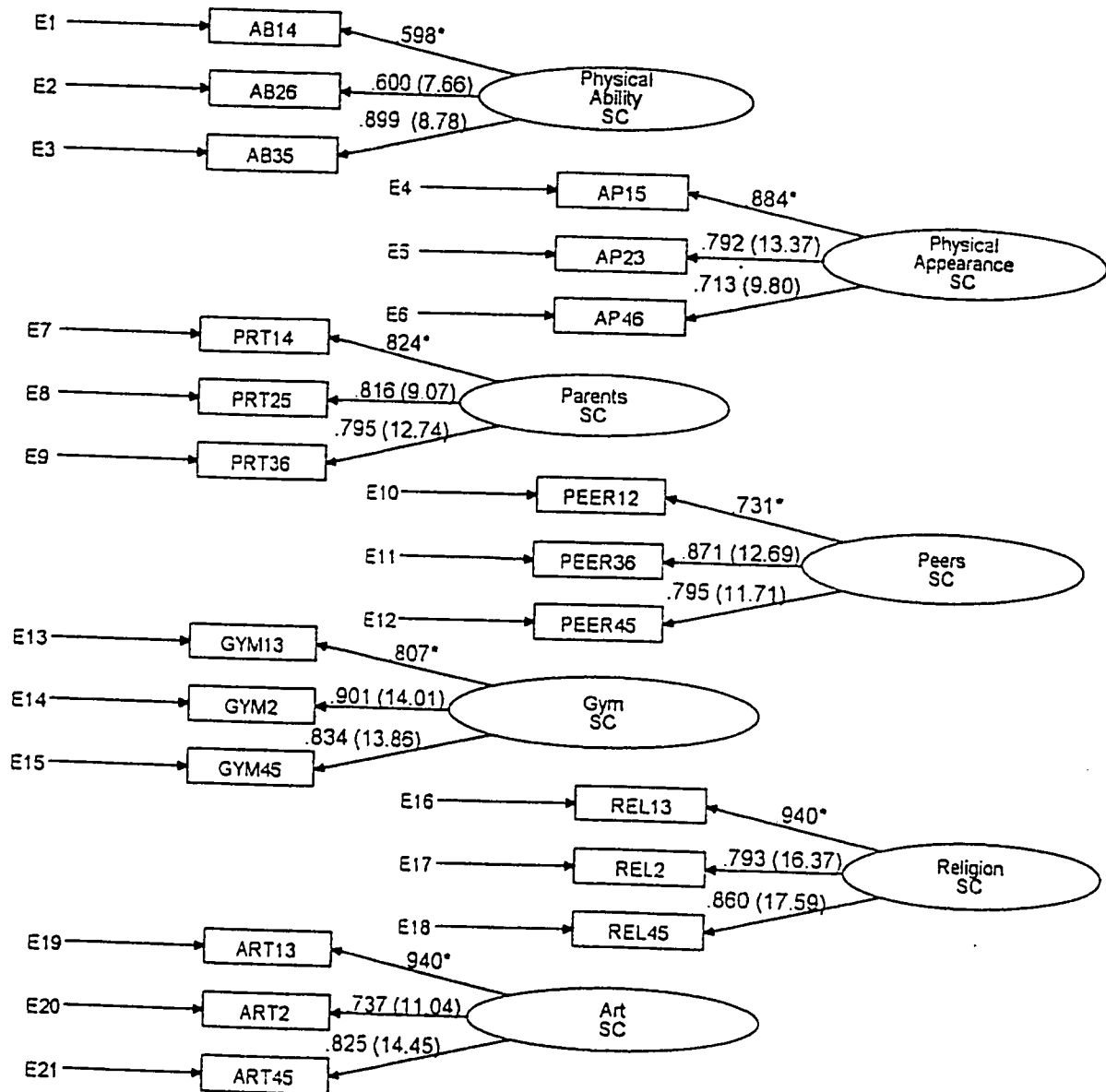


Figure 11. First-order estimates^a for the nonacademic and non-core academic portions of the final hierarchical self-concept (SC) model for girls (Model H4G²).

Note. Boxes represent indicator variables for the corresponding first-order SC factor, with the letters representing this same factor, and the numbers representing the (A)SDQ-I item number(s) from which they were derived. Thus, for example, AB14 represented the indicator variable pair derived from the SDQ-I Physical Ability scale item numbers 1 and 4.

^a All parameter estimates are standardized and statistically significant ($p < .05$). Values in parentheses represent robust z-statistics. * denotes a parameter fixed to 1.00 for purposes of statistical identification.

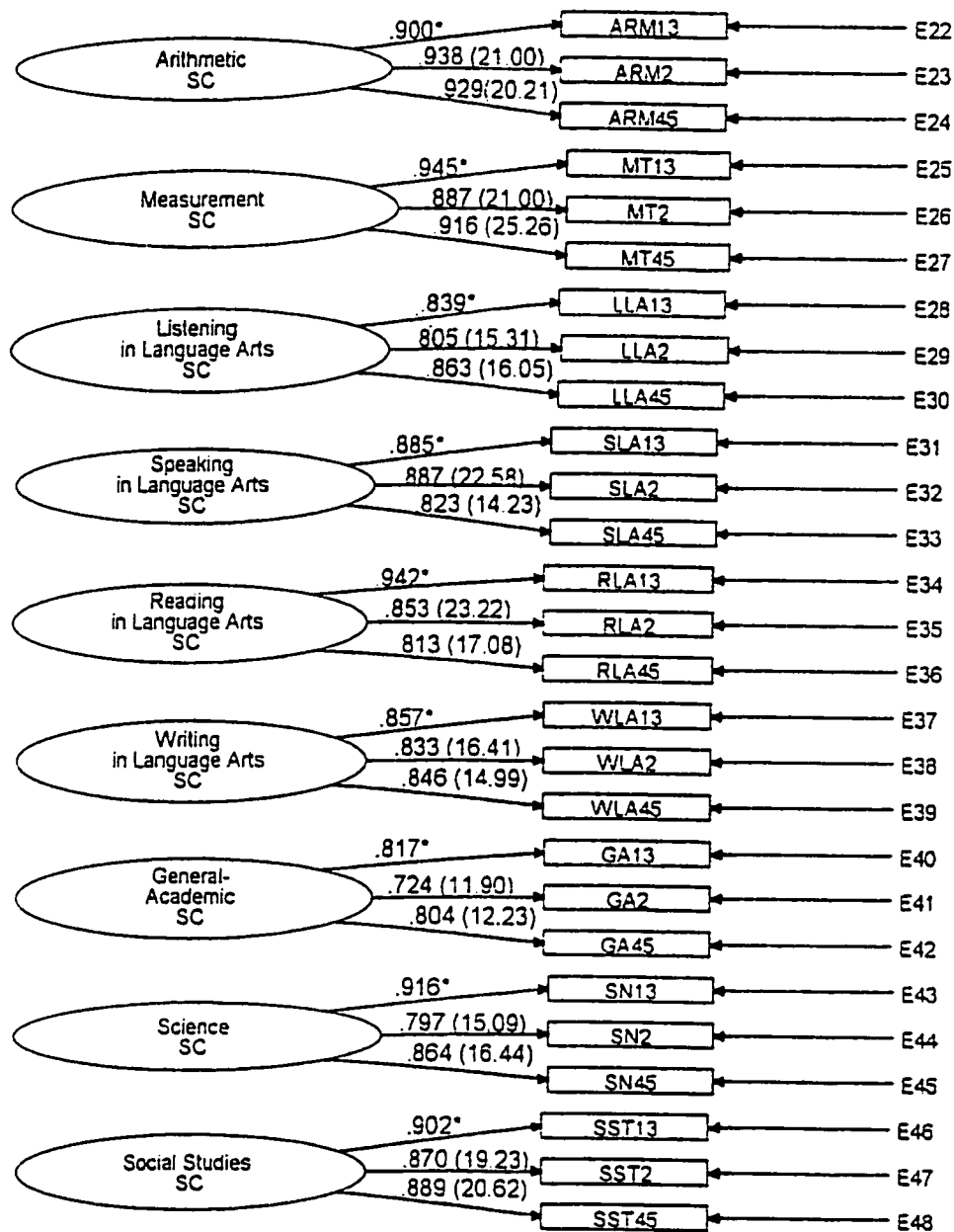


Figure 12. First-order estimates^a for the core academic portions of the final hierarchical self-concept (SC) model for girls (Model H4G²).

Note. Boxes represent indicator variables for the corresponding first-order SC factor, with the letters representing this same factor, and the numbers representing the ASDQ-I item number(s) from which they were derived. Thus, for example, MT13 represented the indicator variable pair derived from the ASDQ-I Measurement scale item numbers 1 and 3.

^a All parameter estimates are standardized and statistically significant ($p < .05$). Values in parentheses represent robust z-statistics. * denotes a parameter fixed to 1.00 for purposes of statistical identification.

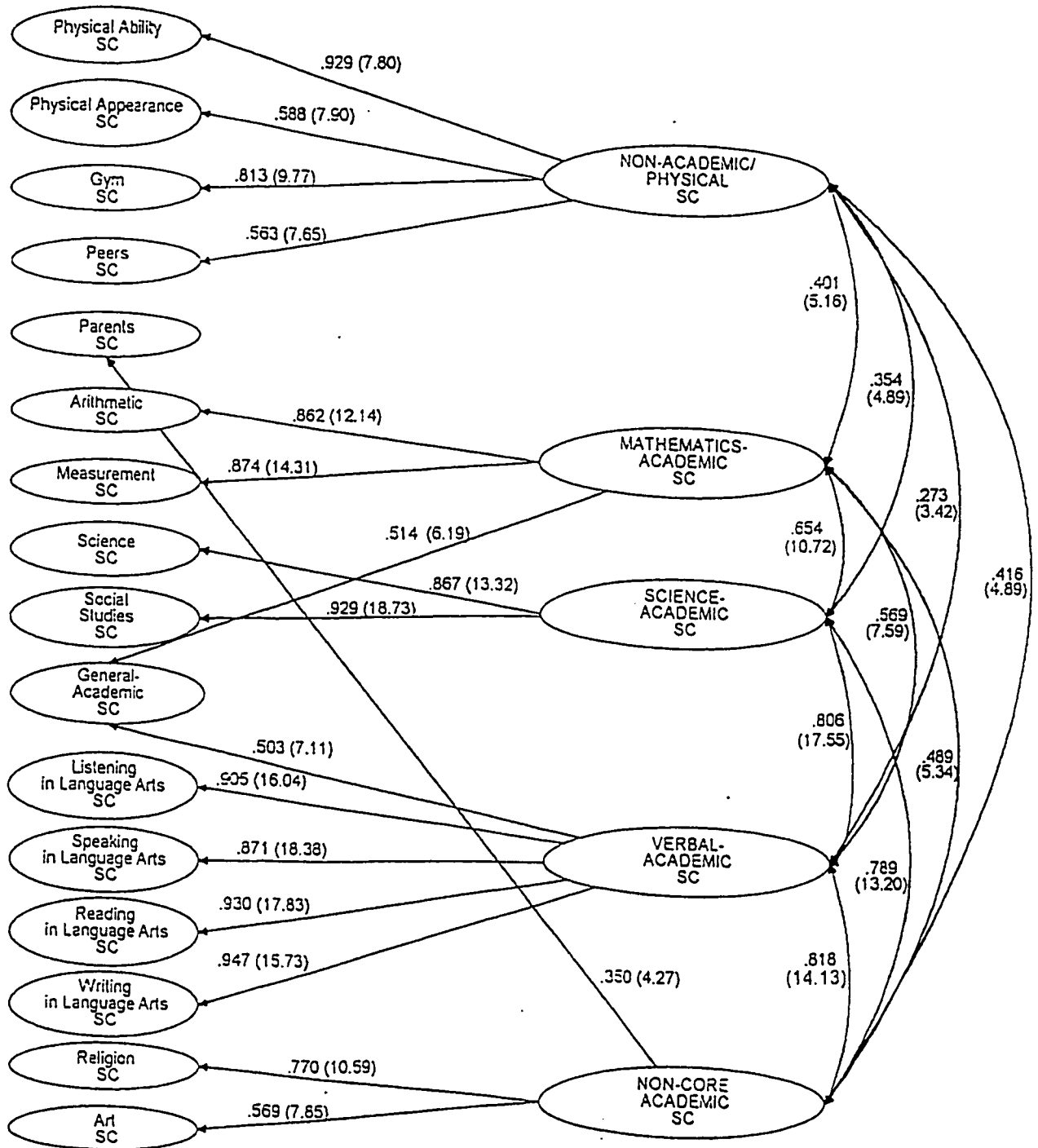


Figure 13. Second-order estimates^a for the final hierarchical self-concept (SC) model for girls (Model H4G²).

^a All parameter estimates are standardized and statistically significant ($p < .05$). Values in parentheses represent robust z-statistics.

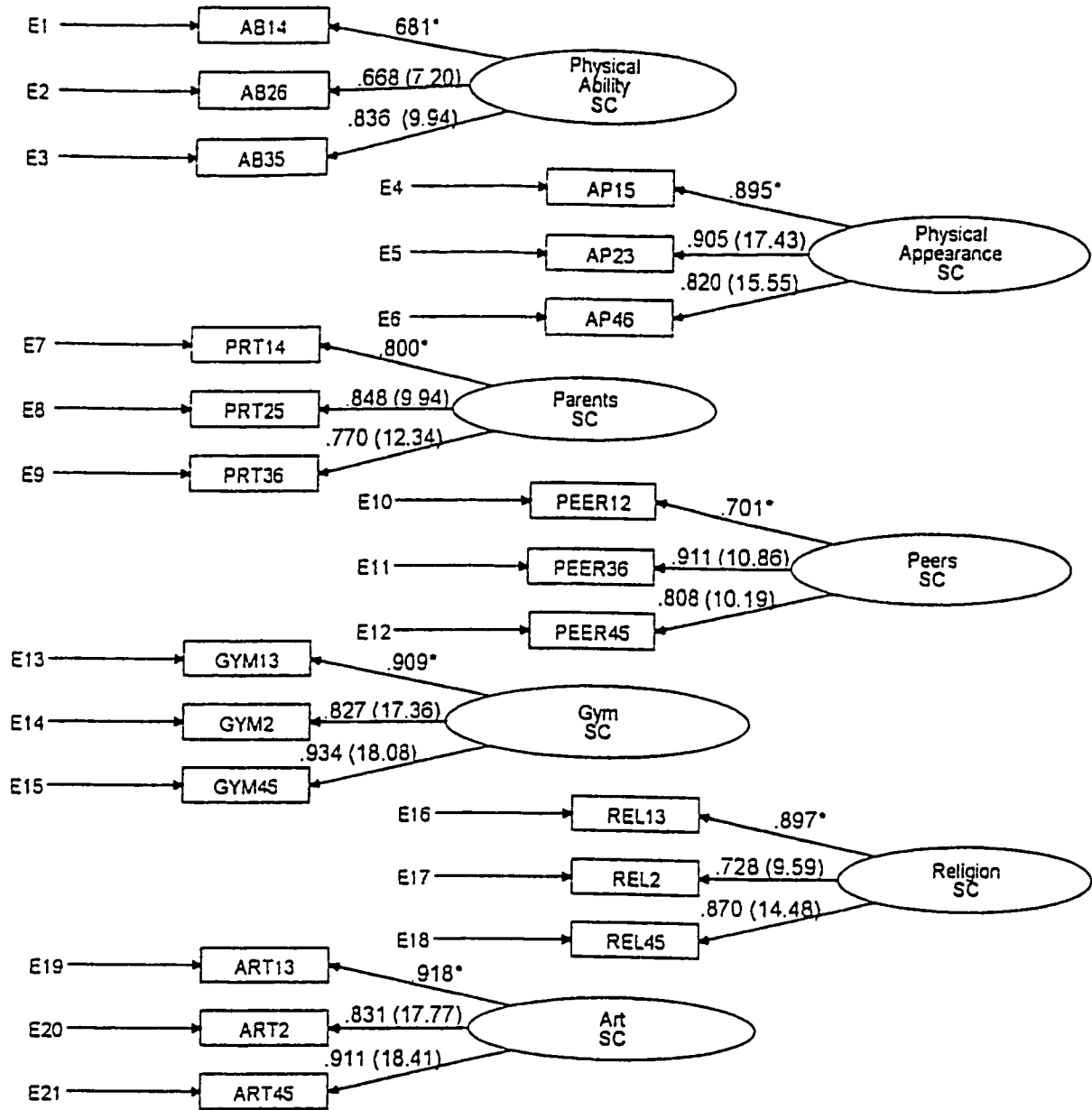


Figure 14. First-order estimates^a for the nonacademic and non-core academic portions of the final hierarchical self-concept (SC) model for boys (Model H4B³).

Note. Boxes represent indicator variables for the corresponding first-order SC factor, with the letters representing this same factor, and the numbers representing the (A)SDQ-I item number(s) from which they were derived. Thus, for example, AB14 represented the indicator variable pair derived from the SDQ-I Physical Ability scale item numbers 1 and 4.

^a All parameter estimates are standardized and statistically significant ($p < .05$). Values in parentheses represent robust z-statistics. * denotes a parameter fixed to 1.00 for purposes of statistical identification.

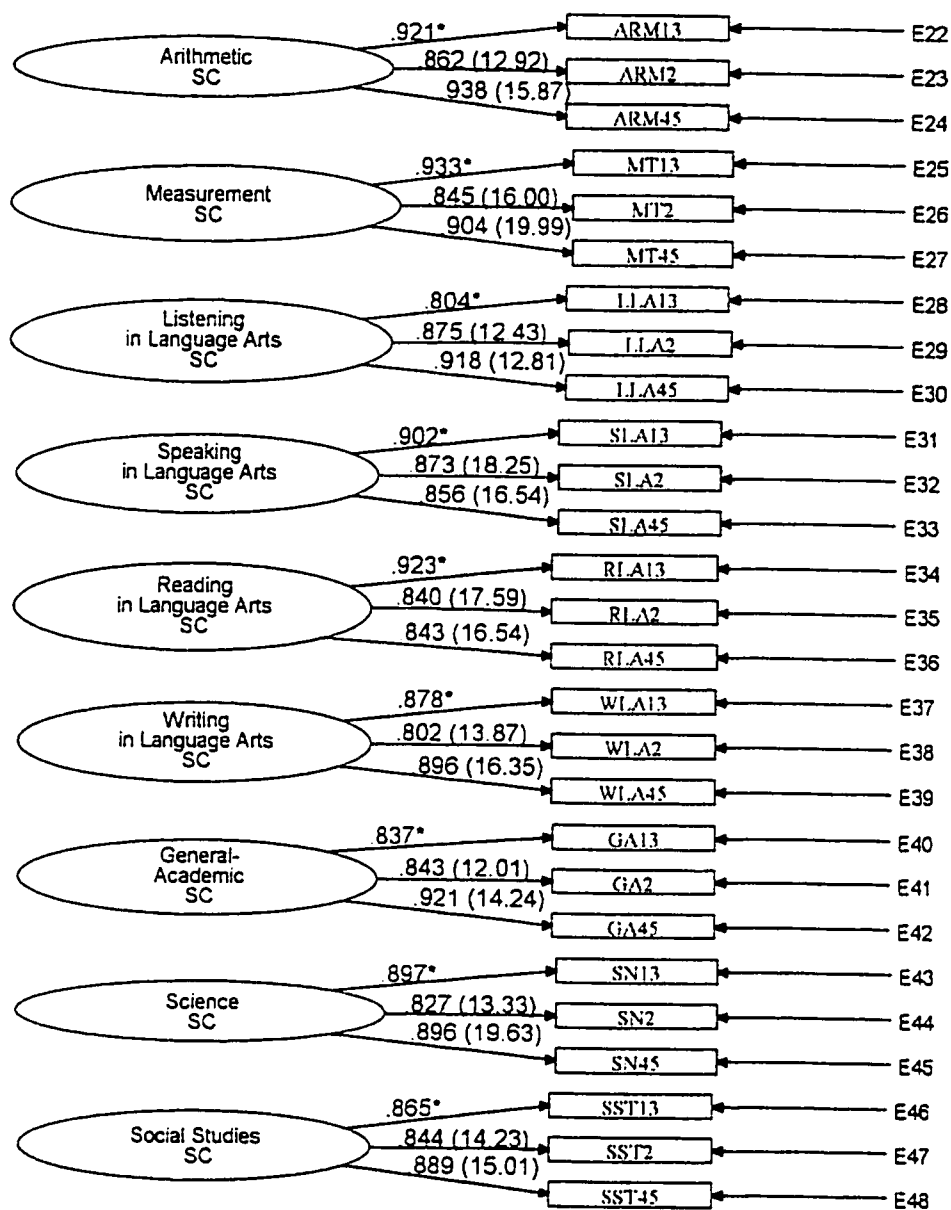


Figure 15. First-order estimates^a for the core academic portions of the final hierarchical self-concept (SC) model for boys (Model H4G³).

Note. Boxes represent indicator variables for the corresponding first-order SC factor, with the letters representing this same factor, and the numbers representing the ASDQ-I item number(s) from which they were derived. Thus, for example, MT13 represented the indicator variable pair derived from the ASDQ-I Measurement scale item numbers 1 and 3.

^a All parameter estimates are standardized and statistically significant ($p < .05$). Values in parentheses represent robust z-statistics. * denotes a parameter fixed to 1.00 for purposes of statistical identification.

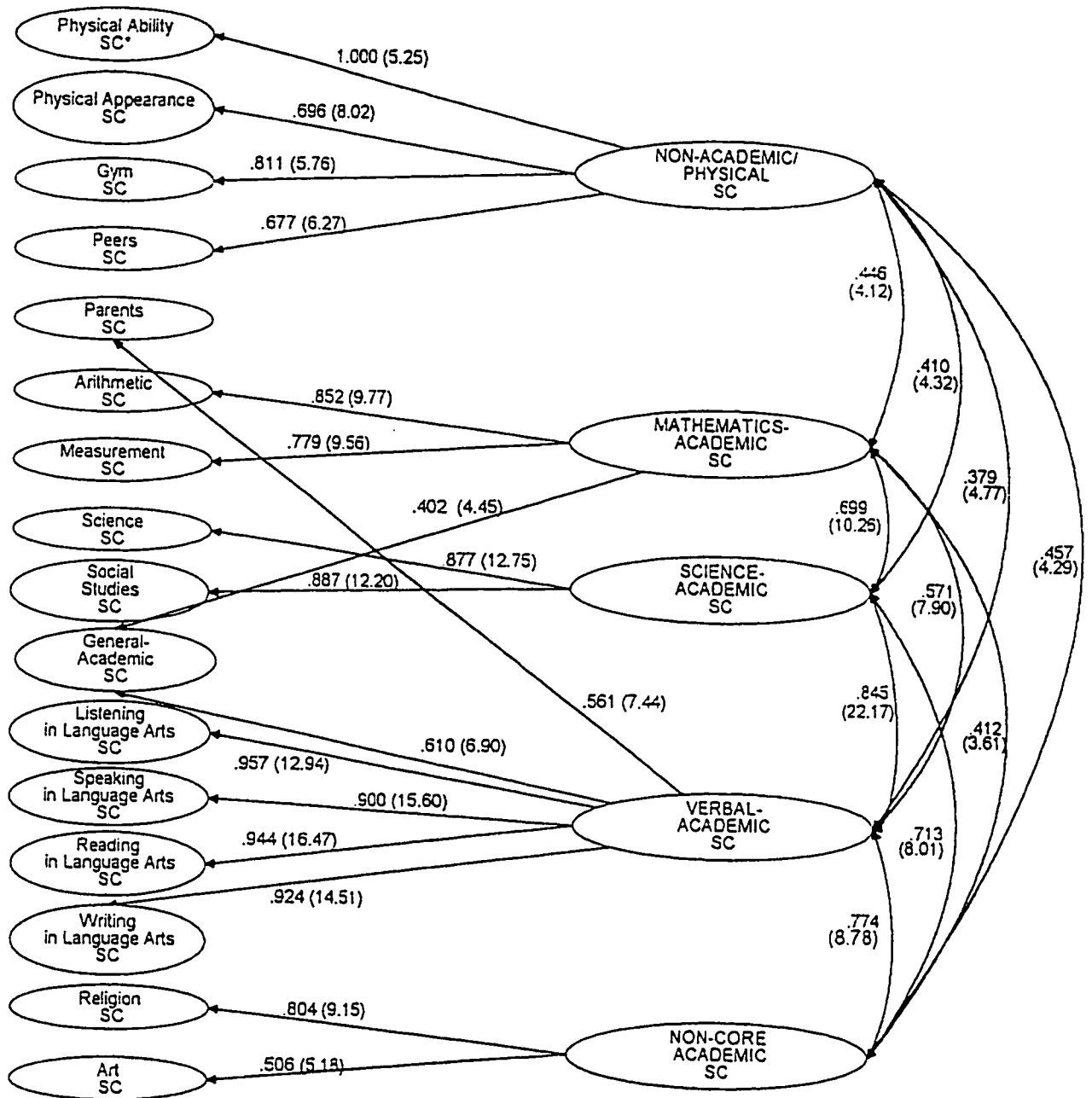


Figure 16. Second-order estimates^a for the final hierarchical self-concept (SC) model for boys (Model H4B³).

