DOCUMENT RESUME

ED 269 429 TM 860 258

AUTHOR Marsh, Herbert W.; Smith, Ian D.
TITLE Cross-National Study of the Structure and Level of Multidimensional Self-Concepts: An Application of Confirmatory Factor Analysis.
PUB DATE 10 Mar 86
NOTE 22p.
PUB TYP Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Cross Cultural Studies; *Factor Analysis; *Factor Structure; Foreign Countries; Foreign Students; Intermediate Grades; Preadolescents; Questionnaires; *Self Concept Measures; *Sex Differences
IDENTIFIERS Australia; *Confirmatory Factor Analyses; England; LISREL Computer Program; *Self Description Questionnaire

ABSTRACT

The purposes of the present investigation are to compare the structure and level of multidimensional self-concepts for boys and girls from England and Australia, and to demonstrate the use of confirmatory factor analysis (CFA) in making such comparisons. Subjects were random samples of 303 English preadolescents and 303 Australians that were matched in terms of age and sex. In the first set of analyses, CFA demonstrated that an a priori factor structure of self-concept fitted responses by English and by Australian students, and that the factor structure was reasonably invariant across the two groups. In the second set of analyses CFA demonstrated that English and Australian students differed significantly on only one of eight self-concept factors, and that all sex differences in the multidimensional self-concepts were similar for English and Australian students. The results of the study provide further evidence of the generality of multidimensional self-concepts as inferred by the Self Description Questionnaire, and demonstrate recent advances in the application of CFA. (Author)

Cross-national Study of the Structure and Level of Multidimensional Self-concepts: An Application of Confirmatory Factor Analysis

Herbert W. Marsh & Ian D. Smith

University of Sydney

10 March, 1986

Running Head: A Cross-National Study

A Cross-national Study of the Structure and Level of Multidimensional Self-concepts: An Application of Confirmatory Factor Analysis

ABSTRACT

The purposes of the present investigation are to compare the structure and level of multidimensional self-concepts for boys and girls from England and Australia, and to demonstrate the use of confirmatory factor analysis (CFA) in making such comparisons. Subjects were random samples of 303 English preadolescents and 303 Australians that were matched in terms of age and sex. In the first set of analyses, CFA demonstrated that an a priori factor structure of self-concept fitted responses by English and by Australian students, and that the factor structure was reasonably invariant across the two groups. In the second set of analyses CFA demonstrated that English and Australian students differed significantly on only one of eight self-concept factors, and that all sex differences in the multidimensional self-concepts were similar for English and Australian students. The results of the study provide further evidence of the generality of multidimensional self-concepts as inferred by the Self Description Questionnaire, and demonstrate recent advances in the application of CFA.

A Cross-national Study of the Structure and Level of Multidimensional Self-concepts: An Application of Confirmatory Factor Analysis

The purposes of the present investigation are both substantive and methodological. The substantive purpose is the comparison of multidimensional self-concept responses by boys and girls from England and Australia. The methodological purpose is to demonstrate the application of recent developments in the use of confirmatory factor analysis (CFA) for comparing responses by different groups.

The Structure and Level of Self-concept.

Self-concept research has suffered from a paucity of theoretical models and psychometrically-sound measurement instruments. Shavelson, Hubner, and Stanton (1976) reviewed theoretical and empirical research, and developed a multifaceted model of self-concept that served as the basis of the Self Description Questionnaire (SDQ) used in the present investigation. Through the mid-1970s self-concept instruments typically consisted of a hodge-podge of self-referent items, and "blind" applications of exploratory factor analysis (EFA) typically failed to identify salient, replicable factors (see Marsh & Smith, 1982; Shavelson, et al., 1976). More recently, researchers have developed instruments to measure specific self facets that are at least loosely based on explicit theoretical models such as that proposed by Shavelson, and then used factor analyses to support these a priori factors (Boersma & Chapman, 1979; Dusek & Flaherty, 1981; Fleming & Courtney, 1984; Harter, 1982; Marsh, 1986a; Marsh, Barnes, Cairns & Tidman, 1984; Marsh & Hocevar, 1985; Marsh, Smith & Barnes, 1985; Soares & Soares, 1977). Recent reviews of such research (Byrne, 1984; Marsh & Shavelson, 1985; Shavelson & Marsh, in press) support the multidimensional structure of self-concept, and emphasize research based on the SDQ.

Sex and Age Effects in SDQ Responses.