^a All parameter estimates are standardized and statistically significant ($p < .05$). Values in parentheses represent robust z-statistics. * denotes that the variance of this parameter was constrained to its lower bound (i.e., equal to a value of 0.0) by the EQS program.

SCs. The baseline models differ for girls and boys only with regard to the relations among the first- and second-order SC facets. Specifically, for boys, Parent SC is linked to the second-order level of the SC hierarchy via only the Verbal-Academic SC facet; for girls, the link is via only the second-order Non-Core Academic SC construct. Thus, the hypothesized paths linking Parent SC with the second-order SC constructs, with the exception of one path for each gender, are not found to be significant parameters in the final best-fitting models. In addition, for both girls and boys, the hypothesized path linking the first-order General-Academic SC factor with the second-order Science-Academic SC factor is not significant. As a result, the first-order General-Academic SC facet defines only two of the three core second-order academic SC facets, namely, Mathematics- and Verbal-Academic SCs for both girls and boys.

A number of points with regard to the final multidimensional and hierarchical baseline models are worthy of note. First, the estimates associated with the parameters indicate that both the measurement and the structural portions of SC are well represented within the final models for both girls and boys. That is, the factor loadings associated with the measurement portion of SC reveal that the magnitude of the parameters are all high. For both girls and boys, the only factor loadings to fall below .70 are for the indicator variable AB14 (factor loadings = .598 and .681, for girls and boys, respectively) and AB26 (factor loadings = .600 and .668, for girls and boys, respectively). Likewise, the structural parameters representing the relations between the first- and second-order SCs are high for both girls and boys. Notably, the magnitude of the regression path associated with the first-order Writing in Language Arts SC is high for both girls and boys, and thus, indicates that it is well represented by its a priori hypothesized second-order Verbal-Academic SC facet.

A consistent pattern is evident across the final models for girls and boys for the regression paths linking the first- and second-order SC facets. Specifically, for both girls and boys, the first-order Religion SC facet (i.e., regression coefficients = .77 and .80, respectively) is more strongly defined by the second-order Non-core/Academic SC facet than is the first-order Art SC facet (i.e., .57 and .51, respectively). Likewise, for both girls and

boys, the Physical Ability (i.e., regression coefficients = .93 and 1.00, respectively) and Gym (i.e., regression coefficients = .81 for both samples) SC facets are more strongly defined by their second-order SC facet than are Physical Appearance and Peer SCs; in fact, the variance associated with the error in the prediction of Physical Ability SC is nonsignificant for girls and approximates zero for boys. Finally, for both girls and boys, the magnitude of the regression path associated with the first-order Parent SC facet is lower than that for the other first-order facets defined under its second-order SC facet.

The pattern of correlations among the second-order core academic SC facets is consistent across the sample of girls and boys. That is, under the final hierarchical model for both girls and boys, the second-order science SC facet is more highly correlated with Verbal-Academic SC (i.e., $r_s = .81$ and $.85$, respectively) than with Mathematics-Academic SC (i.e., $r_s = .65$ and $.70$, respectively). Similarly, the Verbal-Academic SC facet is more highly related to Science- than to Mathematics-Academic SC (i.e., $r_s = .57$ for both samples). In fact, the correlations among the Science-Academic, Verbal-Academic, and Non-Core Academic SC constructs (i.e., $r_s = .79$ to $.82$ for girls, and $.71$ to $.85$ for boys) are higher than are the correlations between these constructs and Mathematics-Academic SC (i.e., $r_s = .49$ to $.65$ for girls, and $.41$ to $.70$ for boys). Thus, of the second-order core academic SC constructs, Mathematics-Academic SC is least correlated with both the core and the non-core academic SC constructs. However, for both girls and boys, Science-Academic SC is more highly correlated with Mathematics-Academic SC than is Verbal-Academic SC. Taken together, the results reveal a pattern of relations in which SC constructs in science are more strongly related to mathematics than are verbal constructs, but more strongly related to verbal constructs than to mathematics constructs.

The correlations among the second-order academic SC factors are, as expected, higher than are the correlations between the academic and nonacademic (i.e., $r_s = .27$ to $.42$ for girls, and $.38$ to $.46$ for boys) second-order SC facets. Indeed, the magnitude of the correlations among the academic and nonacademic second-order SC facets are the lowest, and those among the second-order SCs representing science, verbal, and non-core academic

constructs are the highest (i.e., $r_s = .79$ to $.82$ for girls, and $.71$ to $.85$ for boys). We turn now to a review of results related to tests of the invariance of these baseline SC models across gender (as specified under Hypothesis 3).

Testing for the Invariance of Self-concept Measurement and Structure Across Gender

Covariance structure analysis of the invariance of the measurement and structure of SC were conducted using the baseline H4 model established for each gender. Thus, for the sample of girls, Model H4G² was utilized, and Model H4B³ represented the baseline model for the sample of boys. Given that EQS program limitations precludes the derivation of robust statistics with multigroup analyses, fit indices are expected to be lower than those obtained in single-group analyses. A total of 58 constraints were imposed in tests of the invariance of SC measurement and structure across gender (designated Model EQ-SC). Of these 58 parameters, 32 represent measurement constraints, and 26 are structural constraints. For the measurement parameters, two constraints were imposed for each of the 16 first-order SC facets; these paths represent the estimated factor loadings of the indicator variables on their matching first-order latent SC facet. Of the structural constraints, 16 represent the paths linking the first- and second-order SC facets that are common across the final SC models for girls and boys, and 10 represent the paths depicting the relations among the second-order latent SC constructs. Note that given that the path linking the first-order Parent SC facet with the second-order level of the SC hierarchy is not the same across gender, they could not be constrained equal.

Goodness-of-fit indices for Model EQ-SC indicate that this model provides an acceptable fit to the data ($\chi^2_{(2164)} = 3623.49$, CFI = .915). The LM-Test results for Model EQ-SC identifies five of the equivalence constraints imposed as untenable, namely, three measurement and two structural constraints. The measurement parameters identified as non-equivalent across gender correspond to the factor loading for the indicator variables ARM2, GYM2, and AB35 on their corresponding first-order SC factors (i.e., Arithmetic, Gym, and Physical Ability SCs, respectively). Note that two of the non-equivalent measurement parameters are single-item indicator variables. The two non-equivalent structural parameters

represent the paths linking the second-order Verbal-Academic SC construct with the first-order (a) General-Academic and (b) Listening in Language Arts SCs.

Re-estimation of Model EQ-SC with these five parameters allowed to vary across gender in Model EQ-SCa reveals that this less restrictive model provides an acceptable fit to the data ($\chi^2_{(2159)} = 3583.40$, CFI = .917). The deletion of these equality constraints contributes to an improvement in the fit of Model EQ-SCa over Model EQ-SC ($\Delta\chi^2_{(5)} = 40.09$, $\Delta\text{CFI} = .002$). The LM-Test results from Model EQ-SCa do not reveal any additional constraints as untenable. Thus, with the exception of three measurement and two structural parameters, the measurement and structure of SC is found to be equivalent.

The standardized estimates associated with the parameters found to vary across gender reveal a different pattern for the measurement than for the structural parameters. Specifically, the factor loadings for all three measurement parameters are lower for boys than for girls (i.e., factor loading are, respectively, .84 and .89 for AB35, .87 and .93 for ARM2, and .82 and .90 for GYM2). The regression paths associated with the non-equivalent structural parameters are, in contrast to the measurement parameters, higher for boys than for girls (i.e., coefficients in relation to Verbal-Academic SC are, respectively, .61 and .51 for General-Academic SC, and .96 and .91 for Listening in Language Arts). However, the magnitude of the gender differences is, in all instances, quite small (i.e., $\leq .10$). We turn now to a review of results related to the between-network component of the SC construct.

Between-network Analyses: Self-concept/Academic Achievement Relations

Interest in the relations among the latent SC and AA constructs is centered on two issues. The first focus is the extent to which the pattern of correlations among the latent constructs are consistent with a multidimensional and hierarchically structured SC, as specified under Hypothesis 4. The second issue is the invariance of SC/AA relations across gender, as specified under Hypothesis 5. Estimation of the latent correlations among the eight AA constructs and the 16 first-order SC facets was determined for the sample of girls and boys separately using the 16-factor first-order baseline model of SC established in tests of SC multidimensionality under Hypothesis 1. Likewise, the baseline hierarchical SC

models established under Hypothesis 2 for the sample of girls (Model H4G²) and boys (Model H4B³) were utilized to estimate (for girls and boys, separately) the latent correlations among the eight specific AA constructs and the second-order SC facets. We turn now to a review of the pattern of relations between the latent SC and AA constructs for each gender.

Preadolescent Girls: Testing for Self-concept/Academic Achievement Relations

Estimation of the correlations among the latent AA constructs and the first-order latent SC constructs for the sample of girls reveals correlations among three of the AA constructs to exceed acceptable values. Specifically, latent correlations among the Reading, Writing, and Speaking in Language Arts AA constructs exceed a value of one (i.e., $r_s = 1.02$ to 1.16). Given the shared content area among these three Language Arts AA constructs, their representation as a single construct is considered theoretically feasible. Accordingly, the indicator variables for these three AA constructs were respecified to represent the same latent Language Arts AA construct, with the construct labeled Skills in Language Arts, or Language Arts Skills AA. Thus, a total of six, rather than eight latent AA constructs are represented in analyses of SC/AA relations for girls.

Reestimation of the relations among the 16 first-order SC facets and the six AA constructs do not reveal unacceptably high levels of correspondence among the constructs. The latent factor correlations among the AA constructs and the first-order SC constructs for girls is presented in Table 16. In order to evaluate the tenability of Hypothesis 4, we turn now to a review of the pattern of latent SC/AA correlations. The pattern of latent correlations relevant to the multidimensionality of SC will be reviewed first, and is followed by a review of results relevant to the hierarchical structure of SC.

With regard to the differentiability of the SC and AA constructs, the latent factor correlations reveal that these constructs are intercorrelated to varying degrees, but do not reach unacceptably high levels (i.e., $r_s = -.13$ to $.72$, mean $r = .27$). In addition, the correlations among the AA constructs are generally higher (i.e., $r_s = .50$ to $.82$, mean $r = .70$) than are correlations among all the SC constructs (i.e., $r_s = .06$ to $.90$, mean $r = .42$), and, to

Table 16

Latent Factor Correlations Among Academic Achievement Constructs and First-order Self-concept Constructs for the Sample of Girls (N=220)

	Academic Achievement Constructs					
	ARM	SN	LLA	SkLA	ART	GYM
Academic Achievement Constructs						
SN	0.734	-				
LLA	0.652	0.757	-			
SkLA	0.793	0.821	0.823	-		
ART	0.668	0.691	0.671	0.820	-	
GYM	0.501	0.528	0.604	0.673	0.729	-
Core Academic Self-concepts						
ARM	0.724†	0.248	0.184	0.219	0.151*	0.125*
MT	0.642	0.349	0.268	0.305	0.147*	0.171
SN	0.479	0.722†	0.296	0.424	0.471	0.294
SST	0.450	0.593	0.304	0.375	0.472	0.265
LLA	0.346	0.435	0.420†	0.413	0.377	0.161*
SLA	0.311	0.430	0.259	0.502†	0.414	0.265
RLA	0.411	0.472	0.296	0.537†	0.376	0.177
WLA	0.499	0.523	0.299	0.516†	0.506	0.216
GA	0.681	0.566	0.389	0.531	0.347	0.096*
Non-Core Academic Self-concepts						
REL	0.216	0.303	0.251	0.248	0.197*	0.159
ART	0.078*	0.059*	-0.045*	0.046*	0.311†	0.030*
GYM	0.219	0.059*	-0.068*	-0.005*	0.003*	0.413†
Nonacademic Self-concepts						
AB	0.157*	0.073*	-0.125*	-0.057*	0.041*	0.307
AP	0.228	0.161*	-0.039*	0.117*	0.201	0.070*
PRT	0.118*	0.131*	0.198	0.059*	0.066*	-0.099*
PEER	0.291	0.227	0.047*	0.148*	0.225	0.151*

Note. ARM = Arithmetic; SN = Science; L/S/R/W/SkLA = Listening, Speaking, Reading, Writing, and Skills in Language Arts, respectively; MT = Measurement; SST = Social Studies; GA = General-Academic; REL = Religion; AB = Physical Ability; AP = Physical Appearance; PRT = Parent.

† denotes corresponding self-concept/academic achievement relations.

* denotes a nonsignificant parameter (at $p < .05$).

a modest extent, only the subject-specific core SC constructs (i.e., $r_s = .34$ to $.90$, mean $r = .62$).

With regard to the content-specific multidimensional pattern of SC/AA relations, the AA constructs are, with one exception, more highly correlated with their corresponding first-order SC facet than with the other specific academic SCs (i.e., vertical comparisons within the table). That is, as expected, AA constructs in Arithmetic, Science, Listening in Language Arts, and Gym are more highly correlated with their corresponding SCs (i.e., $r_s = .72, .72, .42$, and $.43$, respectively) than with other specific academic SC facets. Similarly, the Language Arts Skills AA construct is more highly correlated with its three corresponding SC facets (i.e., $r_s = .50, .52$, and $.54$ for Speaking, Writing, and Reading, respectively) than with other specific academic SC facets. Notably, consistent with the distinction between the Listening and Skills in Language Arts AA constructs, Listening in Language Arts AA is more highly correlated with its matching SC facet (i.e., $r = .42$) than with the three SC facets corresponding to Skills in Language Arts (i.e., $r_s = .30, .30$, and $.26$ for Writing, Reading, and Speaking in Language Arts SCs, respectively).

Unexpectedly, however, Art AA is more highly correlated with the specific science and Language Arts SC facets (i.e., $r_s = .38$ to $.51$, mean $r = .44$) than with Art SC (i.e., $r = .31$). In order to more clearly ascertain the nature of this unexpected finding, zero-order correlations were computed among the grades (as obtained from report cards), and the self-reported grades for the sixteen indicator variables representing the AA constructs. Results indicate that self-reported grades in Art are, as expected, more highly correlated with Art grades than with grades in other academic areas. In contrast, grades in Art are equally, or more highly correlated with self-reported grades in most other academic areas than with self-reported Art grades; only self-reported Gym and Arithmetic grades are less highly correlated. Thus, the unexpected pattern of correlations for the Art AA construct appears to be related to report card grades, rather than self-reported grades.

The content-specific pattern of latent SC/AA relations examined on the basis of the SC constructs (i.e., horizontal comparisons within the table) reveals a slightly different

pattern than do the comparisons based on the AA constructs. That is, as expected, latent SC constructs in Arithmetic, Science, Art, Gym, and both Speaking and Reading in Language Arts are more highly correlated with their corresponding AA constructs than with non-corresponding AA constructs. However, Listening in Language Arts SC is essentially equally correlated with AA constructs in Skills in Language Arts and Science (i.e., $r_s = .41$ and $.44$, respectively), as with its corresponding AA construct (i.e., $r = .42$); correlations with Arithmetic and Art AA constructs are also not substantially lower (i.e., $r_s = .35$ and $.38$, respectively). Writing in Language Arts SC is essentially equally correlated with its corresponding Skills in Language Arts AA construct (i.e., $r = .52$), as with Arithmetic, Science, and Art AA constructs (i.e., $r_s = .50$, $.52$, and $.51$, respectively); it is, however, less correlated with Listening in Language Arts and Gym AA constructs (i.e., $r_s = .30$ and $.22$, respectively). Thus, Listening in Language Arts SC is essentially equally, or only slightly less correlated with all AA constructs except Gym AA as with its corresponding AA construct, and Writing in Language Arts SC is essentially equally correlated with all AA constructs except Listening in Language Arts and Gym AA constructs.

In summary, with the exception of the Art dimension, the latent correlations are consistent with a pattern in which AA constructs in particular academic areas are more highly correlated with their first-order SC facets than with non-corresponding SC facets. The latent SC/AA correlations examined from the reverse perspective are consistent with a pattern in which subject-specific SC constructs are more highly correlated with their corresponding than with non-corresponding AA factors for all constructs except Writing and Listening in Language Arts SCs.

A distinction among core and non-core academic areas (Marsh, 1990d) is also useful in examining the pattern of SC/AA relations. That is, for the core AA constructs, a pattern is evident in which these constructs are most correlated with core academic SCs, and least or uncorrelated with nonacademic and non-core academic SCs. Notably, Science AA is less correlated with the specific mathematics SCs (i.e., $r_s = .25$ and $.35$) than with the four specific Language Arts SCs (i.e., $r_s = .43$ to $.52$). One exception to the core/non-core pattern

noted above, however, is that the Science, and the Listening and Skills in Language Arts AA constructs, are equally, or more highly correlated with Religion SC (i.e., $r_s = .25$ to $.30$) than with Arithmetic SC (i.e., $r_s = .19$ to $.25$). Consistent with this pattern, the AA constructs, with the exception of corresponding SC/AA relations, are generally more highly correlated with Religion SC than with the other non-core SC facets. Thus, non-corresponding SC/AA relations are typically of an approximately equal magnitude for Arithmetic and Religion SC, and in addition, Religion SC shares a stronger pattern of relations with non-corresponding AA constructs than do the non-core Art and Gym SC facets. The non-core academic Gym AA facet is, as expected, most correlated with Gym SC (i.e., $r = .41$), followed by Physical Ability SC (i.e., $r = .31$). We turn now to a review of results relevant to an hierarchical pattern of SC/AA relations.

Evaluation of SC/AA relations consistent with an hierarchical pattern involves comparison of the magnitude of the latent correlations between corresponding AA and first-order SC constructs against those between corresponding AA and second-order SCs (i.e., a vertical comparison). The latent factor correlations among each of the AA facets and the second-order SC facets for the sample of girls are presented in Table 17. As expected, Art, Gym, and Listening in Language Arts AA constructs are more strongly correlated with their corresponding first-order SCs (i.e., $r_s = .31$, $.41$, and $.42$, respectively) than with their corresponding second-order SCs (i.e., $r_s = .23$, $.28$, and $.32$, respectively). However, Science, Arithmetic and Language Arts Skills AA constructs are equally, or more strongly correlated with their second-order SCs than with their first-order SCs (i.e., r_s with their first- and second-order SCs, respectively, are $.72$ and $.73$ for Science, $.72$ and $.79$ for Arithmetic, and, for Language Arts Skills, are $.50$ to $.54$ compared to $.56$). Thus, the expected pattern of correlations is evident for Art, Gym, and Listening in Language Arts constructs but not for Science, Arithmetic, and the remaining Language Arts factors.

Table 17

Latent Factor Correlations Among Academic Achievement Constructs and Second-order Self-concept Constructs for the Sample of Girls (N=220)

Academic Achievement	Second-order Self-concept Factors				
	Nonacademic/ Physical	Mathematics- Academic	Science- Academic	Verbal- Academic	Non-core Academic
ARM	0.241	0.790†	0.515	0.445	0.241
SN	0.116*	0.354	0.733†	0.524	0.321
LLA	-0.104*	0.257	0.315	0.320†	0.220
SkLA	-0.001*	0.309	0.446	0.555†	0.251
ART	0.067*	0.164*	0.448	0.401	0.227†
GYM	0.279†	0.098*	0.263	0.199	0.101*

Note. ARM = Arithmetic; SN = Science; L/SkLA = Listening and Skills in Language Arts, respectively.

† denotes corresponding self-concept/academic achievement relations.

* denotes a nonsignificant parameter (at $p < .05$).

Preadolescent Boys: Testing for Self-concept/Academic Achievement Relations

The estimation of the correlations among the latent AA constructs and the first-order latent SC constructs for the sample of boys, as with the results for girls, reveal that latent correlations among the Reading, Writing, and Speaking in Language Arts AA constructs exceed acceptable values. In fact, the latent correlations among these Language Arts AA constructs are slightly higher for boys (i.e., $r_s = 1.05$ to 1.21) than they are for girls. In addition, the correlations between the Listening in Language Arts AA facet and the remaining Language Arts AA facets are unacceptably high (i.e., $r_s = .96$ to 1.05), as are some of the correlations between the Science and the Language Arts AA constructs (i.e., $r_s = .87$ to 1.06). Given that representing all of these latent AA constructs as one construct is not theoretically tenable, only the Reading, Writing, and Speaking in Language Arts latent AA

constructs, as with the sample of girls, were collapsed and represented as a single Skills in Language Arts latent AA construct.

Results from this re-estimated model yields more acceptable, albeit high, correlations between the Skills and Listening in Language Arts AA constructs (i.e., $r = .93$), and between the Science AA construct and both the Skills and Listening in Language Arts AA constructs (i.e., $r_s = .92$ and $.95$, respectively). Thus, the correlations among these AA constructs are higher for boys than for girls. Despite the high level of correspondence among these three latent AA facets, the pattern of correlations between these constructs and the SC constructs (as described below), in addition to the lack of theoretical justification for doing so, argues against their being represented as one AA construct. The latent factor correlations among the 16 first-order SC facets and the six AA constructs for the sample of boys are presented in Table 18.

As Table 18 reveals, the Science AA is substantially more highly correlated with Science SC (i.e., $r = .70$) than with any of the Language Arts SCs (i.e., $r_s = .30$ to $.51$). Similarly, the Listening in Language Arts AA construct is most correlated with Listening in Language Arts SC (i.e., $r = .37$), and generally less correlated with the three other Language Arts SC facets, and the science SC facets (i.e., $r_s = .25$ to $.35$). Finally, the Skills in Language Arts AA facet is generally most correlated with the three corresponding Language Arts SC facets (i.e., $r_s = .43$ to $.49$), and more correlated with the Listening in Language Arts SC facet (i.e., $r = .43$) than with the Science SC facet (i.e., $r = .35$). Thus, although these AA constructs are highly correlated, corresponding SC/AA relations are generally higher than are non-corresponding SC/AA constructs, particularly for science. The pattern of SC/AA relations, therefore, rather than the relations among these AA constructs, mitigates their representation as a unitary AA construct. Accordingly, the Skills in Language Arts, Listening in Language Arts, and Science AA constructs are represented separately.

With regard to the differentiability of SC and AA constructs for the sample of boys, the latent SC/AA factor correlations fall within acceptable bounds (i.e., $r_s = -.09$ to $.79$,

Table 18

Latent Factor Correlations Among Academic Achievement Constructs and First-order Self-concept Constructs for the Sample of Boys (N=205)

	Academic Achievement Constructs					
	ARM	SN	LLA	SkLA	ART	GYM
Academic Achievement Constructs						
SN	0.779	-				
LLA	0.728	0.947				
SkLA	0.751	0.918	0.931	-		
ART	0.385	0.383	0.440	0.409	-	
GYM	0.551	0.391	0.661	0.519	0.433	-
Core Academic Self-concepts						
ARM	0.793†	0.329	0.201	0.244	0.152*	0.139*
MT	0.557	0.365	0.231	0.130*	0.084*	0.087*
SN	0.367	0.695†	0.310	0.345	0.206	0.066*
SST	0.374	0.514	0.285	0.379	0.245	0.187*
LLA	0.281	0.441	0.372†	0.432	0.143*	0.078*
SLA	0.222	0.294	0.251	0.428†	0.105*	0.049*
RLA	0.286	0.511	0.314	0.490†	0.163*	0.002*
WLA	0.347	0.444	0.354	0.461†	0.199	0.113*
GA	0.429	0.482	0.337	0.414	0.193*	0.112*
Non-Core Academic Self-concepts						
REL	-0.022*	0.121*	0.091*	0.121*	0.258	-0.043*
ART	0.033*	0.068*	-0.068*	-0.009*	0.650†	-0.055*
GYM	0.127*	-0.047*	0.074*	-0.012*	0.149*	0.556†
Nonacademic Self-concepts						
AB	0.169*	-0.091*	-0.027*	-0.054*	0.044*	0.448
AP	0.073*	0.012*	-0.043*	0.081*	0.071*	0.234
PRT	-0.025*	0.241	0.131*	0.143*	0.040*	0.034*
PEER	0.139*	0.035*	0.061*	0.180	0.115*	0.336

Note. ARM = Arithmetic; SN = Science; L/S/R/W/SkLA = Listening, Speaking, Reading, Writing, and Skills in Language Arts, respectively; MT = Measurement; SST = Social Studies; GA = General-Academic; REL = Religion; AB = Physical Ability; AP = Physical Appearance; PRT = Parent.

† denotes corresponding self-concept/academic achievement relations.

* denotes a nonsignificant parameter (at $p < .05$).

mean $r = .21$). Consistent with the results for girls, the correlations among the AA constructs are higher (i.e., $r_s = .38$ to $.95$, mean $r = .62$) than are correlations among all the SC constructs (i.e., $r_s = .13$ to $.91$, mean $r = .46$), and correlations among the specific core SC constructs (i.e., $r_s = .38$ to $.91$, mean $r = .63$) are higher than among all SC constructs. However, the magnitude of the correlations among the core academic SC constructs do not differ from those among the AA constructs.

With regard to the content-specific multidimensional pattern of SC/AA relations, the AA constructs are typically more highly correlated with their corresponding first-order SC than with non-corresponding academic SCs (i.e., vertical comparisons within the table). In particular, AA constructs in Arithmetic, Science, Art, and Gym are more highly correlated with their corresponding SC (i.e., $r_s = .79$, $.70$, $.65$, and $.56$, respectively) than with other specific academic SCs. The Skills in Language Arts AA construct is, with one exception, more highly correlated with its three corresponding SC facets ($r_s = .43$, $.46$, and $.49$ for Speaking, Writing, and Reading, respectively) than with other specific academic SCs; it is essentially equally correlated with Listening in Language Arts SC (i.e., $r = .43$). Similarly, with one exception, the Listening in Language Arts AA construct is most correlated with its corresponding SC facet (i.e., $r = .37$); it is essentially equally correlated with Writing in Language Arts SC (i.e., $r = .35$). Thus, the corresponding SC/AA correlations are higher than non-corresponding SC/AA relations for Arithmetic, Science, Gym, and, in contrast to the results for girls, Art AA constructs. The hypothesized pattern of results is not evident for the Skills in Language Arts AA constructs in relation to Listening in Language Arts SC, and for Listening in Language Arts AA in relation to Writing in Language Arts SC.

An examination of the content-specific pattern of latent SC/AA relations on the basis of the SC constructs (i.e., horizontal comparisons within the table), as compared to those based on the AA constructs, reveals a pattern of results for the sample of boys which is similar to that obtained for girls. That is, latent SC constructs in Arithmetic, Science, Gym, Art, and Speaking in Language Arts are more highly correlated with their corresponding AA constructs than with non-corresponding AA constructs. Although not surprising given the

high level of correspondence among Science and Language Arts AA constructs, Listening in Language Arts SC is more highly correlated with AA constructs in Skills in Language Arts and Science (i.e., $r_s = .43$ and $.44$, respectively) than with its corresponding AA construct (i.e., $r = .37$). Writing in Language Arts SC is essentially equally correlated with its corresponding Skills in Language Arts AA construct (i.e., $r = .46$) as with Science AA (i.e., $r = .44$); the same pattern is true of Reading in Language Arts SC. However, consistent with the differentiation between Listening and Skills in Language Arts AA constructs, the three SCs associated with Language Arts skills are less correlated with Listening in Language Arts AA (i.e., $r_s = .25$ to $.35$) than with their corresponding AA construct (i.e., $r_s = .42$ to $.49$).

In summary, the latent factor correlations are generally consistent with a pattern in which corresponding SC/AA relations are higher than non-corresponding relations. However, the pattern of content-specific SC/AA relations within the Language Arts areas is, as for girls, less consistent with expectations. That is, the Skills and Listening in Language Arts AA constructs do not follow the hypothesized pattern in relation to, respectively, Listening and Writing in Language Arts SCs. The pattern of latent SC/AA relations, examined on the basis of the SC constructs, is as expected for all but the Writing, Reading, and Listening in Language Arts SC facets. In particular, Writing and Reading in Language Arts SCs are equally correlated with their corresponding AA construct as with Science AA; they are, however, less correlated with Listening in Language Arts AA. This pattern of results further substantiates the separation between the Listening and Skills in Language Arts constructs. Listening in Language Arts SC is, however, equally, or more highly correlated with Science and Language Arts Skills AA constructs as with its corresponding AA construct.

As with the sample of girls, the distinction among core and non-core academic areas (Marsh, 1990d) is useful in examining the pattern of SC/AA relations. The core AA constructs are most correlated with core academic SCs, and least, or uncorrelated with nonacademic and non-core academic SCs. The only exception to this pattern for the sample of boys is that the Language Arts Skills AA construct is not significantly correlated with

Measurement SC. As with the pattern of correlations for the sample of girls, of the core academic SC constructs, the mathematics SC constructs are generally the least correlated with Language Arts AA constructs. Similar to the results for girls, Science AA is less correlated with the specific mathematics SCs (i.e., $r_s = .33$ and $.37$) than with three of the four specific Language Arts SCs (i.e., $r_s = .44$ to $.51$). In contrast to the results for girls, however, Religion SC is not significantly correlated with any of the core academic areas. The non-core academic Gym AA facet is, as expected and consistent with the results for girls, most correlated with Gym SC (i.e., $r = .56$), followed by Physical Ability SC (i.e., $r = .45$). We turn now to a review of SC/AA relations relevant to an hierarchical pattern.

The hierarchical (i.e., vertical) comparisons of the latent correlations between corresponding AA and first-order SC constructs against those between corresponding AA and second-order SCs reveals a pattern similar to that obtained for girls. The latent factor correlations among the AA facets and the second-order SC facets for the sample of boys are presented in Table 19. Gym AA is moderately more strongly correlated with its corresponding first-order SC (i.e., $r = .56$) than with its corresponding second-order SC (i.e., $r = .48$). The remaining AA constructs are, with few exceptions, essentially equally, or more highly correlated with their corresponding first-order, as with their second-order SC facets. Listening in Language Arts AA is almost equally correlated with its corresponding first- and second-order SC facets (i.e., $r_s = .37$ and $.34$, respectively). Arithmetic, Science, and Art AA constructs are generally equally, or more strongly correlated with their second-order SCs than with their first-order SCs (i.e., r_s with their first- and second-order SCs, respectively, are $.79$ and $.82$ for Arithmetic, $.70$ and $.70$ for Science, and $.65$ and $.73$ for Art). The Skills in Language Arts AA construct is equally correlated with its corresponding Verbal-Academic SC facet (i.e., $r = .48$) as with its first-order Writing and Reading in Language Arts SC facets (i.e., $r_s = .46$ and $.49$), although it is slightly less correlated with its corresponding Speaking in Language Arts SC facet (i.e., $r = .43$). Thus, an hierarchical pattern of SC/AA relations in which AA constructs are more highly correlated with corresponding first-order SCs than with second-order SCs is evident for Gym but not for the remaining academic constructs.

Table 19

Latent Factor Correlations Among Academic Achievement Constructs and Second-order Self-concept Constructs for the Sample of Boys (N=205)

Academic Achievement	Second-order Self-concept Factors				
	Nonacademic/ Physical	Mathematics-Academic	Science-Academic	Verbal-Academic	Non-core Academic
ARM	0.155*	0.817†	0.416	0.291	0.009*
SN	-0.057*	0.398	0.704†	0.467	0.143*
LLA	0.007*	0.244	0.333	0.347†	-0.004*
SkLA	-0.003*	0.244	0.398	0.475†	0.070*
ART	0.094*	0.155*	0.238	0.162*	0.734†
GYM	0.476†	0.140*	0.121*	0.071*	-0.101*

Note. ARM = Arithmetic; SN = Science; L/SkLA = Listening and Skills in Language Arts, respectively.

† denotes corresponding self-concept/academic achievement relations.

* denotes a nonsignificant parameter (at $p < .05$).

Testing for the Multidimensional and Hierarchical Structure of Self-concept/Academic Achievement Relations for Preadolescent Girls and Boys: A Summary

The pattern of results for the sample of boys is generally consistent with that for the sample of girls. The most striking difference between the results obtained for girls and boys occurs for the Art AA construct. The content-specificity of this construct is more consistent with hypotheses for boys than for girls. Additional analyses of the AA construct for the sample of girls, however, suggests that the results obtained are related to a lack of content-specificity for report card, as opposed to self-reported grades.

With regard to the multidimensionality and content-specificity of SC/AA relations, results are generally consistent with a pattern of relations in which AA and first-order SC constructs are higher for corresponding than for non-corresponding content areas. Exceptions to this pattern are evident for SC and AA constructs in the Language Arts

content area for both girls and boys, and for Art AA for girls. That is, Language Arts SC and AA constructs are unexpectedly more highly correlated with some of the non-corresponding core academic achievement and SC constructs, respectively, and most particularly, other Language Arts and Science constructs. A similar pattern is evident for girls with regard to Art AA and its relation to SC constructs in science and Language Arts areas. In contrast to multidimensional SC/AA relations, an hierarchical pattern of SC/AA relations is not consistently demonstrated for either girls or boys.

Testing for the Invariance of Self-concept/Academic Achievement Relations Across Gender

The final set of analyses conducted involve tests of the invariance of the SC/AA relations across gender, as specified under Hypothesis 5. As with the estimation of SC/AA relations, the final baseline hierarchical models of SC established for each gender under Hypothesis 2 were utilized in testing for the gender invariance of SC/AA relations. That is, Model H4G² was utilized for the sample of girls, and Model H4B³ was utilized for the sample of boys. In addition, consistent with analyses under Hypothesis 4, the Language Arts AA constructs are represented as two, rather than four factors for both girls and boys. As noted with the tests of invariance conducted in relation to Hypothesis 3, only the uncorrected chi-square statistic and CFI index are currently available with the EQS program for multigroup comparisons; accordingly, the fit indices are expected to be lower than if robust indices were available.

The model of invariant SC/AA relations across gender includes both a measurement and a structural component. The invariance of the measurement portion of the SC construct was specified consistent with that established under Hypothesis 3 in tests of the invariance of SC measurement and structure across gender. As with Hypothesis 4, tests of the invariance of SC/AA relations for first-order SCs were conducted separately from those for second-order SCs. Tests for the gender invariance of the relations between the AA constructs and second-order SC facets were specified consistent with the results obtained under Hypothesis 3 with regard to the invariance of SC measurement and structure.

With regard to the invariance of relations between AA constructs and first-order SC facets, a total of 125 parameters were constrained equal (designated Model EQ-SC1AA). Of these parameters, 29 represent the measurement constraints found to be invariant across gender under Hypothesis 3. A total of 96 structural constraints were imposed, all of which represent the relations among the AA and first-order SC constructs; that is, six constraints were imposed for each of the sixteen first-order SC factors in relation to the AA constructs.

CFA of Model EQ-SC1AA indicates that this model provides a less than adequate fit to the data across girls and boys ($\chi^2_{(3567)} = 6190.72$, CFI = .882). The LM-Test identifies six of the 125 parameters constrained equal across gender to be significantly different. Of these, one represents a measurement parameter, namely, Item 2 of the Writing in Language Arts ASDQ-I instrument; notably, this is a single-item indicator variable. The remaining five parameters identified as significantly different across gender are structural parameters. Not surprisingly, two of these involve the Art AA construct. In particular, the factor covariances linking Art AA with the Art and Speaking in Language Arts SC factors are identified as non-equivalent across girls and boys. This finding is consistent with the differing pattern of results for the Art AA construct across gender identified earlier in the review of the results relevant to SC/AA relations for each gender. The relation between the first-order Measurement SC facet and the Language Arts Skills AA construct is also found to vary across gender. The two remaining parameters identified as non-equivalent involve the relations between Arithmetic AA and Physical Ability SC, and between Gym AA and General-Academic SC. Note that the Physical Ability SC scale and indicator variables are generally non-normally distributed for the sample of boys, but not for the sample of girls.

The six parameters identified as non-equivalent across girls and boys were freely estimated in Model EQ-SC1AA². This less restrictive model contributes to a statistically significant improvement in model fit, although the practical change is negligible ($\Delta\chi^2_{(6)} = 60.64$, $\Delta\text{CFI} = .002$). Although this fit of the model remains less than adequate ($\chi^2_{(3561)} = 6130.08$, CFI = .884), this is not unexpected given that these values are not based on a correction due to non-normality of the data. An additional constraint is identified as

untenable under the LM-Test for this model (i.e., between Religion SC and Art AA), however, this parameter is not incorporated into a subsequent model given that it was not identified under multivariate conditions with Model EQ-SC1AA.

The standardized estimates associated with the parameters found to vary across gender reveal that, with the exception of parameters involving the Art AA construct, the gender differences associated with these parameters are of a small magnitude (i.e., $\leq .13$). With regard to the Art AA construct, results reveal that for boys, as compared with girls, Art AA is more highly correlated with Art SC (i.e., $r_s = .65$ and $.32$, respectively) but less highly correlated with Speaking in Language Arts SC (i.e., $r_s = .08$ and $.37$, respectively, with the parameter for boys being nonsignificant). The remaining parameters found to be higher for girls than for boys include the factor loading associated with the Writing in Language Arts SC indicator variable (i.e., $.84$ and $.81$, respectively), and the correlation between Language Arts Skills AA and Measurement SC (i.e., $r_s = .28$ and $.15$, respectively). The remaining parameters identified as non-equivalent are higher for boys than for girls in that relations are of a modest and significant positive value for boys but are nonsignificant, albeit positive for girls. These parameters represent the correlation between Physical Ability SC and Arithmetic AA (i.e., $r_s = .25$ and $.12$ for boys and girls, respectively), and between General-Academic SC and Gym AA (i.e., $r_s = .15$ and $.02$ for boys and girls, respectively). The weaker correlation for Art SC/AA relations and the stronger correlation for Art SC and Speaking in Language Arts AA relations obtained for the sample of boys, as compared with girls, is consistent with the differing pattern of SC/AA relations observed for the Art AA construct in analyses of the samples of girls and of boys separately. In sum, with the exception of the two parameters involving Art AA, the gender differences for non-equivalent parameters are all of a small magnitude.

With regard to the invariance of relations between AA constructs and second-order SC facets, a total of 83 parameters were constrained equal (designated Model EQ-SC2AA). Of these parameters, 29 represent the measurement constraints found to be invariant across gender under Hypothesis 3; the measurement parameter identified under Model EQ-SC1AA

was not incorporated in order to determine whether the finding of non-equivalence of this single-item indicator variable across gender would be duplicated. A total of 54 structural equality constraints were imposed, 24 of which represent the SC structural parameters found to be invariant across gender under Hypothesis 3. The remaining 30 parameters represent the relations among the AA and second-order SC constructs; that is, six constraints, one for each of the AA constructs, were imposed for each of the five second-order SC factors.

CFA of Model EQ-SC2AA indicates that this model provides a less than adequate fit to the data across the samples of girls and boys ($\chi^2_{(3843)} = 6909.24$, CFI = .862). This is not unexpected, however, given that these are non-robust indices, and that the AA constructs are not permitted to covary with the first-order SCs. Examination of the LM-Test results reveals chi-square values related to three of the 83 parameters constrained equal across gender to be statistically significant. Of these, one represents a measurement parameter, namely, the indicator variable representing the pairing of Items 4 and 5 of the ASDQ-I Science scale. Notably, the single-item indicator variable found to be non-equivalent in tests of SC/AA relations for first-order SC facets is not identified for Model EQ-SC2AA. The LM-Test also indicates that the parameter representing the relation between Science AA and Nonacademic/Physical SC is not equivalent across gender. In addition, the parameter representing the association between Art AA and Verbal-Academic SC is identified as non-equivalent under Model EQ-SC2AA.

Model EQ-SC2AA was reestimated as Model EQ-SC2AA² with these three parameters freely estimated across the girl and boy samples. The removal of these equality constraints contributes to a statistically significant improvement in the fit of the model, although there is no change in the value of the CFI ($\Delta\chi^2_{(3)} = 17.00$, $\Delta\text{CFI} = .000$); no additional constraints are identified as untenable under the LM-Test for this model. Not unexpectedly, this model nevertheless yields less than adequate fit indices ($\chi^2_{(3840)} = 6892.24$, CFI = .862). Examination of the standardized estimates for this model reveals that the non-equivalent science measurement parameter is higher for boys than for girls (i.e., factor loadings are .90 and .87, respectively). For both structural parameters, the correlations

among the constructs are higher for girls than for boys (i.e., correlations are, respectively, .42 and .13 for Art, with the latter being nonsignificant; and .11 and -.06 for Science, with both parameters being nonsignificant). For the science construct, note that the gender difference for the measurement parameter is of a quite small magnitude, and that the structural parameters are nonsignificant for both girls and boys. Thus, consistent with the results obtained under Model EQ-SC1AA, the only parameter with a gender difference that is of a substantial magnitude involves the Art AA construct.

In summary, tests for the invariance of SC/AA relations suggest few differences between boys and girls in the relations between SC and AA constructs. Moreover, when differences are found, these differences tend to be of a small magnitude. The only gender differences in SC/AA relations that are of a substantial magnitude involve the Art AA construct and are consistent with the unexpected pattern of results found with regard to Art AA in analyses conducted separately for girls and boys. Specifically, tests of invariant SC/AA relations across gender indicate that for girls, as compared with boys, Art AA is less highly correlated with its corresponding Art SC but more highly correlated with both the first-order Speaking in Language Arts SC facet and its corresponding second-order Verbal-Academic SC.

DISCUSSION

The purpose of the present study was, in broad terms, to examine the construct validity of an expansion of the Marsh/Shavelson model of SC, and of SC/AA relations for preadolescents within and across gender. Specifically, the purposes included (1) testing for the multidimensional structure of nonacademic SC and a broad set of academic SCs separately for girls and boys, (2) testing competing hierarchical representations of these multidimensional SC facets separately for girls and boys, (3) testing for the invariance of the multidimensional measurement and hierarchical structure of the SC construct across gender, (4) examining the multidimensional and hierarchical pattern of SC/AA relations separately for girls and boys, and (5) testing for the invariance of SC/AA relations across gender.

As with the review of the results, the findings will be discussed under the rubrics of preliminary, within-network, and between-network results. Likewise, gender differences in relation to within-network analyses will be discussed prior to those relevant to between-network analyses. Finally, the strengths and limitations of the present study, and the research and clinical implications will be discussed along with suggestions for future research.

Preliminary Analyses

Preliminary analyses highlight several issues with regard to the SPPC and SDQ instruments. With regard to specific indicator variables, the descriptive statistics for the item pair derived from Items 2 and 5 of the SDQ-I Parent scale suggest the possibility that either one, and likely both of the items representing this indicator variable did not provide a high level of differentiation for the Parent SC construct for both girls and boys. This finding renders the value of the item(s) comprising this indicator variable in differentiating between children of varying levels of Parent SC questionable.

A similar phenomenon appears to be occurring for boys for the Gym SC scale, particularly for the (A)SDQ-I, and to a lesser extent, the (A)SDQ-I Physical Ability scale. The relative similarity of the means and standard deviations of the Gym SC indicator variables within each of the SC instruments, however, suggest that these results are more a reflection of the nature of the Gym SC construct for preadolescents boys than a lack of scale

cohesion among the indicator variables by which it is represented. That is, the results suggest that Gym SC is generally highly salient for preadolescent boys. This interpretation is consistent with a developmental phenomenon reported in the SC literature in which younger children have higher SCs than do older children (e.g., Marsh, 1989a, 1990a), as well as with empirical research indicating that boys have modestly higher SCs in areas related to physical competence than do girls (e.g., Marsh, 1992b; Marsh, Barnes, et al., 1984; Marsh, Smith, & Barnes, 1985). However, for the SDQ-I Physical Ability scale, the consistency of the values for the descriptive statistics across the indicator variables are lower than for the SDQ-I Gym SC scale, and thus, suggest that the results for this scale reflect a lower level of scale cohesion for the indicator variable by which it is represented.

Taken together, results relevant to the validity of the SC instruments are consistent in suggesting that the (A)SDQ-I instruments offer a more psychometrically sound measure of preadolescent girls' and boys' SC than does the SPPC instrument. That is, findings support the internal consistency of the (A)SDQ-I and SPPC instrument scales for both girls and boys, however, the internal consistency of the SPPC scales are, with the exception of the SDQ-I Physical Ability SC scale, less favorable than those for the (A)SDQ-I. These results indicate that the SPPC scales are less cohesive than are the (A)SDQ-I scales. Nevertheless, the results for the SDQ-I Physical Ability scale suggest the need to more fully examine the measurement portion of the Physical Ability SC construct, and in particular, the SDQ-I Physical Ability SC scale.

Separate CFAs of the (A)SDQ-I and SPPC instruments support the factorial validity of both of these instruments as measures of multidimensional preadolescent SCs for both girls and boys. Combined analyses of these SC instruments, however, suggest the presence of a moderate method effect for the SPPC instrument for both girls and boys, albeit the method effect appears to be slightly stronger for boys. These results are consistent with the moderately lower internal consistencies obtained for the SPPC, as compared to the (A)SDQ-I scales, as well as with a pattern in which the internal consistency of the SPPC scales are generally lower for boys than for girls.