Sex and age differences are two of the most frequently examined influences on self-concept responses (e.g., Wylie, 1979), and have been the focus of much SDQ research (e.g. Marsh, 1985a; Marsh, Barnes, et al., 1984; Marsh, Smith & Barnes, 1985; also see Marsh, 1986a). SDQ research with Australian preadolescents has shown that: a) responses to most self-concept facets decline with age; b) sex differences in specific facets of self-concept appear to be consistent with sex stereotypes (e.g., girls tend to have higher academic self-concepts -- particularly Reading self-concept -- and boys tend to have higher self-concept in physical areas); and c) sex-by-

age interactions are small and typically nonsignificant.

Despite the large number of studies of sex and age effects in mean responses to self-concept instruments, few researchers have examined the age and sex differences in the factor structure of such responses. Unless the factor structures are similar, there may be no basis for comparing mean differences. A few researchers have used conventional exploratory factor analysis (EFA) to illustrate the similarity of factor structures by different groups (e.g., Marsh, Barnes, et al., 1984; Dusek & Flaherty, 1981). However, Marsh (1985b; Marsh & Hocevar, 1985) argued that EFA is not entirely appropriate for the comparison of factor structures, and advocated the use of confirmatory factor analysis (CFA). Similarly, Alwin and Jackson (1981) note that the investigation of factorial invariance with EFA confounds separate and distinct issues such as the invariance of specific aspects of the factor solution, and argue that "the use of exploratory factor analysis in its conventional form to examine issues of factorial invariance is of limited utility" (p. 253). They also advocated the use of CFA instead of EFA. Using CFA, Marsh and Hocevar (1985) demonstrated that the factor structure of responses to the SDQ was relatively invariant across responses by four preadolescent age groups, and Marsh (1986a; Marsh, Smith & Barnes, 1985) demonstrated the factorial invariance of responses by preadolescent males and females. These demonstrations of the factorial invariance of SDQ responses are substantively important, and also provide a justification for the interpretation of sex and age differences in mean responses to the SDQ.

The Relevance of Cross-national Studies of Self-concept.

Few cross-national or cross-cultural investigations of multidimensional self-concepts have been conducted, but there are a variety of reasons why the structure and level of self-concept inferred from the same instrument may vary in such comparisons. First, the specific meaning of the words may differ. Second, even if the meaning of the words is the same, children's willingness to describe themselves with either favorable or unfavorable terms may differ from country to country. Third, the relations between different facets of self-concept may differ. Fourth, differences in the levels of self-concepts for salient groups, for example the sex differences that were summarized earlier, may vary from country to country due to the socialization processes experienced by the groups in each country.

Until recently, SDQ research has been conducted almost exclusively with Australian subjects from urban areas of New South Wales, and this may limit

the generality of the findings. Thus, cross-national comparisons also have important practical implications for the validity of the use of the SDQ in different countries and, perhaps, even in different parts of Australia. If the SDQ factor structure is not reasonably well-defined for responses by students from a different country, then its use may not be justified in that country. If the SDQ factor structure derived from responses by preadolescents from a different country is not reasonably invariant with the structure found for Australian preadolescents, then the relations between SDQ factors and other constructs found in Australian studies may not generalize to research in that country. Even when the factor structures are reasonably invariant, if mean responses to the SDQ scales by students from different countries differ from those by Australian students, then the two sets of responses may not be directly comparable and the norm tables for the SDQ may not be appropriate for use elsewhere. To the extent that the structure and level of self-concept inferred from responses to the SDQ are invariant across responses by students from different countries, then the broader use of the SDQ is justified. Hence, the detailed study of cross-national comparisons has important theoretical and practical implications.

Smith and Marsh (1985) collected responses to the SDQ by English preadolescents. An EFA of responses by the English students identified factors similar to those identified in Australian research, and mean differences between responses by English and Australian students were small and generally nonsignificant. The purpose of the present investigation is to describe a more detailed analysis of this data in which recent advances in the application of CFA are applied to test the invariance of the structure and level of self-concept, and to compare sex differences in the two samples.

Method

The Samples.

Smith and Marsh (1985) collected responses to the SDQ by 303 English preadolescents based on a random sample of nine primary schools drawn from the urban areas of Lancashire in North West England (see Smith & Marsh, 1985, for more details of the sample). The English sample consisted of 171 males and 132 females from 11 Year-4 classes (These Year 4 students are typically in their sixth year of formal schooling.) These students were in their final year of primary school and had a modal age of 10 years. Based on age, this school grade corresponds most closely to fifth grade in Australia. Hence, a random sample of fifth-grade students, 171 males and 132 females, was selected from the normative archive of responses by Australian

preadolescents (Marsh, 1986a) to compare with the English sample.

The Self Description Questionnaire.