The pattern of results relevant to the validity of the SC instruments is hypothesized to reflect a moderate “alternating direction” method effect specific to the SPPC instrument. The method effect is posited to be due to the complexity of the SPPC response format, and in particular, its alternating direction. That is, it is hypothesized that preadolescent children, and boys to a slightly greater extent than girls, were not always able to shift the direction of their responses in a manner consistent with their SCs. As a result, the students had not always been able to respond appropriately to the SPPC items, or stated differently, had not consistently made a successful transition to questions requiring a response in an alternate direction. Hence, the children’s responses across SPPC scale items were less internally consistent, as well as, less consistent with their SC, and therefore, are less valid than are responses to the (A)SDQ-I instruments. Thus, the validity of the SPPC instrument as a consistent and accurate measure of preadolescent SC is believed to be compromised by its response format, and in particular, its alternating response direction.

Although the hypothesized alternating direction method effect appeared to be moderate, as did the difference in the magnitude of this effect across gender, inclusion of the SPPC potentially introduced a confound that would have rendered the interpretability of the results more complex. That is, the interpretation of the results in relation to the nature of the SC construct, and in particular, gender differences in the SC construct, was potentially confounded with the hypothesized SPPC method effect. Thus, inclusion of the SPPC instrument in tests of the measurement and structure of a multidimensional and hierarchical SC construct would have diminished the clarity and validity of SC construct interpretations both within and across gender. Accordingly, although the utilization of multiple independent SC instrument is generally considered to enhance construct interpretations, this advantage was outweighed by the potential confound introduced by the SPPC method effect. Hence, the exclusion of the SPPC from further analyses (i.e., tests of the hypotheses).

The interpretation of the pattern of results relevant to the SPPC instrument within the context of a method effect is consistent with both theoretical concerns and empirical research. That is, a number of researchers have expressed concern with the complexity and

appropriateness of the SPPC response format for younger and less cognitively-able students (e.g., Keith & Bracken, 1996; Marsh, 1990b). Similarly, Marsh hypothesized, and a number of researchers have demonstrated empirical support for a “negative item” response bias in which children’s responses to negatively worded items are less scale consistent than are their responses to positively worded items (e.g., Benson & Hocevar, 1985; Marsh, 1986b). Finally, the differing results for the SC instruments are consistent with earlier empirical research with preadolescents in which a moderate method effect associated with the SPPC instrument was found in the absence of a similar effect for the SDQ-I. On the basis of the present findings, the negative item response bias hypothesized by Marsh (1986b) was extended, refined and labeled an alternating direction response bias in order to explain the results obtained for the SPPC instrument.

The absence of the SPPC instrument method effect for the single-instrument confirmatory factor analyses is believed to be due to two factors: (a) that the SPPC items were collapsed across response directions, and (b) that the method effect was both of a moderate magnitude and generalized across the SPPC indicator variables. Collapsing items across response directions likely attenuated the SPPC method effect. This attenuation of an already modest method effect, combined with its generality across indicator variables, is believed to underlie the absence of large and distinct areas of misfit, or large overall misfit in tests of the single-instrument SPPC model. Logically, however, this misfit was magnified when the SPPC and (A)SDQ-I instruments were both included in analyses. Consequently, combined analyses of the SC instruments more clearly reveal the SPPC method effect. This effect was apparent, as compared to the single-instrument SC models, through (a) a dramatic drop in the fit of the dual-instrument models, (b) the identification of correlated errors primarily or exclusively among the SPPC indicator variables, and (c) an increase in the magnitude of the misfit associated with ill-fitting parameters. Although the magnitude of the misfit (i.e., the chi-square values) associated with ill-fitting parameters increased for the dual-instrument SC models, they nevertheless remained relatively modest and non-distinct. This pattern of results is similarly believed to reflect the modest magnitude of the SPPC

method effect, and its generality across indicator variables. These results highlight the importance of utilizing multiple independent instruments in construct validity research.

The differences in the descriptive statistics for the Gym and Physical Ability SC indicator variables across the SPPC and (A)SDQ-I instruments for boys is also consistent with a method effect related to the SPPC response format. That is, for both of these scales, the descriptive statistics are more, or are exclusively suggestive of non-normality for the (A)SDQ-I than for the SPPC instruments. On the basis of the alternating direction method effect hypothesized to have occurred for the SPPC instrument, a decrease in non-normality (and a regression towards the mean) would be expected for dual-item indicator variables paired across response directions. Accordingly, differences due to this method effect would be most apparent for scales in which indicator variables are non-normally distributed. Hence, the differences in the extremity of the distributions of these indicator variables and scales across the (A)SDQ-I and SPPC instruments for boys.

With regard to the (A)SDQ-I instruments, preliminary analyses demonstrate strong support for the psychometric soundness of the (A)SDQ-I. In particular, the internal consistency and the factorial validity of each of the (A)SDQ-I scales are supported. Nevertheless, for both girls and boys, the Physical Ability scale is the least internally consistent of the (A)SDQ-I scales, and this scale is the only (A)SDQ-I scale to be less internally consistent than its corresponding SPPC scale. Consistent with these findings, and as noted previously, there is, for the sample of boys, less non-normal distributional consistency across the (A)SDQ-I indicator variables for this scale than is true of the Gym SC scale. Thus, although the results support the internal consistency and factorial validity of each of the (A)SDQ-I scales, the pattern of results also suggest that the Physical Ability SC scale is the least cohesive of the (A)SDQ-I scales.

Self-concept Multidimensional and Hierarchical Measurement and Structure:

Within-network Issues

Consistent with the review of the results, the findings bearing on the within-network component of the SC construct will be discussed first with respect to the multidimensional

structure of SC, and then with regard to its hierarchical structure. Gender differences in the measurement and structure of a multidimensional and hierarchical SC construct will be discussed subsequent to a discussion of the results obtained for each gender.

Self-concept Multidimensional Measurement and Structure

CFA results with regard to the competing multidimensional models of SC provide support for a highly differentiated preadolescent SC for both girls and boys. That is, as hypothesized, results demonstrate support for the most differentiated model of multidimensional preadolescent SC structure for both girls and boys. Results demonstrate strong support for each of the 16 a priori SC facets, and in particular, for preadolescent SC differentiation within sub-areas of mathematics (i.e., Arithmetic and Measurement) and Language Arts (i.e., Reading, Writing, Speaking, and Listening). For both girls and boys, the 16-factor model of SC yields an excellent fit to the data, an absence of distinct model misfit, and accounts for a greater proportion of the variance in the data than do the less differentiated 13- and 15-factor models.

In contrast to the 16-factor SC model, the fit of the less differentiated models is poorer and yields correlated measurement errors among the indicator variables represented under the less differentiated factors. Incorporation of the correlated errors improves the statistical and practical fit of these models such that they provide a fit to the data that is less than or equal to the a priori 16-factor model. However, the 16-factor model represents the data in a theoretically and methodologically superior manner than do the less differentiated models, namely, with a priori latent constructs rather than post hoc correlated measurement errors. Taken together, the excellent fit of the a priori 16-factor model of SC, the absence of any distinctly ill-fitting parameters, as well as the nature and the consistency of the misfit of the less differentiated 13- and 15-factor models across gender groups demonstrate support, for both girls and boys, for the 16 hypothesized dimensions of preadolescent SC. Accordingly, the first hypothesis, in which SC was postulated as a multidimensional construct comprised of 16 SC dimensions, is supported for preadolescent girls and boys.

The magnitude of the correlations among the two specific mathematics and the four specific Language Arts SC facets are higher than has typically been found in SDQ research. That is, near-zero correlations have been observed between mathematics and Language Arts SC facets in SDQ research limited to investigations of reading and mathematics SC facets (e.g., Marsh, Barnes, et al., 1984; Marsh, Smith, & Barnes, 1983; for a review of this research, see Marsh, 1986c). However, the present results are consistent with research in which an expanded set of academic SC facets was considered (Marsh, 1990d), and with the suggestion that the comparison processes posited in the I/E model are heightened when limited academic areas are considered (Marsh, 1990d). Thus, the results replicate earlier research in which the academic SC areas considered were expanded, and support Marsh's hypothesis that the differing results across these empirical studies are related to this expansion.

In summary, the pattern of results obtained provide strong support for a highly specific multidimensional preadolescent SC construct for both girls and boys, and in particular, the 16 a priori SC facets hypothesized. These results argue strongly against the value of unidimensional measures of SC. Moreover, consistent with research by Marsh (e.g., 1990d), the findings provide support for a greater level of SC specificity and differentiation than has typically been examined. Indeed, the present study represents the most specific and differentiated investigation of mathematics and Language Arts SC constructs. Thus, consistent with both the criticisms of earlier SC research (e.g., Shavelson et al., 1976; Wylie, 1974, 1979) and the more recent conclusions of SC researchers (e.g., Marsh, 1990a, 1990d), results indicate that support for the multidimensionality of SC can be demonstrated when psychometrically sound measuring instruments are utilized. Finally, from a practical perspective, the present results suggest that educators, clinicians, and researchers are advised to measure and target multidimensional SCs specific to the subject area(s) of interest. That is, the multidimensionality of SC may need to be reflected at an applied level in order to permit the most useful and valid information to be obtained. We turn now to a discussion of the results relevant to tests of the competing hierarchical models of SC and their

implications with regard to the hierarchical structure of preadolescent SC for girls and boys. This is followed by a discussion of the final best-fitting multidimensional hierarchical SC models obtained for girls and boys.

Self-concept Hierarchical Structure

Comparison of the competing a priori multidimensional hierarchical SC models tested in the present study suggest that the hierarchical structure of SC is more differentiated than has previously been suggested. Consistent with that hypothesized, tests of the competing hierarchical models reveal that representing Gym and Physical Ability SCs under the same, as opposed to different second-order SC facets provides a better representation of the hierarchical structure of SC for both girls and boys. Thus, the hypothesis regarding the hierarchical representation of Gym SC under Nonacademic/Physical SC is supported.

Tests of the hierarchical representation of Gym SC indicate that the better-fitting model nevertheless yields a modest degree of misfit for that portion of the SC hierarchy under which Gym SC is represented, namely, second-order Nonacademic/Physical SC. An enhancement in the fit of this portion of the SC hierarchy was potentially offered by an increased level of differentiation between the nonacademic first-order social and physical SC components comprising this second-order SC construct. Tests of an a priori model specifying a second-order differentiation of the physical, social, and artistic SC factors, however, could not be appropriately estimated for either girls or boys. Although it might be tempting to conclude on the basis of these results that separate representation of the physical, social, and artistic components of the SC hierarchy is inappropriate, it is believed that such a conclusion is premature. Research examining the multidimensionality of the physical (e.g., Marsh, 1996b, 1997), social (e.g., Byrne & Shavelson, 1996), and artistic (e.g., Vispoel, 1995) portions of the SC construct, although focusing primarily on adolescents and older individuals, suggest that these SC facets may be more subject-specific than was measured in the present study. In addition, the range of first-order artistic SC facets included in tests of this hierarchical model is narrower than that originally hypothesized for this model. In the absence of the inclusion of a broader range of physical, social, and, in particular, artistic SC

facets, the appropriate and thorough evaluation of this model is not possible. Accordingly, conclusions as to the validity of this more differentiated hierarchical model for preadolescent girls and boys must await further research.

With regard to the core academic portion of the SC hierarchy, the results support the hierarchical separation of the more specific first-order mathematics and science SCs for preadolescent girls and boys. Thus, the a priori specified expansion of the core academic portion of the SC hierarchy is supported. In the absence of this differentiation, correlated disturbance terms among the specific first-order mathematics SC constructs are obtained for both girls and boys. These findings indicate that systematic errors in the prediction of these first-order constructs, and thus, model misfit, occur in the absence of this differentiation. This pattern of results are consistent with previous research which revealed that first-order mathematics SCs are less adequately defined than are science SCs when these constructs are represented under the rubric of the same second-order SC facet (Marsh, 1990d). The present findings indicate that the differentiation of these first-order SC constructs at the second-order level eliminates this model misfit. In addition, consistent with the rationale for the second-order separation of science and mathematics constructs, the measurement and the structure of the core academic portions of the SC construct are exceptionally well-defined for girls and boys with this differentiation. Thus, the results suggest that the hierarchical structure of the core academic portion of preadolescent SC for girls and boys is more differentiated than that represented in the Marsh/Shavelson model. In particular, the results indicate that specific science and mathematics SC constructs are best represented separately at the second-order level of the SC hierarchy for preadolescent girls and boys.

The pattern of correlations among the second-order core academic SC facets obtained for the final best-fitting hierarchical SC models for both girls and boys further substantiates the separation of the Science- and Mathematics-Academic SC constructs. That is, for both girls and boys, the second-order science SC facet is more highly related to Verbal-Academic SC than to Mathematics-Academic SC. Similarly, Verbal-Academic SC is more highly related to Science- than to Mathematics-Academic SC. In fact, the Science-Academic,

Verbal-Academic, and Non-Core Academic SC constructs are more highly related to each other than to Mathematics-Academic SC for both girls and boys. However, for both girls and boys, Science-Academic SC is more highly correlated with Mathematics-Academic SC than is Verbal-Academic SC.

Taken together, the pattern of correlations support the separation of verbal, science, and mathematics SC constructs, and a pattern in which SC constructs in science are more strongly related to mathematics than are verbal constructs, but more strongly related to verbal than to mathematics constructs. The greater differentiation between mathematics and verbal SCs than between mathematics and science SCs is consistent with both within- and between-network SC research (e.g., Marsh, 1990d; Marsh, Parker, & Smith, 1983). A closer association between science and verbal constructs than between science and mathematics constructs is, however, contrary to the implications of earlier SC/AA research measuring limited academic subject areas (i.e., Marsh, Parker, & Smith, 1983). In contrast, the results are consistent with those obtained for preadolescent boys when a broad range of academic areas were considered. Specifically, Science and Social Studies SCs were found to be less highly correlated with mathematics SC than with Reading SC, and equally or less highly correlated than with Spelling SC (Marsh, 1990d).

The pattern of relations among the core academic SCs are interpreted within the context of the nature of the subjects areas which they reflect (see Appendix B). That is, SCs in mathematics subjects areas might reasonably be conceptualized as primarily driven by perceptions of competence in relation to the manipulation of numbers. In contrast, SCs in verbal areas might reasonably be conceptualized as primarily based on perceptions relevant to one's language abilities. However, language/verbal skills would be expected to play a more significant role in science areas than would skills in relation to the manipulation of numbers. The distinctiveness of the nature of the content areas underlying the SC constructs would, thus, yield the greatest distinctiveness between mathematics and verbal SCs, and the least distinctiveness between science and verbal SCs, with the distinctiveness between

science and mathematics SCs falling somewhere in between. Hence, the observed pattern of second-order core academic SC relations for both girls and boys.

The specific parameters in the final best-fitting models differ for girls and boys only with regard to the relations among the first- and second-order SC facets. Specifically, results suggest that Parent SC is uniquely defined by higher-order verbal SC for boys, and by higher-order non-core SCs related to art and religion for girls. Thus, previous empirical research demonstrating an association between Parent SC and second-order nonacademic and academic SCs (e.g., Marsh & Shavelson, 1985) was not replicated. Although quite speculative, perhaps the level and quality of parental interactions, and thus, Parent SC is more strongly related to verbal SCs for preadolescent boys than girls. In contrast, for girls, it is perhaps the nature of their parental interactions (i.e., the religious/moral values, and/or the artistic emphasis transmitted) that is the central impact of Parent SC. Alternatively, it may be that the lack of differentiation provided by one of the indicator variables for the Parent SC construct (as noted above) led to the unexpected results with regard to this factor. Nevertheless, future research is required to ascertain whether the gender difference observed for this construct is replicated.

Results for the final hierarchical models for both girls and boys reveal that both the measurement and structural components of all portions of the models are well-defined. That is, the factor loadings representing the measurement portion of the SC facets are well above acceptable values, including those for the Physical Ability SC indicator variables, despite their lower magnitude as compared to the other SC constructs. Similarly, the structural parameters linking the first- and second-order SC facets indicate that the first-order parameters are adequately represented by their second-order SC facets. This was true despite a modest degree of misfit in the nonacademic/physical portion of the SC hierarchy.

Nevertheless, a consistent pattern was evident across the final models for girls and boys which indicates that some of the nonacademic and non-core academic first-order SC facets are less strongly defined by their second-order facets than is true for the core academic portion. Specifically, Religion SC facet is more strongly defined by its hierarchical Non-core

Academic SC than is the first-order Art SC facet; these results are consistent with those found by Marsh (1990d). Likewise, for both girls and boys, the Physical Ability and Gym SC facets are more strongly defined by their second-order SC facet than are Physical Appearance and Peer SCs; this finding is consistent with the increased physical competence emphasis of this portion of the SC hierarchy with the inclusion of Gym SC. Given the inability to be confident that the first-order SC facets measured in the present study permit an adequate test of the limits of the non-core academic and nonacademic portions of the SC hierarchy, this aspect of the final hierarchical models must be viewed more tentatively. A more thorough investigation of the hierarchical structure of the non-core academic and nonacademic portions of the SC hierarchy, and the extent to which the representation of their first-order SC facets can be enhanced will be possible with the inclusion of an expanded set of physical, social, and artistic SC dimensions.

It is noteworthy that the Writing in Language Arts SC facet is well represented under its hypothesized Verbal-Academic SC construct in the final hierarchical model for both girls and boys. These results are in contrast to previous results in which Handwriting SC was poorly represented within the verbal portion of the SC hierarchy (Marsh, 1990d). It is possible that the Handwriting SC facet measured by Marsh is not equivalent, and/or was interpreted differently by respondents than was the Writing in Language Arts SC facet in the present study. Whether these differing empirical findings are related to the inclusion of a Language Arts label, the non-inclusion of the word “hand”, and/or other factors cannot be differentiated. However, the school boards participating in the present study provide a separate Handwriting category for teachers to provide grades, and this category is not represented under Language Arts. Handwriting would seem to convey components relevant to neatness, and artistic attention to, or interest in the aesthetic aspects of writing, whereas writing, and particularly writing in Language Arts, would convey more of a verbal skills component. The present results nevertheless provide support for a first-order Writing in Language Arts SC facet that is appropriately conceptualized hierarchically as a construct that is primarily verbal in nature.

The magnitude of the correlations among the second-order SC facets indicate that despite the distinctiveness of these constructs, they are nevertheless highly associated for preadolescent girls and boys. The strength of the correlations among the SC constructs are interpreted within the context of the hypothesized developmental characteristic of SC (e.g., Harter, 1985; Shavelson et al., 1976). That is, although the SC constructs have emerged as distinct constructs for preadolescent girls and boys, given their highly specific nature, they are not as distinct as would be expected for older students. Accordingly, the SC differentiation process is expected to continue and yield a greater level of distinctiveness among the SC constructs for older students than is true for the present sample. This interpretation is consistent with the pattern of results obtained in research examining an expansion of social SC to include more specific multidimensional SC constructs (i.e., Byrne & Shavelson, 1996). That is, comparison of the correlations among these more specific social SC constructs across samples of Grades 3, 7, and 11 students revealed a pattern in SC facets were more differentiated with increasing grade level.

It is interesting to note that for both girls and boys, General-Academic SC was linked hierarchically to Mathematics- and Verbal-Academic SCs but not to Science-Academic SC. Although General- and Science-Academic SC facets were hypothesized to be associated hierarchically, this hypothesis is not supported. However, the obtained hierarchical structure for General-academic SC is consistent with the assumption that presumably was underlying the focus on mathematics and verbal SC constructs of early SDQ research (e.g., Marsh, Relich, & Smith, 1983). Namely, that mathematics and verbal constructs are most central to academic SC. Hence, for both girls and boys, the results suggest that the hierarchical structure of SC includes a link between General-Academic SC and both mathematics and verbal SCs, but not with the academically less central, but still core, higher-order Science-Academic SC. Thus, the results suggest, consistent with the academic areas forming the focus of early SC research, that mathematics and verbal areas are the most central to core academic SCs.

In summary, the present results support a multidimensional and hierarchical SC structure for preadolescent girls and boys that is well defined at both the measurement and structural levels. For both girls and boys, a higher level of differentiation is indicated for the core science and mathematics academic portion of the SC hierarchy, as is an amalgamation of Gym and nonacademic SC facets, as compared to their separate hierarchical representation. Nevertheless, the results with regard to the Nonacademic/Physical and the Non-core academic portions of the SC hierarchy suggest the need for research examining alternate conceptualizations of these portions of the SC hierarchy so as to enhance the representation of their associated first-order SC facets. As a result, conclusions regarding these portions of the SC hierarchy must necessarily be tentative. We turn now to a review of results related to tests of the invariance of these baseline multidimensional and hierarchical SC models across gender.

Gender Invariance: Self-concept Multidimensional Measurement and Hierarchical Structure

With the exception of three measurement and two structural parameters, results support the hypothesis of equivalent SC measurement and structure across gender. Furthermore, in all instances, the gender differences found are of a small magnitude. The results for the measurement parameters indicate that these indicator variables more strongly define their corresponding first-order SC factors for girls than for boys. Several possible explanations are considered to explain the gender differences found for the three measurement parameters.

First, it is notable that two of three non-equivalent measurement parameters (i.e., ARM2 and GYM2) are single-item indicator variables. The rationale against the use of single-items is that sample-specific idiosyncrasies reflecting trivial artifacts are more likely (e.g., Marsh, 1992b, 1993b). Thus, it is possible that the finding of non-equivalence for these single-item indicator variables are a reflection of this type of phenomenon rather than true gender differences that are generalizable to preadolescent girls and boys. Alternatively, ASDQ-I Item 2 refers to whether the children have “always done well”, and it may be that girls and boys interpreted and responded to this question differently. That is, what constitutes

“done well” may differentially relate to the measurement of preadolescent girls’ and boys’ SCs in these areas. However, under this explanation, differences would also be expected for this item in other academic areas (e.g., science). Accordingly, the former explanation is believed to be more likely than the latter.

Another common link among the measurement parameters found to vary across gender is that between the Gym and Physical Ability SC indicator variables (i.e., GYM2 and AB35). Both of these measures tap first-order factors represented under the rubric of the second-order Nonacademic/Physical SC facet. The non-equivalence of these parameters may be related to the nature of the items comprising these indicator variables, and in particular, the Physical Ability indicator variable. Specifically, this indicator variable is comprised of items which refer to how good the children are at sports and as athletes, respectively. In contrast, the remaining Physical Ability scale items refer to more concrete and specific activities and characteristics (e.g., running speed and endurance, muscles, and throwing a ball). Perhaps the more abstract items operated differently for girls than for boys. In addition, it may be that the non-equivalence of these abstract items is related to their inclusion within a context in which Gym SC was also measured. That is, perhaps there are gender differences in the manner in which aspects of sports and athletic performance define preadolescents’ Physical Ability and Gym SCs. In previous SDQ-I research, the items were collapsed without apparent regard to similarity of content and thus, may have masked the non-equivalence found in the present study. Thus, single-item analyses of the Gym and Physical Ability SC facets are needed to more clearly discern the equivalence or nonequivalence of the measurement of these SC facets across gender, as well as potential gender differences in the nature of the components defining these two constructs.

Consideration of the finding of non-equivalence of the Gym and Physical Ability measurement parameters must also include an awareness of the fact that the tests of invariance were not based on robust estimates. In both instances, these indicator variables are generally normally-distributed for girls but are relatively highly non-normal for boys. In the absence of a correction factor provided by robust estimates with the current version of

the EQS program, perhaps these measurement parameters were found to be non-equivalent when comparison of robust estimates would not have yielded a difference. Future advances in methodology will permit an evaluation of this possibility. Nevertheless, additional research is needed to determine whether the parameters identified as non-equivalent are replicated, and if so, to more clearly discern the processes underlying these differences.

The results with regard to the non-equivalent structural parameters suggest that the first-order Listening in Language Arts and General-Academic SC facets are more strongly defined by their higher-order Verbal-Academic SC for boys than for girls. Perhaps Listening in Language Arts is more reflective of behavioral aspects (see Appendix B) relevant to cooperation and compliance in Language Arts that is a more important and/or stronger component of verbal SC for boys than for girls. This suggestion is consistent with research demonstrating a link between reading and behavior problems among primary school children (e.g., Horn & Packard, 1985) and, in particular, a higher rate of behavior problems for preadolescents with reading difficulties among boys than among girls (e.g., Maughan, Pickles, Hagell, Rutter, & Yule, 1996; Smart, Sanson, & Prior, 1996). Thus, SCs of listening (behavior) in Language Arts might reasonably be found to be more strongly defined by higher-order verbal SC constructs for boys than for girls.

The gender difference for the General-Academic SC parameter, however, is less consistent with both theory and empirical research. That is, theoretical assumptions and empirical research suggest that general-academic and verbal constructs are more highly related for girls than for boys (i.e., Byrne & Shavelson, 1987). Within a SC enhancement model, gender differences in the relations among SC constructs are expected to follow a pattern which permits the development of higher SCs. Thus, gender differences in the magnitude of the correlations between first- and second-order SC constructs are expected to be parallel gender differences in SC construct means. Given that preadolescent SC research has typically found boys to have lower SCs in verbal areas than girls (e.g., Marsh, Relich, & Smith, 1983), gender differences in the relation between general-academic and verbal constructs, consistent with a SC enhancement hypothesis, would be expected to favor girls

(i.e., be higher for girls), as was found for adolescents (Byrne & Shavelson, 1987). Thus, results for this parameter are not consistent with the direction that would be expected on the basis of previous theory or empirical research. However, it may be that the higher Listening in Language Arts and Verbal-Academic SC association for boys, in combination with their common Language Arts component, similarly enhanced the association between General- and Verbal-Academic SCs.

Alternatively, the direction of the gender difference for the General- and Verbal-Academic SC constructs may be related to the difference across girls and boys in the final SC hierarchical models. That is, this portion of the hierarchical model for boys includes a path linking the Parent and Verbal-Academic SC facets that was not maintained for the final model for girls. Perhaps the inclusion of this parameter for boys, in the absence of the same path in the model for girls, had the effect of accentuating the associations among the first- and second-order verbal constructs for boys, and thus, led to gender differences in this portion of the hierarchical model that are reflected in General-Academic, and perhaps, Listening in Language Arts SCs as well.

Tests of the competing hierarchical models conducted separately for each gender suggest that the change in the representation of Gym SC is more critical for boys than for girls, and that the second-order separation of the science and mathematics SC constructs is more critical for girls than for boys. Interestingly, the only gender differences in the SC construct relevant to the parameters for which the hierarchical structure of SC changed for the final, as compared to earlier, hierarchical models are measurement parameters. That is, tests of gender differences did not support any structural differences in SC across girls and boys for the nonacademic/physical portion of the SC hierarchy; only measurement differences were identified for this portion of the SC hierarchy (i.e., GYM2 and AB35 indicator variables). Similarly, tests for gender invariance reveal a difference at only the measurement level of the SC hierarchy for the portion relevant to the second-order separation of science and mathematics SC constructs (i.e., for MT2). Thus, the present

results highlight the important advantage permitted by empirical tests of invariance across groups under covariance structure analysis.

Taken together, the results of the present study suggest that there are few gender differences in the measurement and structure of SC; furthermore, any gender differences that may exist are of a small magnitude. Nevertheless, tests of invariance reveal that gender differences may exist in the measurement of SCs in mathematics and physical competence (i.e., gym and physical ability) areas, and the structure of SCs in verbal and general-academic areas. Thus, comparisons of observed means across preadolescent girls and boys for these constructs may not be valid. Given that gender differences in observed means have typically been identified for physical ability, verbal, and mathematics SC constructs (e.g., Marsh, 1992b; Marsh, Relich, & Smith, 1983; Marsh, Smith, & Barnes, 1983, 1985), the present findings indicate that conclusions derived on the basis of such comparisons must be made cautiously. Accordingly, educators, clinicians and researchers need to be cognizant of these issues when making interpretations of SC scale scores.

Between-network Issues: Self-concept/Academic Achievement Relations

For both girls and boys, results did not support the differentiation of Reading, Writing, and Speaking in Language Arts AAs. Although Listening in Language Arts AA is moderately distinct from each of these three Language Arts AAs for girls, this finding is less true for boys. In addition, for boys, the relations among Science and Language Arts AAs provide less support for the differentiability of these constructs from each other for boys than for girls, although the differentiated pattern of relations between these AA constructs and SC factors mitigates their representation as a single unitary construct. The lack of theoretical justification for collapsing these constructs further argues against their unitary representation.

The lack of differentiation among the Reading, Writing, and Speaking in Language Arts AA facets can be considered reasonable within the context of SC theory and previous empirical research. That is, the comparison processes posited in the I/E model predict that SCs are more differentiated than are their corresponding AA indicators (e.g., Marsh, 1987b).

Research examining reading and mathematics SC/AA relations has substantiated that children differentiate less among achievement constructs than among corresponding SC constructs (e.g., Marsh, 1992a; for reviews of this and related research, see, for example, Byrne, 1996a; Marsh, 1990a). Given an already high level of correspondence among the Language Arts SCs, a higher level of correspondence among these AA constructs renders the differentiation among at least three of the Language Arts AA facets superfluous.

The pattern of results for Reading, Writing, and Speaking in Language Arts AA constructs is also understood within a conceptual perspective that distinguishes between Language Arts skills and behavior. Specifically, it is hypothesized that the Reading, Writing, and Speaking in Language Arts AA constructs might reasonably be conceptualized as skill-based aspects of Language Arts, whereas Listening in Language Arts might more appropriately be designated as reflecting more behaviorally-linked aspects of Language Arts classroom behavior (see Appendix B). This differentiation is consistent with research suggesting that reading SC is itself multidimensional and that SCs relevant to reading competence can be differentiated from those related to attitudes towards reading (Chapman & Tunmer, 1995); listening might thus represent a behavioral manifestation of the attitude construct. From a perspective distinguishing between skills and listening behavior, the lack of differentiation among the former three Language Arts AA constructs is considered theoretically feasible.

With regard to the multidimensionality and content-specificity of SC/AA relations for both the first- and second-order SC constructs, results generally support the hypothesis that relations between AA and first-order SC constructs are stronger for corresponding than for non-corresponding content areas. The exception to this pattern is for the Language Arts content area for both girls and boys, and the Art AA construct for girls; the content-specificity of Art SC is, however, supported. In particular, to varying degrees, the pattern of latent SC/AA factor correlations for the Language Arts constructs indicate that the content-specificity of their relations is not restricted to the Language Arts areas, but rather that they are equally correlated with latent factors in other core academic areas, and in particular,

science. However, the results are generally consistent with the Language Arts listening and skill-based differentiation. That is, SC/AA relations reveal that Language Arts constructs are frequently equally correlated with their corresponding constructs as with that of other “skill-based” academic areas; they are, however, typically less correlated with the hypothesized “behavioral” aspects of academic constructs (i.e., Listening in Language Arts). Thus, although the pattern of results substantiate the separation between listening and skill-based Language Arts constructs, they fail to support a content-specific pattern across the Language Arts constructs.

Taken together, the pattern of SC/AA relations suggest that the Language Arts constructs are more general than their designation implies. The results are interpreted to suggest that SC and AA constructs in Language Arts areas share a less specific pattern of relations than implied by the content-specificity of their Language Arts designation. Consistent with this, the results are also interpreted to suggest that content-specific skill-based and Listening in Language Arts SCs are related to performance and behavior, respectively, across a wider variety of academic contexts than merely Language Arts. Given that writing skills and listening behavior are exhibited across a wider variety of academic areas than merely Language Arts, such a hypothesis appears theoretically feasible. Thus, while the pattern of results suggest that both the skill- and behaviorally-based Language Arts constructs have a content-specific Language Arts element, they also suggest that these constructs have an academic element that is common to other core academic areas and, in particular, to science.

A theoretical distinction among core and non-core academic areas (Marsh, 1990d) is also useful in examining the pattern of SC/AA relations. That is, SC/AA relations among core constructs are generally higher than those between core constructs and either non-core academic or nonacademic constructs. However, consistent with the earlier differentiation of the mathematics areas from the science and Language Arts areas, of the core academic constructs, mathematics SC facets are least correlated with non-corresponding science and verbal AA constructs for both girls and boys. Similarly, Science AA is less highly correlated

with SC facets in mathematics than in other academic areas. Thus, the pattern of results support both the core/non-core academic distinction and the separation of mathematics constructs from science constructs.

With regard to the Art AA construct, post hoc analyses suggest that the lack of support for the content-specificity of this construct is related to report card grades, as provided by teachers, rather than self-reported grades. This finding is consistent with research which suggests the presence of a method/halo effect across subject areas for grades provided by the same teacher, in the absence of a similar effect for grades across subject areas provided by different teachers (Byrne & Worth Gavin, 1996). Presumably, teacher grades are less independent, and thus content-specific, across subject areas when provided by the same, as opposed to different teachers. Such a hypothesis seems reasonable, and is consistent with the difference found for girls in the content-specificity of self-reported grades, as compared to report card grades in Art in the present study. Thus, perhaps the results for the Art AA construct for girls are due to a non-independence among teacher ratings, as reflected in report card grades in Art, rather than a phenomenon specific to the Art AA construct. However, given that the present study represents the only empirical attempt to examine relations among these SC and AA constructs, further research is needed to more clearly evaluate the SC/AA nomological network for preadolescent girls and boys for art constructs, particularly in relation to the nature of the AA construct utilized.

The pattern of results relevant to the hierarchical pattern of SC/AA relations obtained for both girls and boys is, as with previous SC research, inconsistent with theorizing as to an hierarchically structured SC construct. Thus, the results did not generally support the hypothesis that SC/AA relations would follow an hierarchical pattern. That is, SC/AA relations at the first-order level of the SC hierarchy are frequently equal to, or smaller than those at the second-order level of the SC hierarchy. Despite the incongruence between SC theory and empirical research with regard to the between-network hierarchical characteristic of the SC construct, a theoretical explanation of this apparent inconsistency has not been provided in the literature. In the present study, the previously proposed distinction between

language arts skills and behavior, and the proposed generality of verbal skills and behaviors across academic areas is utilized to reconcile the apparent contradiction between empirical research and a theoretical model specifying an hierarchical pattern of SC/AA relations.

The explanation offered is believed to reconcile the apparently contradictory nature of theorizing and empirical research relevant to an hierarchical pattern of SC/AA relations, with the exception of the results for the Art AA construct for girls. As previously noted, the results with regard to Art AA are believed to be specific to the non-independence of the teachers providing report card grades across academic subject areas. It is believed that the explanation offered also more appropriately recognizes the overlapping content component of both general academic and verbal constructs across academic areas, and in particular, science academic areas.

It is proposed that subject-specific core academic areas, and to a lesser extent, non-core academic areas, while containing a content-specific element, also share a common academic element. Such a proposition is consistent with the emerging distinctiveness of the academic constructs, as reflected in the magnitude of their latent factor correlations. In addition, consistent with the pattern of correlations and earlier arguments with regard to the underlying nature of the core academic constructs, the Language Arts skills constructs are hypothesized to be tapping an academic component common to the academic areas to a greater extent than are science constructs, and in particular, mathematics and non-core academic constructs. Again, this is not surprising given that Language Arts skills might reasonably play a role in science areas to a greater extent than in mathematics or non-core academic areas. Finally, Listening in Language Arts constructs are hypothesized to be tapping behavioral components specific to Language Arts, as well as a behavioral element common across academic contexts.

As a result of the dual content-specific and common general-academic components of the academic constructs, support for a content-specific multidimensional pattern of SC/AA relations would be expected for those constructs sharing a smaller general-academic component. In contrast, a more generalized pattern of SC/AA relations would be expected

for areas sharing a large common academic component. Thus, support for a content-specific pattern of relations for Language Arts constructs would be expected to be less consistent given their greater generality across academic areas for which verbal skills are relevant (e.g., science). A similar principle would be true of listening behavior, both in relation to its Language Arts content-specificity and its generality across academic areas.

With regard to an hierarchical pattern of relations, one would expect a reduction in the strength of the specific components with a concomitant increase in the extent to which a general component is shared. Depending upon the relative change in the extent to which specific, as compared to general, components are shared among constructs, a vertical pattern of hierarchical SC/AA relations may or may not be supported. That is, if the increase in the common academic component exceeds the loss of the specific component, a decrease in the magnitude of SC/AA relations as the SC hierarchy is ascended would be mitigated. Given that verbal constructs are hypothesized to generalize the most to other academic areas, one would expect to find the least support for the content-specificity of these academic constructs, and the most support for the mathematics constructs. Hence, an apparently contradictory pattern of multidimensional and hierarchical pattern of SC/AA relations.

The theoretical explanation offered for the pattern of SC/AA relations obtained in the present study suggest that theoretical models and hypotheses relevant to the within- and between-network SC nomological network may benefit from incorporating several distinctions. One of these distinctions is that between skill-based and behaviorally-based aspects of academic constructs. In addition, the core/non-core distinction proposed by Marsh (1990d) appears to be a relevant dimension in understanding SC/AA relations. The extent to which academic constructs share common academic components, particularly in relation to verbal elements, and are distinct as a result of elements more relevant to mathematics (i.e., numbers) than verbal areas also appears to be a useful distinction. Thus, the present results suggest that with an expansion of the academic portion of the SC construct, more refined distinctions are needed to explain its nomological network.

Gender Invariance: Self-concept/Academic Achievement Relations

A total of six non-equivalent parameters were identified in tests of invariant SC/AA relations for first-order SC facets, and a total of three were identified in tests related to the second-order SC factors. Thus, with the exception of these nine parameters, the hypothesis of invariant SC/AA relations across gender is supported. However, of these nine parameters, five involve gender differences that are of a small magnitude, and one is a parameter that is nonsignificant for both girls and boys. We turn now to a discussion of the findings for these six non-equivalent parameters, beginning first with those that are measurement parameters and following this with a discussion of the non-equivalent structural parameters. The results for the three parameters that involve gender differences that are not of a small magnitude, all of which are structural parameters, will then be discussed.

The non-equivalent measurement parameters involve gender differences that are of a small magnitude and include the single-item indicator variable representing Writing in Language Arts SC (i.e., WLA2), and one of the indicator variables representing the Science SC facet (i.e., SN45). The parameters were identified as non-equivalent across gender under invariance tests related, respectively, to the first- and second-order SC facets. The former parameter is higher for girls, and the latter is higher for boys. Drawing from research indicating that model misfit of a small, as compared to a large magnitude is less likely to replicate across independent samples (e.g., MacCallum, 1986), it seems reasonable to assume that this would also be true of gender differences that are of a small magnitude, as is the case for these indicator variables. In addition, given that these parameters were not identified in initial tests of the invariance of the measurement and structure of SC across gender, it seems particularly unlikely that their non-equivalence across gender will replicate across independent samples. Finally, one of these indicator variables is represented by a single-item and, as noted earlier, such indicator variables are generally more vulnerable to sample-specific idiosyncrasies, and thus, are less likely to replicate across samples than are dual-item indicator variables. Thus, it is believed that the likelihood of replicating a gender

difference for these indicator variables across independent variables is quite low, and that their non-equivalence is likely a reflection of a sample-specific artifact.

Under invariance tests for relations between AA and first-order SC constructs, three non-equivalent structural parameters involve gender differences that are of a small magnitude. The results with regard to one of these parameters suggest that the relation between Language Arts and mathematics constructs differs across gender. In particular, the relation between Measurement SC and Language Arts Skills AA is slightly higher for girls than for boys. This finding suggests the possibility that preadolescent boys differentiate more between Arithmetic AA and Language Arts skills than do girls, and thus that performance in Language Arts may be more independent of Arithmetic SC for boys than for girls. This finding, however, contradicts gender differences in these constructs found with adolescents in the only known study to examine this aspect of SC construct validity (i.e., Byrne & Shavelson, 1987). That is, grades in Language Arts were found to be correlated equally with Mathematics and Language Arts SCs for boys, whereas a content-specific pattern of relations was found for girls, as well as for mathematics grades for boys.

The results in the present study, however, are consistent with a SC enhancement hypothesis. In particular, the present pattern of results would permit the development of higher SCs in Language Arts areas for girls, and minimize an attenuation of mathematics SC for boys. That is, research examining gender differences in mean SC levels has demonstrated that preadolescent boys have higher SCs in mathematics areas and lower SCs in verbal areas (e.g., Hyde et al., 1990; Marsh, Relich, & Smith, 1983; Marsh, Smith, & Barnes, 1983, 1985; Meece et al., 1982). This pattern was evident despite girls having concomitantly higher achievements in both mathematics and verbal areas (i.e., Marsh, Smith, & Barnes, 1985). Thus, a greater interdependence of Language Arts AA and mathematics SC constructs would permit girls to maintain a higher mean SC in both these academic areas. In contrast, a greater independence of Language Arts AA constructs from mathematics SCs would permit boys to maintain a higher mathematics SC than would be possible if their Language Arts performance was significantly related to their SC in mathematics areas. Although it is

recognized that correlation does not equate with causation, the establishment of a correlational relation between constructs is a presumptive element of a causative relation. Nevertheless, given the contradictory nature of the present findings and those obtained by Byrne and Shavelson (1987), as well as potential gender-related developmental differences between preadolescent and adolescent samples, further research is needed before any definitive conclusions can be reached as to gender differences for preadolescents in the relations between verbal and mathematics SC and AA constructs.

Two of the structural parameters identified as non-equivalent across gender involve parameters in which differences are of a small magnitude, and are of a low positive value for boys but are nonsignificant, albeit positive, for girls. Specifically, these are the relation between Gym AA and General-Academic SC, and between Arithmetic AA and Physical Ability SC. These results suggest that performance in Gym is associated with General-Academic SC for boys but not for girls, and likewise, that performance in Arithmetic is associated with Physical Ability SC for boys but not for girls. Stated differently, these findings suggest that the relation between these constructs is more differentiated or independent for girls than is true for boys.

The pattern of results with regard to the relation between Gym AA and General-Academic SC is again consistent with a SC enhancement hypothesis. That is, gender role stereotypes would suggest that if differences exist in Gym AA, they are likely to favor boys rather than girls. Thus, an interdependence between Gym AA and General-Academic SC for boys would permit a higher level of General-Academic SC. The results with regard to the relation between Arithmetic AA and Physical Ability SC are not, however, consistent with a SC enhancement hypothesis. That is, in light of the trend toward higher mathematics achievement for girls noted earlier, a greater interdependence among Arithmetic AA and Physical Ability SC would not permit higher Physical Ability SC for boys than if these constructs were independent. In addition, research suggests that gender differences in mean Physical Ability SC levels among preadolescents favor boys (e.g., Marsh, Relich, & Smith, 1983; Marsh, Smith, & Barnes, 1983, 1985). However, it may be that a pattern in which the

performance of boys in mathematics is superior relative to language arts, and an internal comparison of their relative performance in these areas, as posited in the I/E model, permits higher Physical Ability SC for boys with the interdependence of Arithmetic AA and Physical Ability SC. Future research is required to more clearly establish gender differences in the relations among these constructs.

It is important to note that the magnitude of the gender difference for the relation between Arithmetic AA and Physical Ability SC, and between Gym AA and General-Academic SC is small. Accordingly, future research is required to more clearly ascertain whether these gender differences are replicated. In addition, it is also important to consider the results for the Physical Ability SC construct within the context of the non-normal distributional differences between gender groups for the scale measuring this construct. Given that tests of invariance within the current version of the EQS program do not allow for a correction for non-normally distributed data, as was the case for boys but not girls, whether the gender difference for this construct would be replicated with such a correction remains unanswered. Thus, results bearing on gender differences relevant to this construct must necessarily be viewed more tentatively than if there were no distributional gender differences.

The relation between Science AA and the second-order Nonacademic/Physical SC construct was also identified as non-equivalent across gender. However, these parameters are not significant for both girls or boys, albeit the relation is positive for girls and negative for boys. Although this finding suggests the possibility that Science AA shares a low positive association with a more general Nonacademic/Physical SC for girls, and a low negative association for boys, it is probable that the association between these two constructs is negligible for both gender groups. We turn now to a discussion of the results related to the invariance of SC/AA relations for which gender differences are of a larger magnitude than is true for the parameters discussed above.

Consistent with the differing pattern of results obtained across gender for the Art AA construct in analyses of SC/AA relations conducted for girls and boys, gender differences

that are of the largest magnitude involve relations between SC constructs and Art AA. Not surprisingly, the Art SC/AA relation is higher for boys than for girls. In addition, the relation between Art AA and both the first-order Speaking in Language Arts SC and the second-order Verbal-Academic SC is higher for girls than for boys. The Art SC/AA gender difference suggests that the relation between Art SC and Art AA is higher for boys than for girls. Gender differences in the relation between Art AA and both Speaking in Language Arts and Verbal-Academic SCs suggest that Art AA is more highly related to these SC facets for girls than for boys. Indeed, the relation between Speaking in Language Arts SC and Art AA was not significant for boys.

The gender differences with regard to the Art AA construct are puzzling. However, the findings are consistent with the differing pattern of results obtained for girls and boys in analyses conducted separately for each gender. The results suggest that the lack of support for a content-specific multidimensional pattern of SC/AA relations are due to both a low association with its corresponding Art SC, and a high association with verbal SCs. In addition, the results are consistent with earlier conclusions that the lack of support for the content-specificity of the Art AA construct is related to the generality of this construct to other academic area for ratings provided by teachers. Thus, perhaps the gender differences for the Art AA construct are due to a non-independence among teacher ratings, as reflected in report card grades in Art for girls. However, given that the present study represents the only empirical attempt to examine relations among these SC and AA constructs across gender, further research is needed to more clearly evaluate the SC/AA nomological network within and across gender for art constructs, particularly in relation to the nature of the AA construct utilized.

In summary, tests of the invariance of SC/AA relations suggest that there are few differences between girls and boys in the relations between SC and AA constructs. Moreover, when differences are found, these differences tend to be of a small magnitude. Gender differences that are of a small magnitude are less likely to replicate across independent samples. The only gender differences in SC/AA relations that are of a

moderately substantial magnitude are associated with the Art AA construct. With regard to Art AA, results suggest that this construct is less strongly associated with Art SC, and more strongly associated with Speaking in Language Arts and Verbal-Academic SCs for girls than for boys. Post-hoc analyses suggest that the results are related to report card grades in Art, rather than to self-reports of Art grades, and thus, may be due to the non-independence of teacher ratings in Art.