The SDQ assesses three areas of academic self-concept and four areas of nonacademic self-concept derived from the Shavelson model of self-concept (Shavelson, et al. 1976) and a General-self scale derived from the Rosenberg (1965) self-esteem scale. On the SDQ, preadolescent children are asked to respond to simple declarative sentences (e.g., I'm good at mathematics, I make friends easily) using one of five response categories: false; mostly false; sometimes false, sometimes true; mostly true; true. Each of the eight SDQ scales is inferred on the basis of eight positively worded items. For both the English and Australian samples, the SDQ was administered to intact classes of students during regular school hours according to standardized administration procedures that are presented in the test manual (Marsh, 1986a). A brief description of the eight SDQ scales is as follows:

1) Physical Abilities/Sports (Phys) -- student ratings of their ability and enjoyment of physical activities, sports and games;

2) Physical Appearance (Appr) -- student ratings of their own attractiveness, how their appearance compares with others, and how others think they look;

3) Peer Relations (Peer) -- student ratings of how easily they make friends, their popularity, and whether others want them as a friend;

4) Parent Relations (Prnt) -- student ratings of how well they get along with their parents and whether they like their parents;

5) Reading (Read) -- student ratings of their ability in and their enjoyment/interest in reading;

6) Mathematics (Math) -- student ratings of their ability in and their enjoyment/interest in mathematics;

7) General-school (Schl) -- student ratings of their ability in and their enjoyment/interest in "all school subjects;"

8) General-self (Genr) -- student ratings of themselves as effective, capable individuals, who are proud and satisfied with the way they are.

Descriptions of the instrument, the theoretical definition of self-concept upon which it is based, the eight scales, internal consistency estimates of reliability, numerous EFAs and CFAs of responses to the SDQ, and construct validity studies are summarized in the test manual (Marsh, 1986a; also see Marsh, 1985a; 1986b; Marsh, Barnes, et al., 1984; Marsh & Hocevar, 1985; Marsh & Parker, 1984; Marsh, Parker & Smith, 1983; Marsh, Relich & Smith, 1984; Marsh & Richards, 1986; Marsh & Shavelson, 1985;

Marsh, Smith & Barnes, 1983; 1984; 1985; Marsh, Smith, Barnes & Butler, 1983). This research has shown the SDQ scales to be well defined, reliable (coefficient alphas in the .80s and .90s), moderately correlated with measures of corresponding academic abilities (.3 to .7), in agreement with self-concepts inferred by others, affected by experimental manipulations designed to enhance self-concepts, and logically related to other constructs.

Statistical Analyses.

Two sets of analyses were conducted. First, the commercially available LISREL V program (Joreskog & Sorbom, 1981) was used to test the factorial invariance of responses to the SDQ by English and Australian students. Second, LISREL was used to test for multivariate group differences, sex differences, and sex-by-group interactions across the set of eight SDQ factors.

Tests of factorial invariance. In CFA the researcher defines the specific factor structure to be examined, and is able to test its ability to fit the data (see Bagozzi, 1980; Joreskog, 1983; Joreskog & Sorbom, 1981; Long, 1983; Pedhauzur, 1982; Wolfle, 1981 for overviews of the LISREL approach to CFA). Recent methodological advances in the application of CFA (Alwin & Jackson, 1981; Joreskog & Sorbom, 1981; Marsh, 1985; Marsh & Hocevar, 1985) provide a more rigorous comparison of the factor structures resulting from multiple groups for which the researcher is able to test the fit of a model in which any specified group of parameter estimates are constrained to be invariant across groups. This allows the researcher to specify the factor structure to be examined, to uniquely identify parameters in the solutions, and to test hypotheses of invariance for particular components of the factor solution. Here, the researcher is not only examining the similarity of the pattern of parameters from two different groups, but is testing the hypothesis that the parameter values are the same across groups.

A description of the technical details of the application of CFA and tests of factorial invariance is beyond the scope of this study, but the general procedures are similar to those described by Marsh and Hocevar (1985) in their analysis of the factorial invariance of responses to the SDQ by children from different age groups. The general approach is to first establish the ability of the hypothesized factor structure to fit responses from each group without requiring any of the parameter estimates -- factor loadings, factor variances, etc. -- to be the same across groups. Then, a series of increasingly restrictive factor models are tested in which

different sets of parameter estimates are required to be the same across groups. To the extent that a more restrictive model -- one that requires more parameter estimates to be invariant -- is able to fit the data as well as a less restrictive model, then there is support for the invariance of those parameter estimates.