Strengths and Limitations of the Present Study

The present study represents the first attempt to investigate within- and between-network components of the nomological network of the nonacademic and academic portions of the SC construct with a broad range of SC and AA constructs. In addition, it represents the only empirical effort to examine the construct validity of the SC construct for a broad range of academic and nonacademic content areas within and across gender. Accordingly, the present study represents an important expansion of previous SC research.

The current study, however, is also characterized by several limitations. Broadly, these include the sampling frame, the utilization of non-independent SC instruments, the absence of samples with which to cross-validate findings, the limited breadth of the nonacademic and non-core academic subject areas measured, and the limited AA indicators obtained. We turn now to an elaboration of each of these limitations.

Sampling Frame

The students who participated in the present study were selected on the basis of the school they attended. Accordingly, the students were not randomly selected. In addition, the participants were not selected in such a manner as to ensure that the students were representative of Grades 5 and 6 students at either a national or local level. Accordingly, this represents a limitation on the generalizability of the present results to other Grades 5 and 6 students. Nevertheless, attempts were made at the level of the school boards to obtain schools that were located in middle-class suburban neighborhoods in order to provide information as to the demographic characteristics of the neighborhoods in which the schools were located.

Use of Non-independent Self-concept Instruments

Ideally, tests of the validity of a construct involve the use of multiple independent measuring instruments. Although potential problems with the SPPC instrument precluded the prudent use of two independent measures of multidimensional SC, the generality of the results is limited by the use of only the (A)SDQ-I measuring instruments. As previously noted, the generalizability of construct interpretations is increased when construct validation research is based on multiple independent measure of the construct of interest. Thus, future research examining the within- and between-network aspect of the SC construct can be strengthened with the use of multiple independent measures of SC.

Sample Size: Cross-validation and Replication

Although it is recommended in the literature that covariance structure analysis ideally include cross-validation comparisons (e.g., Browne & Cudeck, 1989), this is rarely the case in practice. The availability of a cross-validation sample allows researchers to more clearly ascertain whether findings are replicated across samples, and thus, whether they are sample-specific or generalizable; this has important implications for the validity of the results, and the confidence a researcher can have in the conclusions offered. The sample size originally targeted for the current study was 500 Grades 5 and 6 students for each gender group. As previously noted, it is recommended, based on sound statistical theory and empirical research, that covariance structure analyses be based on a minimum sample size of 100, and preferably, on samples sizes of 200 or more subjects (e.g., Boomsma, 1982). Given the size of the samples of girls and boys in the present study, splitting the samples for cross-validation purposes was precluded. Accordingly, there is a need for research examining the extent to which the pattern of results obtained in the present study are replicated for independent samples of preadolescent girls and boys.

The Breadth of Self-concept Dimensions

As noted previously, the current study represents the first and currently the only research project to consider both nonacademic SC dimensions, and a wide range of academic SC and achievement measures. However, the results of the present study indicate the need

for research examining a broader and more specific range of nonacademic and non-core academic dimensions of the SC construct, particularly in relation to their hierarchical structure. For example, the present study could not ascertain whether the lack of differentiability obtained between Social and Art SC facets for preadolescent girls and boys is due to a true lack of differentiation in these areas, or the absence of an expanded set of nonacademic and/or non-core academic SC facets. Thus, inclusion of a more comprehensive set of these SC facets in future research will allow a more thorough examination of the hierarchical structure of preadolescent SC for girls and boys.

Academic Achievement Indicator Variables

The AA constructs in the present study are represented with report card grades and self-reports of these grades. However, the results with regard to the Art AA construct for the sample of girls highlight the necessity of future research examining the viability of representing these indicator variables as a single construct. Future research might compare the fit of models in which various AA indicator variables are represented as separate constructs against those in which the constructs are represented as collapsed constructs. In addition, tests of the invariance of SC/AA relations across different indicator variables might similarly be conducted. The different indicator variables might be grades provided by the same and by different teachers, teacher ratings made by the same and different teachers, self-reported grades, as well as achievement tests scores.

Implications and Directions for Future Research and Clinical Intervention Efforts

The findings of the present study hold several implications for future SC research and potential implications with regard to clinical intervention efforts. With regard to the findings relevant to the instruments utilized in the present study, several recommendations are made with regard to the (A)SDQ-I and SPPC instruments. In particular, it is recommended that the extent to which the (A)SDQ-I Parent SC scale Items 2 and 5 allow for differentiation among children varying in the level of their Parent SC be examined in greater depth. Unless these items offer some measure of distinction between children having low Parent SC, as compared with children having high Parent SC, the usefulness of these items is

questionable. The items comprising the (A)SDQ-I Physical Ability and Gym SC scales also require further investigation, particularly with regard to gender differences in the nature of the physical competence components and/or characteristics which define these constructs. Nevertheless, given the strong support for the psychometric soundness of the (A)SDQ-I instruments, its use and adaptation as a measure of multiple dimensions of SC is recommended.

Results with regard to the SPPC instrument suggest that the reliability and validity of preadolescent girls' and boys' responses to this instrument are compromised by its current response format. The findings in the present study further substantiate the concerns expressed by SC researchers as to the complexity of the response format of this instrument. In addition, the results are consistent with previous research with the SPPC, and with research suggesting that negatively worded items yield less consistent and valid measures, particularly with younger and less cognitively able preadolescents. That is, in the present study, an alternating direction response bias was proposed as an extension and refinement of the negative item response bias in order to explain the SPPC method effect. Thus, the present results add to a growing body of theoretical and empirical literature highlighting the need to examine more fully the impact of differing response formats on the reliability and validity of children's responses, and in particular, to simplify the SPPC response format.

The results provide for support for a multidimensional and hierarchical SC structure for preadolescent girls and boys. Most importantly, and consistent with the conclusions of more recent reviews of SC research (e.g., Byrne, 1996a), the present study adds to the growing body of research (e.g., Byrne & Worth Gavin, 1996; Marsh, 1990d) indicating that SC, and indeed, academic SC cannot be understood without adequate attention to, and measurement of the multiple dimensions of which it is comprised. Moreover, the findings suggest the need to examine the theoretical tenability of an even wider range of SC dimensions than has typically been examined in SC research, particularly in relation to alternate hierarchical representations of the non-core academic and nonacademic portions of

the SC hierarchy. However, research is needed in order to determine the stability and the predictive validity of specific multidimensional SC constructs.

The pattern of multidimensional and hierarchical relations obtained among the SC and AA constructs in the present study suggest that a set of principles recognizing both the common and distinct components of these constructs may be required to reflect more accurately the nature of their relations. Finally, although results with regard to the invariance of SC and SC/AA relations across gender suggest that few differences exist among preadolescents, it is incumbent upon researchers to be aware of, and consider the possibility that the validity of construct interpretations of scale (or factor) scores for dimensions associated with gender stereotypes may not be equivalent across girls and boys.

With regard to the potential implications of the present study in clinical situations, it is important to recognize that generalization of the results of the present study to clinical contexts is speculative given differing contexts. Nevertheless, the strong support for the multidimensionality of SC suggests that it is important to measure content-specific multidimensional SC facets, and thus, to target these dimensions in intervention efforts. The use of more general measures may mask important characteristics of the SC constructs, as well as clinical changes and outcomes. Similarly, intervention programs and strategies designed to enhance SC may need to direct attention to more than merely general SC facets; more specifically, they may need to target, also, specific SC facets. In doing so, clinicians may effect change at more than one level of the SC hierarchy, and thereby enhance the effectiveness of their interventions.

REFERENCES

- Abu-Hilal, M. M., & Aal-Hussain, A. A. (1997). Dimensionality and hierarchy of the SDQ in a non-western milieu: A test of self-concept invariance across gender. *Journal of Cross-Cultural Psychology*, *28*(5), 535-553.
- Bearden, W. O., Sharma, S., & Teel, J. E. (1982). Sample size effects on chi square and other statistics used in evaluating models. *Journal of Marketing Research*, *19*, 425-430.
- Benson, J., & Hocevar, D. (1985). The impact of item phrasing on the validity of attitude scales for elementary children. *Journal of Educational Measurement*, *22*, 231-240.
- Bentler, P. M. (1990). Comparative fit indices in structural models. *Psychological Bulletin*, *107*, 238-246.
- Bentler, P. M. (1992). *EQS: Structural Equations Program manual*. Los Angeles: BMDP Software.
- Bentler, P. M. (1995). *EQS: Structural equations program manual*. Encino, CA: Multivariate software.
- Bentler, P. M., & Bonnett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*(3), 588-606.
- Bentler, P. M., & Wu, E. J. C. (1995). *EQS for Windows User's Guide: Beta Release*. Encino, CA: Multivariate software.
- Boomsma, A. (1982). The robustness of LISREL against small sample sizes in factor analysis models. In K. G. Joreskog & H. Wold (Eds.), *Systems under indirect observation: causality, structure, prediction* (Part 1, pp. 149-173). Amsterdam: North-Holland.
- Boomsma, A. (1985). Nonconvergence, improper solutions, and starting values in LISREL maximum likelihood estimation. *Psychometrika*, *50*, 229-242.
- Bracken, B. A. (1992). *Multidimensional Self Concept Scale*. Austin, TX: Pro-Ed.
- Bracken, B. A. (1996). Clinical applications of a context-dependent, multidimensional model of self-concept. In B. A. Bracken (Ed.), *Handbook of self-concept: Developmental, social, and clinical considerations* (pp. 463-503). New York: Wiley.

- Bracken, B. A., Bunch, S., Keith, T. Z., & Keith, P. B. (1992, August). *Multidimensional self-concept: A five instrument factor analysis*. Paper presented at the annual conference of the American Psychological Association, Washington, DC.
- Browne, M. W., & Cudeck, R. (1989). Single sample cross-validation indices for covariance structures. *Multivariate Behavioral Research*, *24*(4), 445-455.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. S. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.
- Burns, R. B. (1979). *The self-concept: Theory, measurement, development and behavior*. London: Longman.
- Byrne, B.M. (1984). The general/academic self-concept nomological network: A review of construct validation research. *Review of Educational Research*, *54*, 427-456.
- Byrne, B.M. (1986). Self-concept/academic achievement relations: An investigation of dimensionality, stability, and causality. *Canadian Journal of Behavioural Science*, *18*, 173-186.
- Byrne, B.M. (1989). *A primer of LISREL: Basic applications and programming for confirmatory factor analytic models*. New York: Springer-Verlag.
- Byrne, B.M. (1990a). Investigating gender differences in adolescent self-concept: A look beneath the surface. *Applied Measurement in Education*, *3*, 255-274.
- Byrne, B.M. (1990b). Methodological approaches to the validation of academic self-concept: The construct and its measures. *Applied Measurement in Education*, *3*(2), 185-207.
- Byrne, B.M. (1994). *Structural equation modeling with EQS and EQS/Windows: Basic concepts, applications, and programming*. Thousand Oaks, CA: Sage.
- Byrne, B. M. (1996a). Academic self-concept: Its structure, measurement, and relation to academic achievement. In B. A. Bracken (Ed.), *Handbook of self-concept: Developmental, social, and clinical considerations* (pp. 287-316). New York: Wiley.
- Byrne, B. M. (1996b). *Measuring self-concept across the life span: Issues and instrumentation*. Washington, DC: American Psychological Association.

Byrne, B. M. (1998). *Structural Equation Modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Mahwah, NJ: Erlbaum.

Byrne, B. M., & Schneider, B. H. (1988). Perceived Competence Scale for Children: Testing for factorial validity and invariance across age and ability. *Applied Measurement in Education, 1*, 171-187.

Byrne, B. M., & Shavelson, R. J. (1986). On the structure of adolescent self-concept. *Journal of Educational Psychology, 78*, 474-481.

Byrne, B. M., & Shavelson, R. J. (1987). Adolescent self-concept: Testing the assumption of equivalent structure across gender. *American Educational Research Journal, 24*, 365-385.

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 (1976) Model. *Journal of Personality and Social Psychology, 70*(3), 599-613.

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105*, 456-466.

Byrne, B. M., & Worth Gavin, D. A. (1996). The Shavelson model revisited: Testing for the structure of academic self-concept across pre-, early, and late adolescents. *Journal of Educational Psychology, 88*(2), 215-228.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.

Chapman, J. W., & Tunmer, W. E. (1995). Development of young children's reading self-concepts: An examination of emerging subcomponents and their relationship with reading achievement. *Journal of Educational Psychology, 87*(1), 154-167.

Coopersmith, S. A. (1967). *The antecedents of self-esteem*. San Francisco: W. H. Freeman.

Crain, R. M., & Bracken, B. A. (1994). Age, race, and gender differences in child and adolescent self-concept: Evidence from a behavioral-acquisition, context-dependent model. *School Psychology Review, 23*(3), 496-511.

Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), *Educational Measurement* (2nd ed., pp. 443-507). Washington, DC: American Council on Education.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281-302.

Dance, K. A., & Kuiper, N. A. (1987). Self-schemata, social roles, and a self-worth contingency model of depression. *Motivation and Emotion*, *11*(3), 251-268.

Delugach, R. R., Bracken, B. A., Bracken, M. J., & Schicke, M. C. (1992). Self-concept: Multidimensional construct exploration. *Psychology in the Schools*, *29*, 213-223.

Fleming, J. S., & Watts, W. A. (1980). The dimensionality of self-esteem: Some results for a college sample. *Journal of Personality and Social Psychology*, *39*, 921-929.

Frey, K. S., & Ruble, D. N. (1987). What children say about classroom performance: Sex and grade differences in perceived competence. *Child Development*, *58*, 1066-1078.

Gerbing, D. W., & Anderson, J. C. (1993). Monte Carlo evaluations of goodness-of-fit indices for structural equation models. In K. A. Bollen, & J. S. Long (Eds.), *Testing Structural Equation Models* (pp. 40-65). Newbury Park, CA: Sage.

Hammen, C., Marks, T., deMayo, R., & Mayol, A. (1985). Self-schemas and risk for depression: A prospective study. *Journal of Personality and Social Psychology*, *49*(5), 1147-1159.

Hansford, B. C., & Hattie, J. A. (1982). The relationship between self and achievement/performance measures. *Review of Educational Research*, *52*, 123-142.

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 (Ed.), *Handbook of child psychology* (4th ed., Vol. 4, pp. 275-385). New York: Wiley.

Harter, S. (1985). *Manual for the Self-perception Profile for Children*. Denver, CO: University of Denver.

Harter, S. (1990). Issues in the assessment of the self-concept of children and adolescents. In A. M. La Greca (Ed.), *Through the eyes of the child: Obtaining self-reports from children and adolescents* (pp. 292-325). Boston: Allyn & Bacon.

Harter, S. (1996). Historical roots of contemporary issues involving self-concept. In B. A. Bracken (Ed.), *Handbook of self-concept: Developmental, social, and clinical considerations* (pp. 1-37). New York: Wiley.

Hattie, J., & Marsh, H. W. (1996). Future directions in self-concept research. In B. A. Bracken (Ed.), *Handbook of self-concept: Developmental, social, and clinical considerations* (pp. 421-462). New York: Wiley.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, *94*(3), 319-340.

Higgins, E. T., Bond, R. N., Klein, R., & Strauman, T. (1986). Self-discrepancies and emotional vulnerability: How magnitude, accessibility, and type of discrepancy influence affect. *Journal of Personality and Social Psychology*, *51*(1), 5-15.

Higgins, E. T., Klein, R., & Strauman, T. J. (1985). Self-concept discrepancy theory: A psychological model for distinguishing among different aspects of depression and anxiety. *Social Cognition*, *3*(1), 51-76.

Higgins, E. T., Klein, R. L., & Strauman, T. J. (1987). Self-discrepancies: Distinguishing among self-states, self-state conflicts, and emotional vulnerabilities. In K. Yardley & T. Honess (Eds.), *Self and identity: Psychosocial perspectives* (pp. 173-186). New York: Wiley.

Hoge, D. R., & McCarthy, J. D. (1984). Influence of individual and group identity salience in the global self-esteem of youth. *Journal of Personality and Social Psychology*, *47*, 403-414.

Honess, T., & Yardley, K. (1987). Cognitive, affective, and contextual aspects of self: An introductory review. In K. Yardley & T. Honess (Eds.), *Self and identity: Psychosocial perspectives* (pp. 149-156). New York: Wiley.

Horn, W. F., & Packard, T. (1985). Early identification of learning problems: A meta-analysis. *Journal of Educational Psychology*, *77*, 597-607.

Hoyle, R. H. (1995). The structural equation modeling approach: Basic concepts and fundamental issues. In R. H. Hoyle (Ed.), *Structural Equation Modeling: Concepts, issues, and applications* (pp. 1-15). Thousand Oaks, CA: Sage.

Hoyle, R. H., & Panter, A. T. (1995). Writing about structural equation models. In R. H. Hoyle (Ed.), *Structural Equation Modeling: Concepts, issues, and applications* (pp. 158-176). Thousand Oaks, CA: Sage.

Hu, L., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural Equation Modeling: Concepts, issues, and applications* (pp. 76-99). Thousand Oaks, CA: Sage.

- Hu, L., Bentler, P. M., & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin*, *112*(2), 351-362.
- Hyde, J. S., Fennema, E., Ryan, M., Frost, L. A., & Hopp, C. (1990). Gender comparisons of mathematics attitudes and affect: A meta-analysis. *Psychology of Women Quarterly*, *14*, 299-324.
- James, W. (1890/1950). *The Principles of Psychology*. New York: Holt, Rinehart & Winston.
- Joreskog, K. G., & Sorbom, D. (1993). *LISREL VIII: User's reference guide*. Chicago: Scientific Software.
- Keith, L. K., & Bracken, B. A. (1996). Self-concept instrumentation: A historical and evaluative review. In B. A. Bracken (Ed.), *Handbook of self-concept: Developmental, social and clinical considerations* (pp. 91-170). New York: Wiley.
- Kuiper, N. A., & Higgins, E. T. (1985). Social cognition and depression: A general integrative perspective. *Social Cognition*, *3*(1), 1-15.
- La Du, T. J., & Tanaka, J. S. (1989). Influence of sample size, estimation method, and model specification on goodness-of-fit assessments in structural equation models. *Journal of Applied Psychology*, *74*(4), 625-635.
- Linville, P. W. (1985). Self-complexity and affective extremity: Don't put all of your eggs in one cognitive basket. *Social Cognition*, *3*(1), 94-120.
- Linville, P. W. (1987). Self-complexity as a cognitive buffer against stress-related illness and depression. *Journal of Personality and Social Psychology*, *52*(4), 663-676.
- Long, J. S. (1983a). *Confirmatory Factor Analysis: A preface to LISREL*. (Volume 33). Beverly Hills: Sage.
- Long, J. S. (1983b). *Covariance Structure Models: An Introduction to LISREL*. (Volume 34). Beverly Hills: Sage.
- MacCallum, R. (1986). Specification searches in covariance structure modeling. *Psychological Bulletin*, *100*(1), 107-120.
- MacCallum, R. (1995). Model specification: Procedures, strategies, and related issues. In R. H. Hoyle (Ed.), *Structural Equation Modeling: Concepts, issues, and applications* (pp. 16-36). Thousand Oaks, CA: Sage.

- MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, 111(3), 490-504.
- MacCallum, R. C., Wegener, D. T., Uchino, B. N., & Fabrigar, L. R. (1993). The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin*, 114(1), 185-199.
- Markus, H. (1977). Self-schemata and processing information about the self. *Personality and Social Psychology*, 35(2), 63-78.
- Markus, H. (1983). Self-knowledge: An expanded view. *Journal of Personality*, 51(3), 543-565.
- Markus, H. (1990). Unresolved issues of self-representation. *Cognitive Therapy and Research*, 14(2), 241-253.
- Markus, H., Cross, S., & Wurf, E. (1990). The role of the self-system in competence. In R. J. Sternberg & J. Kolligian (Eds.), *Competence considered* (pp. 205-225). London: Yale University.
- Markus, H., & Kunda, Z. (1986). Stability and malleability of the self-concept. *Journal of Personality and Social Psychology*, 51(4), 858-866.
- Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, 41(9), 954-969.
- Markus, H., & Nurius, P. (1987). Possible selves: The interface between motivation and self-concept. In K. Yardley & T. Honess (Eds.), *Self and identity: Psychological perspectives* (pp. 157-172). New York: Wiley.
- Markus, H., & Ruvolo, A. (1989). Possible selves: Personalized representations of goals. In L. A. Pervin (Ed.), *Goal concepts in personality and social psychology* (pp. 211-241). Hillsdale, NJ: Erlbaum.
- Markus, H., & Sentis, K. (1982). The self in social information processing. In J. Suls (Ed.), *Psychological perspectives on the self* (Vol. 1, pp. 41-58). Hillsdale, NJ: Erlbaum.
- Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology*, 38, 299-337.

- Marsh, H. W. (1984). Self-concept: The application of a frame of reference model to explain paradoxical results. *Australian Journal of Education*, **28**, 165-181.
- Marsh, H. W. (1985). Age and sex effects in multiple dimensions of preadolescent self-concept: A replication and extension. *Australian Journal of Psychology*, **37**(2), 197-204.
- Marsh, H. W. (1986a). Global self-esteem: Its relation to specific facets of self-concept and their importance. *Journal of Personality and Social Psychology*, **51**(6), 1224-1236.
- Marsh, H. W. (1986b). Negative item bias in rating scales for preadolescent children: A cognitive-developmental phenomenon. *Developmental Psychology*, **22**, 37-49.
- Marsh, H. W. (1986c). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal*, **23**, 129-149.
- Marsh, H. W. (1987a). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology*, **79**, 280-295.
- Marsh, H. W. (1987b). The factorial invariance of responses by males and females to a multidimensional self-concept instrument: Substantive and methodological issues. *Multivariate Behavioral Research*, **22**, 457-480.
- Marsh, H. W. (1987c). The hierarchical structure of self-concept and the application of hierarchical confirmatory factor analysis. *Journal of Educational Measurement*, **24**(1), 17-39.
- Marsh, H. W. (1988). *The Self Description Questionnaire-I: SDQ manual and research monograph*. San Antonio, TX: The Psychological Corporation.
- Marsh, H. W. (1989a). Age and sex effects in multiple dimensions of self-concept: Preadolescence to early adulthood. *Journal of Educational Psychology*, **81**(3), 417-430.
- Marsh, H. W. (1989b). Sex differences in the development of verbal and mathematics constructs: The High School and Beyond study. *American Educational Research Journal*, **26**, 191-225.
- Marsh, H. W. (1990a). A multidimensional, hierarchical model of self-concept: Theoretical and empirical justification. *Educational Psychology Review*, **2**, 77-172.

- Marsh, H. W. (1990b). Confirmatory factor analysis of multitrait-multimethod data: The construct validation of multidimensional self-concept responses. *Journal of Personality*, **58**(4), 661-692.
- Marsh, H. W. (1990c). Influences of internal and external frames of reference on the formation of math and English self-concepts. *Journal of Educational Psychology*, **82**(1), 107-116.
- Marsh, H. W. (1990d). The structure of academic self-concept: The Marsh/Shavelson model. *Journal of Educational Psychology*, **82**(4), 623-636.
- Marsh, H. W. (1992a). Content specificity of relations between academic achievement and academic self-concept. *Journal of Educational Psychology*, **84**(1), 35-42.
- Marsh, H. W. (1992b). *Self Description Questionnaire (SDQ) I: A theoretical and empirical basis for the measurement of multiple dimensions of preadolescent self-concept: A test manual and research monograph*. Macarthur, NSW Australia: Faculty of Education, University of Western Sydney.
- Marsh, H. W. (1992c). *Self Description Questionnaire (SDQ) II: A theoretical and empirical basis for the measurement of multiple dimensions of adolescent self-concept: An interim test manual and research monograph*. Macarthur, NSW Australia: Faculty of Education, University of Western Sydney.
- Marsh, H. W. (1992d). *Self Description Questionnaire (SDQ) III: A theoretical and empirical basis for the measurement of multiple dimensions of late adolescent self-concept: An interim test manual and research monograph*. Macarthur, NSW Australia: Faculty of Education, University of Western Sydney.
- Marsh, H. W. (1993a). Academic self-concept: Theory, measurement, and research. In J. Suls (Ed.), *Psychological perspectives on the self* (Vol. 4, pp. 59-98). Hillsdale NJ: Erlbaum.
- Marsh, H. W. (1993b). The multidimensional structure of academic self-concept: Invariance over gender and age. *American Educational Research Journal*, **30**(4), 841-860.
- Marsh, H. W. (1996a). Construct validity of Physical Self-Description Questionnaire responses: Relations to external criteria. *Journal of Sport & Exercise Psychology*, **18**, 111-131.
- Marsh, H. W. (1996b). Physical Self Description Questionnaire: Stability and discriminant validity. *Research Quarterly for Exercise and Sport*, **67**(3), 249-264.

Marsh, H. W. (1997). The measurement of physical self-concept: A construct validation approach. In K. R. Fox (Ed.), *The physical self: From motivation to well-being* (pp. 27-58). Champaign, IL: Human Kinetics.

Marsh, H. W., Barnes, J., Cairns, L., & Tidman, M. (1984). Self-Description Questionnaire: Age and sex effects in the structure and level of self-concept for preadolescent children. *Journal of Educational Psychology*, *76*, 940-956.

Marsh, H. W., Barnes, J., & Hocevar, D. (1985). Self-other agreement on multidimensional self-concept ratings: Factor analysis and multitrait-multimethod analysis. *Journal of Personality and Social Psychology*, *49*(5), 1360-1377.

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., & Craven, R. G. (1991). Self-other agreement on multiple dimensions of preadolescent self-concept: Inferences by teachers, mothers, and fathers. *Journal of Educational Psychology*, *83*(3), 393-404.

Marsh, H. W., & Gouvernet, P. J. (1989). Multidimensional self-concepts and perceptions of control: Construct validation of responses by children. *Journal of Educational Psychology*, *81*(1), 57-69.

Marsh, H. W., & Hattie, J. (1996). Theoretical perspectives on the structure of self-concept. In B. A. Bracken (Ed.), *Handbook of self-concept: Developmental, social, and clinical considerations* (pp. 38-90). New York: Wiley.

Marsh, H. W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First- and higher-order factor models and their invariance across groups. *Psychological Bulletin*, *97*, 562-582.

Marsh, H. W., & MacDonald Holmes, I. W. (1990). Multidimensional self-concepts: Construct validation of responses by children. *American Educational Research Journal*, *27*(1), 89-117.

Marsh, H. W., & O'Neill, R. (1984). Self-Description Questionnaire III: The construct validity of multidimensional self-concept ratings by late adolescents. *Journal of Educational Measurement*, *21*, 153-174.

Marsh, H. W., & Parker, J. W. (1984). Determinants of self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? *Journal of Personality and Social Psychology*, *47*, 213-231.

Marsh, H. W., & Parker, J., & Barnes, J. (1985). Multidimensional adolescent self-concepts: Their relationship to age, sex, and academic measures. *American Educational Research Journal*, *22*, 422-444.

Marsh, H. W., & Parker, J., & Smith, I. D. (1983). Preadolescent self-concept: Its relation to self-concept as inferred by teachers and to academic ability. *British Journal of Educational Psychology*, *53*, 60-78.

Marsh, H. W., Relich, J. D., & Smith, I. D. (1983). Self-concept: The construct validity of interpretations based upon the SDQ. *Journal of Personality and Social Psychology*, *45*(1), 173-187.

Marsh, H. W., Richards, G. E., Johnson, S., Roche, L., & Tremayne, P. (1994). Physical Self-Description Questionnaire: Psychometric properties and a multitrait-multimethod analysis of relations to existing instruments. *Journal of Sport & Exercise Psychology*, *16*, 270-305.

Marsh, H. W., & Shavelson, R. J. (1985). Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist*, *20*, 107-125.

Marsh, H. W., & Smith, I. D. (1982). Multitrait-multimethod analyses of two self-concept instruments. *Journal of Educational Psychology*, *74*, 430-440.

Marsh, H. W., & Smith, I. D. (1987). A cross-national study of the structure and level of multidimensional self-concepts: An application of confirmatory factor analysis. *Australian Journal of Psychology*, *39*, 61-77.

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*(3), 333-357.

Marsh, H. W., Smith, I. D., & Barnes, J. (1984). Multidimensional self-concepts: Relationships with inferred self-concepts and academic achievement. *Australian Journal of Psychology*, *36*(3), 367-386.

Marsh, H. W., Smith, I. D., & Barnes, J. (1985). Multidimensional self-concept: Relations with sex and academic achievement. *Journal of Educational Psychology*, *77*, 581-596.

Marsh, H. W., Smith, I. D., Barnes, J., & Butler, S. (1983). Self-concept: Reliability, stability, dimensionality, validity, and the measurement of change. *Journal of Educational Psychology, 75*(5), 772-790.

Marx, R. W., & Winne, P. H. (1978). Construct interpretations of three self-concept inventories. *American Educational Research Journal, 15*, 99-108.

Maughan, B., Pickles, A., Hagell, A., Rutter, M., & Yule, W. (1996). Reading problems and antisocial behaviour: Developmental trends in comorbidity. *Journal of Child Psychology and Psychiatry, 37*(4), 405-418.

Meece, J. L., Parsons, J. E., Kaczala, C. M., Goff, S. B., & Futterman, R. (1982). Sex differences in math achievement: Toward a model of academic choice. *Psychological Bulletin, 91*, 324-348.

Mulaik, S. A., James, L. R., Van Alstine, J., Bennett, N., Lind, S., & Stilwell, C. D. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin, 105*(3), 430-445.

Muthén, B., & Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology, 38*, 171-189.

Muthén, B., Kaplan, D., & Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. *Psychometrika, 52*(3), 431-462.

Nurius, P. S. (1986). Reappraisal of the self-concept and implications for counseling. *Journal of Counseling Psychology, 33*(4), 429-438.

Nurius, P. S., & Markus, H. (1990). Situational variability in the self-concept: Appraisals, expectancies, and asymmetries. *Journal of Social and Clinical Psychology, 9*(3), 316-333.

Piers, E. V. (1984). *Piers-Harris Children's Self-Concept Scale: Revised manual*. Los Angeles: Western Psychological Services.

Pietromonaco, P. R. (1985). The influence of affect on self-perception in depression. *Social Cognition, 3*(1), 121-134.

Satorra, A., & Bentler, P. M. (1988). Scaling corrections for chi-square statistics in covariance structure analysis. *ASA Proceedings, Business and Economic Statistics Section* (pp. 393-398). Washington, DC: American Statistical Association.

Satorra, A., & Bentler, P. M. (1994). Corrections to test statistic and standard errors in covariance structure analyses. In A. Von Eye & C. C. Clogg (Eds.), *Analysis of latent variables in developmental research* (pp. 399-419). Newbury Park, CA: Sage.

Satorra, A., & Saris, W. E. (1985). The power of the likelihood ratio test in covariance structure analysis. *Psychometrika*, *50*, 83-90.

Shavelson, R. J., & Bolus, R. (1982). Self-concept: The interplay of theory and methods. *Journal of Educational Psychology*, *74*, 3-17.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, *46*, 407-441.

Shavelson, R. J., & Marsh, H. W. (1986). On the structure of self-concept. In R. Schwarzer (Ed.), *Anxiety and cognitions* (pp. 305-330). New York: Erlbaum.

Silon, E. L., & 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.

Silvia, E. S. M., & MacCallum, R. C. (1988). Some factors affecting the success of specification searches in covariance structure modeling. *Multivariate Behavioral Research*, *23*, 297-326.

Singer, J. L., & Kolligian, J., Jr. (1987). Personality: Developments in the study of private experience. *Annual Review of Psychology*, *38*, 533-574.

Sink, C. A., Barnett, J. E., & Pool, B. A. (1993). Perceptions of scholastic competence in relation to middle-school achievement. *Perceptual and Motor Skills*, *76*, 471-478.

Smart, D., Sanson, A., & Prior, M. (1996). Connections between reading disability and behavior problems: Testing temporal and causal hypotheses. *Journal of Abnormal Child Psychology*, *24*(3), 363-383.

Smith, I. D., & Marsh, H. W. (1985). A cross-cultural study of self-concept in primary school children. *Collected papers of the 1985 Annual Conference of the Australian Association for Research in Education* (pp. 350-353).

Song, I. S. (1982). *The dimensionality and relationships between home environment, self-concept and academic achievement*. Unpublished doctoral dissertation, University of New England.

Song, I. S., & Hattie, J. (1985). Relationship between self-concept and achievement. *Journal of Research in Personality, 19*, 365-372.

Steiger, J. H., & Lind, J. M. (1980). *Statistically based tests for the number of common factors*. Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.

Stipek, D. J., & Daniels, D. H. (1988). Declining perceptions of competence: A consequence of changes in the child or in the educational environment? *Journal of Educational Psychology, 80*(3), 352-356.

Stipek, D. J., & MacIver, D. (1989). Developmental change in children's assessment of intellectual competence. *Child Development, 60*, 521-538.

Sugawara, H. M., & MacCallum, R. C. (1993). Effect of estimation method on incremental fit indexes for covariance structure models. *Applied Psychological Measurement, 17*(4), 365-377.

Tanaka, J. S. (1987). "How big is big enough?": Sample size and goodness of fit in structural equation models with latent variables. *Child Development, 58*, 134-146.

Tanaka, J. S. (1993). Multifaceted conceptions of fit in structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 10-39). Newbury Park, CA: Sage.

Tanzer, N., Marsh, H., & Sim, C. (July, 1992). *Self-concept of Singaporean children. A cross-cultural comparison*. Poster presented at the 25th International Congress of Psychology, Brussels, Belgium.

Tanzer, N., & Sim, C. (September, 1991). *Self-concept and achievement attributions. A study of Singaporean primary school students*. Unpublished research report, Institut für Psychologie der Karl-Franzens-Universität Graz.

Van Dongen-Melman, J. E. W. M., Koot, H. M., & Verhulst, F. C. (1993). Cross-cultural validation of Harter's Self-Perception Profile for Children in a Dutch sample. *Educational and Psychological Measurement, 53*, 739-753.

Vispoel, W. P. (1995). Self-concept in artistic domains: An extension of the Shavelson, Hubner, and Stanton (1976) model. *Journal of Educational Psychology, 87*(1), 134-153.

Watkins, D., & Akande, A. (1992). The internal structure of the Self Description Questionnaire: A Nigerian investigation. *British Journal of Educational Psychology*, *62*, 120-125.

Watkins, D., & Gutierrez, M. (1989). The structure of self-concept: Some Filipino evidence. *Australian Psychologist*, *24*, 401-410.

Watkins, D., Lam, M. K., & Regmi, M. (1991). Cross-cultural assessment of self-esteem: A Nepalese investigation. *Psychologia*, *34*, 98-108.

Watkins, D., & Mpofu, E. (1994). Some Zimbabwean evidence of the internal structure of the Self Description Questionnaire-1. *Educational and Psychological Measurement*, *54*(4), 967-972.

Wells, L. E., & Marwell, G. (1976). *Self-esteem: Its conceptualization and measurement*. Beverly Hills, CA: Sage.

West, S. G., Finch, J. F., & Curran, P. J. (1996). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural Equation Modeling: Concepts, issues, and applications* (pp. 56-75). Thousand Oaks, CA: Sage.

West, C. K., Fish, J. A., & Stevens, R. J. (1980). General self-concept, self-concept of academic ability and school achievement: Implications for "causes" of self-concept. *Australian Journal of Education*, *24*, 194-213.

Wheaton, B. (1988). Assessment of fit in overidentified models with latent variables. In J. S. Long (Ed.), *Common problems/proper solutions: Avoiding error in quantitative research* (pp. 193-225). Newbury Park, CA: Sage.

Williams, L. J. & Holahan, P. J. (1994). Parsimony-based fit indices for multiple-indicator models: Do they work? *Structural equation modeling*, *1*(2), 161-189.

Winne, P. H., Marx, R. W., & Taylor, T. D. (1977). A multitrait-multimethod study of three self-concept inventories. *Child Development*, *48*, 893-901.

Wurf, E., & Markus, H. (1991). Possible selves and the psychology of personal growth. In R. Hogan (Series Ed.) & D. Ozer, J. M. Healy, Jr., & A. J. Stewart, (Vol Eds.), *Perspectives in Personality: A research annual, Vol. 3 (Part A). Self and emotion* (pp. 39-62). London: Jessica Kingsley.

Wylie, R. C. (1974). *The self-concept: A review of methodological considerations and measuring instruments*. (Rev. ed., Vol. 1). Lincoln, NE: University of Nebraska Press.

Wylie, R. C. (1979). *The self-concept (Rev. ed., Vol. 2): Theory and research on selected topics*. Lincoln, NE: University of Nebraska Press.

**APPENDIX A: INFORMED CONSENT ORAL PRESENTATIONS,
AND PARENTAL PERMISSION LETTER AND FORMS**

**Student Oral Presentation: Request to Obtain Student Consent
and to Forward Parental Permission Letter and Forms**

I am here today to tell you about some work that we are doing at the University of Ottawa, and to ask you if you would like to help us with this work. We would like to find out about two types of things. First, we would like to know how you think about yourself in a variety of different areas, such as, how you do in school, how you look, and how you get along with your friends, or with your parents. Second, we would like to find out what kind of grades you get in different school subjects, and how important doing well in each of these subjects is to you.

It will take us about two class visits to ask you these questions. These questions are not part of a test; there are no right or wrong answers. In addition, we would like to get a copy of your most recent grades from your school. We will keep your answers and your grades private and will not show them to anyone.

Participation in this study is completely voluntary. You do not have to participate, and if you decide to participate, you can stop participating at any time, or refuse to answer any questions you do not want to answer. If you would like to participate, or you do not want to participate, or you change your mind about participating, this will not affect your school marks.

Does anyone have any questions? Would you like to participate? I will pass out a sheet which has all of your names typed on it. Beside each of your names is a blank line. If you would like to participate, please sign your name on the blank line beside your name. If you do not want to participate, you can put an "X", or the word "No" on the line beside your name. If you would like to participate, I have a hand-out which I would like you to bring to your parents today. The hand-out includes a letter to your parents explaining the study we are doing, and two permission forms for them to indicate whether they agree to your participating in this study or not. I would like you to return one of the permission forms to your teacher later this week (or next week, if presented towards the end of the week), after your parents have had a chance to complete it and sign it. Your parents can keep the letter explaining the study and one of the permission forms. Are there any questions?

OK, here is the sheet with the list of your names. When you have finished with the sheet, please pass it on to the next student. (The form containing the list of students is passed out). As that sheet is going around the class, I will pass out the handout for your parents. If you would like to participate, please take a copy for your parents and pass the rest on to the next student. If you do not want to participate, you can simply pass the handout on to the next student. (The hand-out is passed out).

For those of you who want to participate, please remember to give the handout to your parents today and to return one of the permission forms, after it has been completed and signed by one of your parents, to your teacher by the end of the/next week. Thanks a lot for your help. I will be back in the next few weeks to ask you the types of questions I told you about today.

First Administration Session: Oral Presentation to Students

I came to your class not long ago to tell you about some work that we are doing at the University of Ottawa, and to ask you if you would like to help us with this work. I explained to you then that we would like to find out about two types of things. First, we would like to know how you think about yourself in a variety of different areas, such as, how you do in school, how you look, and how you get along with your friends, or with your parents. Second, we would like to find out what kind of grades you get in different school subjects, and how important doing well in each of these subjects is to you.

I also explained to you that it would take us about two class visits to ask you these questions. These questions are not part of a test; there are no right or wrong answers. In addition, we would like to get a copy of your most recent grades from your school. We will keep your answers and your grades private and will not show them to anyone.

Participation in this study is completely voluntary. You do not have to participate, and if you decide to participate, you can stop participating at any time, or refuse to answer any questions you do not wish to answer. If you would like to participate, or you do not want to participate, or you change your mind about participating, this will not affect your school marks.

Do you have any questions? If you wish to participate and you have obtained permission from your parents to do so, we will read the questions together. Those of you who are not participating in the study can do some schoolwork. Everyone must be perfectly quiet while I am reading the questions. OK?

Parent Permission Letter

Dr. Barbara M. Byrne
Lisa Larocque
School of Psychology
University of Ottawa
Ottawa, Ontario K1N 6N5
231-7925/564-6875

Dear Parent(s):

As researchers from the University of Ottawa, we are conducting a study designed to understand how students' academic performance is related to their academic, social, and physical perceptions of themselves. Students who volunteer to participate will be asked to complete two types of questionnaires (**not tests**). The first type of questionnaire, of which there are three, will tap children's perceptions of self in a variety of social, physical, and academic areas. The second type of questionnaire will ask children to report their most recent school performance in a variety of academic areas, and to rate how important doing well in each of these academic areas is to them. These questionnaires will be completed during regular school hours, sometime during October, or November, 1991^a, and should take no more than one 30-minute and one 45-minute class period. Whenever possible, these two sessions will be held on the same day; recess or the lunch break will separate the two sessions. Alternately, the two sessions will take place a maximum of one week apart.

As part of the study, we also request your permission to obtain your child's most recent grades from his/her school. These grades, together with the questionnaires, will assist us in understanding how childrens' school performance contributes to their self-perceptions in particular academic areas.

All records will be kept strictly confidential; only the investigators (Dr. Barbara Byrne and Lisa Larocque) and trained research assistants will have access to the information. Furthermore, all information will be examined on a group basis, thereby ensuring that individual students cannot be identified.

Participation is completely voluntary. Students who do not wish to take part in the study will not be penalized. Those students who do not wish to take part in the study will be allowed to engage in an alternate school-related activity, possibly in another classroom. Students who agree to participate may discontinue their participation, or refuse to answer any questions at any time and for any reason. In any case, if children do not wish to participate in this study, either in whole or in part, this will not affect their school marks.

If you would like your son or daughter to participate, please sign one of the attached permission forms and return it to your child for his/her teacher by Friday; you may keep the other form for your records. Should you wish to obtain a summary of the results, please indicate this on the permission from you are returning to us with your child, and include a return address.

Thank you in advance for your assistance and cooperation. Should you have further questions concerning this project, please feel free to call Lisa Larocque or Dr. Barbara Byrne at 231-7925.

Parent Permission Form

Student's Name: _____

Grade: _____

Age: _____

Gender: Male _____ Female _____

Teacher's name: _____

School: _____

_____ I would like my son/daughter to take part in the study described above

_____ I would not like my son/daughter to take part in the study described above

(Signature of parent/guardian)

(Date)

(Signature of witness)

(Date)

_____ I do not wish to receive a summary of the study

_____ I would like to receive a summary of the study at this address:

APPENDIX B:
ACADEMIC SUBJECT AREA DESCRIPTIONS OR GRADING CATEGORIES
PROVIDED BY THE PARTICIPATING SCHOOL BOARDS

<u>Academic subject area</u>	School Board A	School Board B	School Board C
LLA	listen to and extract information from oral sources; acquire and demonstrate ability to listen critically; listen to and increase range of prose and poetry		listens attentively for a period of time; listens to follow instructions; listens to others during group discussion
SLA	express ideas, thoughts and feelings through speech, participate in discussion (large and small groups); appraise and improve personal speech habits (oral reading)		participates in group discussions; expresses an idea clearly
RLA	refine and extend reading skills; use reading as a source of information and vocabulary development; develop an appreciation for different forms of literature (poetry, stories, myths)		reads a variety of material; reads with fluency and expression; is able to follow written instructions; reads with understanding

Appendix B continued.

Academic subject area	School Board A	School Board B	School Board C
WLA	develop vocabulary and writing fluency; adapt writing style to suit purposes and audience; extend understanding of punctuation and grammar; demonstrate a growing knowledge of the patterns of spelling; write fluently and legibly using correct letter formation		expresses ideas and feelings effectively; uses appropriate sentence structure; attempts to use new words; applies mechanics of writing; writes legibly; enjoys writing
ARM	master relevant number facts; add, subtract, multiply, and divide; use decimals, fractions, and simple algebraic notations; draw conclusions from evidence obtained through experimentation or logical reasoning	basic operational skills using whole numbers, decimals, and fractions	computation

Appendix B continued.

Academic subject area	School Board A	School Board B	School Board C
MT	estimate with a reasonable degree of precision; match, compare and order length, area, volume, mass, time, and temperature; acquire and demonstrate measuring skills using instrument; solve meaningful problems involving measurement and calculation	problem solving techniques	measurement
SN	(Canadian Studies, Geography, Science, Health). Explore and investigate their social and scientific environment; refine and extend ability to describe environment using observing, communicating and interpreting techniques; refine and extend understanding of relationships in environment	study of the parts and systems of plants, animals, and insects; concepts of matter and energy; study of the environment	natural science

Appendix B continued.

Academic subject area	School Board A	School Board B	School Board C
SST	study of man in his environment, culture, globalism, spacial interaction; map and globe skills	study of ancient and modern civilizations and culture in both Eastern and Western hemispheres; map and globe skills; government study	social studies
REL	become familiar with the life and mission of Jesus; increase familiarity with the scripture; enter more fully into the life of the church, particularly sacramental life; continue growth in personal faith; understand and appreciate that people are: created by God, sexual, living in relationships, called to commitment, growing persons	emphasis on moral and Christian formation; presentation of Jesus through the study of the Gospels and the church's liturgical celebrations	religion or moral education
ART	Extends concepts of space, color and pattern; experience a variety of materials and techniques	develop an awareness of elements of design (e.g., line, shape, color, texture, unity); presentation of techniques of painting, printing, drawing, construction, display	Arts

Appendix B continued.

Academic subject area	School Board A	School Board B	School Board C
GYM	develop physical fitness and an understanding of movement; develop a positive attitude toward vigorous physical activity; develop sportsmanship	major game skills, movement principles, and social experiences are developed through the knowledge and practices common to games, gymnastics, and dance	physical education

Note. L/S/R/WLA = Listening/Speaking/Reading/Writing in Language Arts, respectively; ARM = Arithmetic; MT = Measurement; SN = Science; SST = Social Studies; REL = Religion.

APPENDIX C: INSTRUMENT STANDARDIZED INSTRUCTIONS,
RESPONSE FORMAT, AND ITEMS

Standardized Instructions and Response Format for
the Self-Description Questionnaire-I (SDQ-I) and
the Academic Self-Description Questionnaire-I (ASDQ-I)^a
(Marsh, 1988; Marsh, 1990d)

This is a chance to look at yourself. It is not a test. There are no right answers and everyone will have different answers. Be sure that your answers show how you feel about yourself. **PLEASE DO NOT TALK ABOUT YOUR ANSWERS WITH ANYONE ELSE.** We will keep your answers private and not show them to anyone.

I will be reading the sentences aloud to you. I am presenting the material this way instead of having you read them to yourselves in order that everyone spends the same amount of time on each question.

When you are ready to begin, please read each sentence and choose an answer. (You may read quietly to yourself as I read aloud). There are five possible answers for each question: "True", "False", and three answers in between. There are five boxes next to each sentence, one for each of the answers. The answers are written at the top of the boxes. Choose your answer to a sentence and make a check mark in the box under the answer you choose. **DO NOT** say your answer out loud or talk about it with anyone else.

Before you start, there are three examples below. A student, Bob, has already answered two of these sentences to show you how to do it. In the third example, you must choose your own answer and put in your own check mark.

EXAMPLES

	FALSE	MOSTLY FALSE	SOME- TIMES FALSE/ SOME- TIMES TRUE	MOSTLY TRUE	TRUE
	1	2	3	4	5
1. I like to read comic books 1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Bob checked the box under the answer “True”. This means that he really likes to read comic books. If Bob did not like to read comic books very much, he would have answered “FALSE”, or “MOSTLY FALSE”.

1. In general, I am neat and tidy . . 2	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
---	--------------------------	--------------------------	-------------------------------------	--------------------------	--------------------------

Bob answered “SOMETIMES FALSE, SOMETIMES TRUE” because he is not very neat, but he is not very messy either.

3. I like to watch T.V. 3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
-------------------------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------

For this sentence you have to choose the answer that is best for you. First you must decide if the sentence is “TRUE”, or “FALSE”, or somewhere in between. If you really like to watch T.V. a lot, you would answer “TRUE” by making a check mark in the last box. If you hate watching T.V., you would answer “FALSE” by making a check mark in the first box. If your answer is somewhere in between, then you would choose one of the other three boxes.

If you want to change an answer you have marked, you should cross out the check mark and put a new check mark in another box on the same line.

For all sentences be sure that your check mark is on the same line as the sentence you are answering. You should have one answer and only one answer for each sentence. Do not leave out any of the sentences.

We will be going quite fast, and you will have to mark your answer immediately. Then listen to the next sentence. If you fall behind, leave out the sentences you have not done. Listen to the sentence I am reading and answer that one. I will allow you time at the end to go back to any sentences that you have left out.

Once you have started, **PLEASE DO NOT TALK**. Turn over the page and begin.

^a Note that the standardized instructions were presented on one legal size sheet, and that the response format is as presented in the examples for the standardized instructions.

Standardized Instructions and Response Format
for the Self-Perception Profile for Children (SPPC)^a
(Harter, 1985)

I am interested in what each of you is like. The questions I will be asking you are not part of a test. There are no right or wrong answers. Since kids are very different from one another, each of you will be answering something different. Be sure that your answers show how you feel about yourself. **PLEASE DO NOT TALK ABOUT YOUR ANSWERS WITH ANYONE ELSE.** We will keep your answers private and will not show them to anyone.

First, I will explain how these questions work. There is an example below which we will do together. I'll read it outloud and you can follow along with me.

Really True for me	Sort of True for me		BUT		Sort of True for me	Really True for me
<input style="width: 40px; height: 25px;" type="checkbox"/>	<input style="width: 40px; height: 25px;" type="checkbox"/>	Some kids would rather play outdoors in their spare time	BUT	Other kids would rather watch T.V.	<input style="width: 40px; height: 25px;" type="checkbox"/>	<input style="width: 40px; height: 25px;" type="checkbox"/>

This question describes two kinds of kids, and we want to know which kids are most like **you**.

So, what I want you to decide first is whether **you** are more like the kids on the left side who would rather play outdoors, or whether you are most like the kids on the right side who would rather watch T.V.. Don't mark anything yet, but first decide which kind of kid is **most like you**, and go to that side of the sentence.

Now, the **second** thing I want you to think about, now that you have decided which kind of kids are most like you, is whether that is **sort of true for you**, or **really true for you**. If it's only sort of true, then put an X in the box under sort of true; if it's really true for you, then put an X in that box, under really true.

For each sentence you only check one box. Sometimes it will be on one side of the page, another time it will be on the other side of the page, but you can only check **one box** for each sentence. You **don't** check both sides, just the one side most like you.

Okay, that one was just for practice. Now we have some more sentences which I'm going to read out loud. For each one, just check one box, the one that goes with what is true for you, what you are most like. **DO NOT** say your answers out loud or talk about it with anyone else.

If you want to change an answer you have marked you should cross out the check mark and put a new check mark in another box on the same line. You should have one answer and one answer only for each question. Do not leave out any of the sentences.

We will be going quickly, and you will have to mark your answer immediately. Then listen to the next sentence. If you fall behind, leave out the sentences you have not done. Listen to the sentence I am reading and answer that one. I will allow you time at the end to go back to any sentences that you have left out.

Once you have started, **PLEASE DO NOT TALK**. Turn over the page and begin.

^a Note that the standardized instructions were presented on one legal size sheet, and that the response format is as presented in the examples for the standardized instructions.

Standardized Instructions and Response Format for the Harter Importance Rating Scale^a

(adapted from Harter, 1985)

I am interested in how important doing well in a variety of school subjects is to you. This is **not** a test. There are no right or wrong answers. Since kids are very different from one another, each of you will be answering something different. Be sure that your answers show how important doing well in the different school subjects is to you. **PLEASE DO NOT TALK ABOUT YOUR ANSWERS WITH ANYONE ELSE.** We will keep your answers private and will not show them to anyone.

First, I will explain how these questions work. There is an example below which we will do together. I'll read it out loud and you can follow along with me.

Really True for me	Sort of True for me		Sort of True for me	Really True for me		
<input type="checkbox"/>	<input type="checkbox"/>	Some kids think it is important to do well in games in order to feel good about themselves	BUT	Other kids don't think how well they do in games is that important	<input type="checkbox"/>	<input type="checkbox"/>

This question describes two kinds of kids, and we want to know which kids are most like **you**.

So, what I want you to decide first is whether **you** are more like the kids on the left side who think doing well in games is important, or whether you are most like the kids on the right side who think doing well in games is not that important. Don't mark anything yet, but first decide which kind of kid is **most like you**, and go to that side of the sentence.

Now, the **second** thing I want you to think about, now that you have decided which kind of kids are most like you, is whether that is **sort of true for you**, or **really true for you**. If it's only sort of true, then put an X in the box under sort of true; if it's really true for you, then put an X in that box, under really true.

For each sentence you only check one box. Sometimes it will be on one side of the page, another time it will be on the other side of the page, but you can only check **one box** for each sentence. You **don't** check both sides, just the **one side** most like you.

Okay, that one was just for practice. Now we have some more sentences which I'm going to read out loud. For each one, just check one box, the one that goes with what is true for you, what you are most like. **DO NOT** say your answer out loud or talk about it with anyone else.

If you want to change an answer you have marked you should cross out the check mark and put a new mark in another box on the same line. You should have one answer and one answer only for each question. Do not leave out any of the sentences.

We will be going quickly, and you will have to mark your answer immediately. Then listen to the next sentence. If you fall behind, leave out the sentences you have not done. Listen to the sentence I am reading and answer that one. I will allow you time at the end to go back to any sentences that you have left out.

One you have started, **PLEASE DO NOT TALK**. Turn over the page and begin.

^a Note that the standardized instructions were presented on one legal size sheet, and that the response format was as presented in the examples for the standardized instructions.

The Self-Description Questionnaire-I (SDQ-I) Items
(Marsh,1988)

Physical Abilities (PhAb) Scale Items:

1. I can run fast.
2. I have good muscles.
3. I am good at sports.
4. I can run a long way without stopping.
5. I am a good athlete.
6. I am good at throwing a ball.
7. I like to run and play hard.*
8. I enjoy sports and games.*

Physical Appearance (PhAp) Scale Items:

1. I am good looking.
2. I have a pleasant looking face.
3. I am a nice looking person.
4. Other kids think I am good looking.
5. I have a good looking body.
6. I am better looking than most of my friends.
7. I like the way I look.*
8. I have nice features like nose, and eyes and hair.*

Peer Relations (PEER) Scale Items:

1. I make friends easily.
2. I get along with kids easily.
3. I am easy to like.
4. Other kids want me to be their friend.
5. I am popular with kids my own age.
6. Most other kids like me.
7. I have lots of friends.*
8. I have more friends than most other kids.*

Parent Relations (PRT) Scale Items:

1. My parents understand me.
2. My parents like me.
3. My parents and I spend a lot of time together.
4. My parents are easy to talk to.
5. I get along well with my parents.
6. My parents and I have a lot of fun together.
7. I like my parents.*
8. If I have children of my own, I want to bring them up like my parents raised me.*

General-Self (GSC) Scale Items:

1. Overall, I have a lot to be proud of.
2. I can do things as well as most other people.
3. Other people think I am a good person.
4. A lot of things about me are good.
5. I'm as good as most other people.
6. When I do something, I do it well.
7. I do lots of important things.*
8. In general, I like being the way I am.*

* denotes excluded items (i.e., items that were not administered).

The Academic Self-Description Questionnaire-I (ASDQ-I) Items

(Marsh, 1990d)

1. Compared to others my age, I am good at _____.
2. I have always done well in _____.
3. I get good marks in _____.
4. I learn things quickly in _____.
5. Work in _____ is easy for me.
6. I'm hopeless when it comes to _____.

Note. The academic subject areas listed in Table 2 in the method section for each of the school boards were substituted once for each item. The word “school” was substituted for the items designed to measure General-Academic SC. As noted earlier, analyses did not include Item 6.

The Self-Perception Profile for Children (SPPC) Items

(Harter, 1985)

Social Acceptance (PEER) Scale Items:

7. Some kids find it hard to make friends BUT other kids find it's pretty easy to make friends.
8. Some kids have a lot of friends BUT Other kids don't have very many friends.
9. Some kids would like to have a lot more friends BUT Other kids have as many friends as they want.
10. Some kids are always doing things with a lot of kids BUT Other kids usually do things by themselves.
11. Some kids wish that more people their age liked them BUT Other kids feel that most people their age do like them.
12. Some kids are popular with other kids their age BUT Other kids are not very popular.

Athletic Competence (PhAb) Scale Items:

7. Some kids do very well at all kinds of sports BUT Other kids don't feel that they are very good when it comes to sports.
8. Some kids wish they could be a lot better at sports BUT Other kids feel they are good enough at sports.
9. Some kids think they could do well at just about any new sports activity they haven't tried before BUT Other kids are afraid they might not do well at sports they haven't tried before.
10. Some kids feel that they are better than others their age at sports BUT Other kids don't feel that they can play as well.
11. In games and sports, some kids usually watch instead of play BUT Other kids usually play rather than watch.
12. Some kids don't do well at new outdoor games BUT Other kids are good at new games right away.

Physical Appearance (PhAp) Scale Items:

7. Some kids are happy with the way they look BUT Other kids are not happy with the way they look.
8. Some kids are happy with their height and weight BUT Other kids wish their height or weight were different.
9. Some kids wish their body was different BUT Other kids like their body the way it is.
10. Some kids wish their physical appearance (how they look) was different BUT Other kids like their physical appearance the way it is.
11. Some kids wish that something about their face or hair looked different BUT Other kids like their face and hair the way it is.
12. Some kids think that they are good looking BUT Other kids think that they are not very good looking.

Global Self-Worth (GSC) Scale Items:

7. Some kids are often unhappy with themselves BUT Other kids are pretty pleased with themselves.
8. Some kids don't like the way they are leading their life BUT Other kids do like the way they are leading their life.
9. Some kids are usually happy with themselves as a person BUT Other kids are often not happy with themselves.
10. Some kids like the kind of person they are BUT Other kids often wish they were someone else.
11. Some kids are very happy being the way they are BUT Other kids wish they were different.
12. Some kids are not happy with the way they do a lot of things BUT Other kids think the way they do things is fine.

Scholastic Competence Scale Items:

7. Some kids feel that they are very good at (_____) BUT Other kids worry about whether they can do the school work assigned to them (in _____).
8. Some kids feel like they are just as smart as other kids their age (in _____) BUT Other kids aren't so sure and wonder if they are as smart (in _____).
9. Some kids are pretty slow in finishing their school work (in _____) BUT Other kids can do their school work (in _____) quickly.
10. Some kids do very well at (_____) BUT Other kids don't do well at (_____).
11. Some kids often forget what they learn in school BUT Other kids remember things easily in school.*
12. Some kids have trouble figuring out the answers in school BUT Other kids can almost always figure out the answers in school.*

Note. The academic subject areas listed in Table 2 in the method section for each of the school boards were substituted once for each of the Scholastic Scale items included in the present study. Where necessary, "school" was substituted for the items designed to measure General-Academic SC.

* denotes excluded items (i.e., items that were not administered).

Children's Report of School Grades

Try to remember, as best as you can, your most recent grades. There are five possible answers for each school subject: "exceptional", "unsatisfactory", and three answers in between. 1 is "exceptional", 5 is "unsatisfactory"^a. These numbers are the same numbers that your teacher used on your report card. Pick the number which best describes your most recent grades in each of the following areas:

GRADE POINT VALUE^a

	EXCEP- TIONAL	VERY GOOD	SATIS- FACTORY	EXPERIENCING DIFFICULTY	UNSATIS- FACTORY
	1	2	3	4	5
Social Studies	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Speaking in Language Arts . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Computation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Natural Science . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Writing in Language Arts . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Physical Education .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	EXCEP- TIONAL	VERY GOOD	SATIS- FACTORY	EXPERIENCING DIFFICULTY	UNSATIS- FACTORY
	1	2	3	4	5
Measurement	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Religion or Moral Education . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Reading in Language Arts	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Listening in Language Arts	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Arts	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Geometry	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
School in general	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

^a The particular descriptive labels utilized varied with the school board. That is, the labels utilized corresponded to those used by the particular school board. Thus, for example, one school board used the letters A to E, with the corresponding labels “very good”, “good”, “satisfactory”, “fair”, and “unsatisfactory”; the same category designation and labels were used for the students from this school board. The academic content areas measured within each of the school boards are listed in Table 2 in the method section.

Children's Ratings of the Importance of School Subjects

We would like you to tell us how important doing well in a variety of school subjects is to you. There are no right or wrong answers, so everyone will have different answers. There are five possible answers for each school subject: “**not at all** important to me”, “**extremely** important to me”, and three answers in between. 1 is “**not at all** important to me”, and 5 is “**extremely** important to me”. Pick the number which best describes how important doing well in each of the following areas is to you:

Importance Scale

	not at all important to me 1	a little important to me 2	important to me 3	quite important to me 4	extremely important to me 5
Speaking in Language Arts	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Religion or Moral Education . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Social Studies	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Reading in Language Arts . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Computation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Natural Science	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	not at all important to me 1	a little important to me 2	important to me 3	quite important to me 4	extremely important to me 5
Writing in Language Arts . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Physical Education .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Measurement	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Listening in Language Arts . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Arts	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Geometry	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
School in general	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Note. The academic content areas measured within each of the school boards are listed in Table 2 in the method section.

Harter Importance Rating Scale^a

(adapted from Harter, 1985)

Really True for me	Sort of True for me		BUT		Sort of True for me	Really True for me
1. <input type="checkbox"/>	<input type="checkbox"/>	Some kids think it is important to do well _____ in order to feel good about themselves	BUT	Other kids don't think how well they do _____ is that important	<input type="checkbox"/>	<input type="checkbox"/>
2. <input type="checkbox"/>	<input type="checkbox"/>	Some kids don't think that getting good grades in _____ is all that important to how they feel about themselves	BUT	Other kids think that getting good grades in _____ is important	<input type="checkbox"/>	<input type="checkbox"/>

^a For the General-Academic Importance scale, at school work (for Item 1) or school (for Item 2) was substituted; for the specific academic importance scales, the specific academic subject areas, as listed in Table 2 in the method section, were substituted (with the word "in" preceding the subject area for Item 1).

APPENDIX D: LATENT FACTOR CORRELATIONS AMONG FIRST-ORDER SELF-CONCEPT (SC) CONSTRUCTS
FOR MODEL 16G FOR GIRLS (N = 220) AND MODEL 16B FOR BOYS (N = 205)

	Core Academic SCs										NIC AC SC					Nonacademic SCs				
	ARM	MT	SN	SST	LLA	SLA	RLA	WLA	GA	REL	ART	GYM	PHAB	PHAB	PHAP	PRT	PEER			
Core Academic SCs																				
ARM	.669	.486	.525	.459	.422	.422	.465	.474	.650	.293	.209	.317	.359	.255	.342	.323				
MT	.758	-.486	.517	.432	.397	.380	.380	.432	.561	.227	.150*	.270	.320	.242	.238	.255				
SN	.528	-.528	-.778	.718	.646	.705	.705	.671	.732	.460	.413	.235	.299	.363	.483	.350				
SST	.478	.503	-.807	-.703	.653	.731	.697	.690	.690	.500	.357	.286	.322	.352	.428	.369				
LLA	.449	.506	.625	.712	-.848	.906	.888	.824	.824	.593	.344	.204	.264	.397	.513	.509				
SLA	.338	.472	.612	.650	.795	-.844	.855	.759	.759	.603	.340	.295	.287	.420	.489	.480				
RLA	.462	.474	.613	.660	.804	.839	-.877	.765	.765	.619	.300	.128*	.208	.380	.558	.464				
WLA	.437	.468	.663	.729	.843	.806	.899	-.754	.754	.550	.327	.272	.317	.412	.465	.456				
GA	.693	.689	.719	.721	.768	.638	.748	.746	-.523	.279	.313	.359	.400	.538	.467					
Non-Core Academic SCs																				
REL	.299	.273	.466	.585	.674	.594	.575	.581	.479	-.408	.318	.251	.329	.426	.362					
ART	.342	.212	.379	.458	.435	.319	.378	.402	.385	.430	-.296	.252	.237	.242	.195*					
GYM	.299	.319	.235	.250	.174	.188	.192*	.133*	.291	.206	-.304	-.880	.464	.281	.417					
Non-Academic SCs																				
PhAB	.212	.262	.220	.197	.167*	.192	.108*	.059*	.227	.163	.230	.784	-.662	.328	.653					
PhAP	.297	.287	.294	.287	.319	.385	.322	.308	.325	.250	.302	.358	.541	-.297	.727					
PRT	.278	.279	.245	.254	.398	.210	.206	.267	.291	.255	.268	.073*	.103*	.206	-.460					
PEER	.360	.360	.434	.467	.492	.441	.397	.440	.378	.368	.355	.375	.449	.577	.230					

Note. ARM = Arithmetic; MT = Measurement; SN = Science; SST = Social Studies; L/S/R/WLA = Listening/Speaking/Reading/Writing in Language Arts, respectively; GA = General-Academic; REL = Religion; PhAb = Physical Ability; PhAp = Physical Appearance; PRT = Parent. Correlations above and below the diagonal are for the sample of boys, and of girls, respectively.

* denotes a nonsignificant parameters (at p < .05).

APPENDIX E: COPYRIGHT PERMISSION LETTERS

I agree for Lisa Larocque to use the
SDA instruments for purposes of her
research and to include a copy of the
instrument in her thesis.

Robert W. Mundy

January 12, 1999

Dana Cruikshank
Permissions Department
American Educational Research Association
1230 17th Street NW
Washington, DC

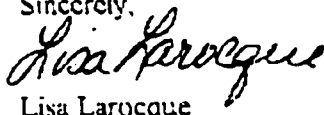
Dear Mr. Cruikshank,

As we discussed earlier today, I am a Ph.D. student in clinical psychology at the University of Ottawa in Ottawa, Ontario, Canada. I am requesting permission to include a copy of a figure in my doctoral dissertation for the purpose of submission to the School of Graduate Studies and Research of the University of Ottawa. In particular, I am requesting permission to include a copy of the model of self-concept represented pictorially in Figure 1 of the article "Self-concept: Validation of construct interpretations" by R. J. Shavelson, J. J. Hubner, and G. C. Stanton, 1976, *Review of Educational Research*, 46, p. 413. I understand that my copy of this model must include this bibliographic citation, and a credit line acknowledging the copyright year, and the American Educational Research Association as the copyright holder.

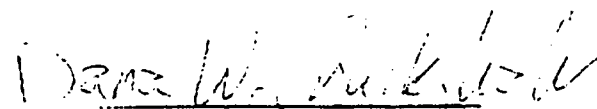
It is my understanding from our conversation that permission to reprint a figure is not needed under the circumstances I described above. Nevertheless, you indicated your willingness to provide formal permission to reprint the aforementioned figure in the doctoral dissertation I will be submitting to the School of Graduate Studies and Research of the University of Ottawa. For your convenience, I have provided for you to indicate formal permission by signing below.

I thank you for your time and assistance in this matter.

Sincerely,



Lisa Larocque

 1-12-98

Dana Cruikshank

January 12, 1999

Rights and Permissions Department
Lawrence Erlbaum Associates, Incorporated
10 Industrial Avenue
Mahwah, New Jersey
07430-2262

To Whom It May Concern,

I am a Ph.D. student in clinical psychology at the University of Ottawa in Ottawa, Ontario, Canada. As part of the requirements for the Ph.D. program, I am completing a thesis entitled "Proadolescent self-concept and self-concept/academic achievement relations: Investigating multidimensional and hierarchical structures within and across gender".

As part of the thesis I submit to the School of Graduate Studies and Research of the University of Ottawa, I would like to include an adaptation of a figure for which Lawrence Erlbaum Associates is copyright holder. Accordingly, I am requesting permission to include an adaptation of a figure in my doctoral dissertation for the purpose of submission to the School of Graduate Studies and Research of the University of Ottawa. In particular, I am requesting permission to include an adaptation of one of the models (Model 7) of self-concept represented pictorially in Figure 3 of the article "Self-concept: Its multifaceted hierarchical structure" by H. W. Marsh, and R. J. Shavelson, 1985, *Educational Psychologist*, 20(3), p. 114. I understand that my copy of this model must include this bibliographic citation, and a credit line acknowledging the copyright year, Lawrence Erlbaum Associates, Inc. as the copyright holder, and that the figure was adapted by permission of the publisher. Please find enclosed a copy of the adaptation of the figure for which I am requesting permission.

I thank you in advance for your time and attention in this matter. Please feel free to contact me should any additional information be required.

Sincerely,

Lisa Larocque
Lisa Larocque

PERMISSION GRANTED provided that material has appeared in our work without credit to another source; you obtain the consent of the author(s); you credit the original publication; and reproduction is confined to the purpose for which permission is hereby given.

E. D'Attilio

1-20-99

Date

Office of Rights/Permissions
Lawrence Erlbaum Associates, Inc.

I, Herb Marsh

provide consent

do not provide consent

for Lisa Larocque to include, in her doctoral dissertation submitted to the School of Graduate Studies and Research of the University of Ottawa, an adaptation of one of the models (Model 7) of self-concept represented pictorially in Figure 3 of the article "Self-concept: Its multifaceted hierarchical structure" by H. W. Marsh and R. J. Shavelson, 1985, *Educational Psychologist*, 20(3), p. 114.

Herb Marsh
(Signature)

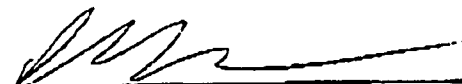
3 Feb 1999
(Date)

I, Rubens Shavelson

provide consent

do not provide consent

for Lisa Larocque to include, in her doctoral dissertation submitted to the School of Graduate Studies and Research of the University of Ottawa, an adaptation of one of the models (Model 7) of self-concept represented pictorially in Figure 3 of the article "Self-concept: Its multifaceted hierarchical structure" by H. W. Marsh and R. J. Shavelson, 1985, *Educational Psychologist*, 20(3), p. 114.


(Signature)

2/2/99
(Date)