Analyses of factorial invariance were performed on covariance matrices derived from responses to the 32-item pairs that are designed to measure the eight SDQ factors (see Marsh, 1986a; Marsh & Hocevar, 1985; and Marsh, Smith & Barnes, 1985 for a description of the use of item-pairs). A restrictive, "simple structure" was specified in which each item-pair was allowed to load on only the factor it was designed to measure and all other factor loadings were required to be zero (see Table 1 presented in the Results section). One factor loading for each factor was set at 1.0 to serve as a reference indicator (see Long, 1983) so that the model was identified, while the remaining three factor loadings for each factor, factor variances, factor covariances, and error/uniquenesses were estimated.

Tests of mean differences. The analysis of mean differences was based on analyses of a 35 x 35 covariance matrix derived from responses by all 606 subjects from England and Australia. In addition to the 32 variables used to define the SDQ factors, three dichotomous variables were used to define the main effect of country, the main effect of sex, and the sex-by-country interaction. Since the number of male and female students from each country was the same, the sex-by-country interaction term was statistically independent of the main effects of sex and country (see Cohen & Cohen, 1975). In order to test the multivariate effects of the sex and country on the eight SDQ factors, a series a priori models was formulated in which the various sets of parameters representing these effects were estimated or constrained to be zero.

Results

The Factorial Invariance of Responses by English and Australian Students.

In Model 1, the a priori factor structure was fitted to responses by English and Australian students without requiring any of the parameters to be the same across the two groups. The assessment of goodness-of-fit in CFA is not well established, but the general procedure is to examine the parameter estimates in relation to the substantive model, to evaluate the statistical significance of differences between the observed and predicted covariance matrices, and to evaluate subjective indicators of goodness of fit. For both samples the factor structure is well-defined (see Table 1) in

that every factor loading is statistically significant and each factor accounts for a statistically significant portion of the variance (all t-values > 5, $p < .001$). While the X^2s for each of the samples is statistically significant, the goodness-of-fit indicators demonstrate that the fit is reasonable (Table 2). These results suggest that the same pattern of parameters is able to describe responses by students from both countries, but does not necessarily imply that the actual values for these parameters are the same or even similar.

Insert Tables 1 & 2 About Here

In the Models 2 - 5, the invariance of factor loadings, factor correlations, factor variances, error/uniquenesses, and various combinations of these parameters are tested (see Table 2). The minimal condition of factorial invariance is the invariance of factor loadings (Model 2), and Models 3 - 5 each constrains the factor loadings to be invariant in combination with other parameters. Constraining the factor loadings (Model 2), and the factor variances and covariances (Model 3) to be the same for the English and Australian samples does not result in a X^2 that is significantly different from the unconstrained Model 1. Model 5, which constrains all 92 estimated parameters to be the same, differs from Model 1 in a strict statistical sense. However, the goodness-of-fit indices for Models 1 and 5 are nearly the same, and indicate that the factor structure is reasonably invariant across responses by English and Australian students.

In summary, these results provide strong support for the invariance of the factor structure across the responses by English and Australian students, and also provide a justification for comparing mean differences in responses by the two groups.

Tests of Country and Sex Differences In Mean Responses.

For purposes of this analysis, three dichotomous variables representing sex (1=male, 2=female), group (1=Australia, 2=England), and the group-by-sex interaction were added to the 32 item-pairs from the SDQ. Because the design is balanced (i.e., the number of males and females is equal in each group), it was possible to construct uncorrelated variables to represent each of these effects (see Cohen & Cohen, 1975). A 35 x 35 covariance matrix was derived from responses by all 606 students from both countries. All analyses were conducted with an 11-factor model consisting of the 8 SDQ factors, and three single-item factors representing the independent variables to be tested (see Table 3). The set of 8 x 3 = 24 covariances between the 8 SDQ factors and the 3 independent variables indicate the size and statistical

Summary and Implications.

The Substantive Issue.

The substantive purpose of this study was to compare responses to the SDQ by groups of Australian and English students. In the first set of analyses, the factor structure for the two groups was found to be reasonably invariant. In the second set of analyses, Australian and English students were shown to have similar self-concepts for 7 of the 8 SDQ factors, and sex differences in SDQ factors were found to be similar for English and Australian preadolescents. Taken together these results provide strong support for the generalizability of the SDQ responses across responses by English and Australian students.

Several limitations in these results need to be discussed. First, while the two samples may be representative of the geographical areas from which they were chosen (i.e., urban areas of north-west England and of New South Wales), they may not be representative of children in their respective countries as a whole. Second, while the two samples represent different countries, there are many cultural similarities between Australia and England. Further research is needed to determine if these results generalize to other English-speaking Western countries, and to other countries where language and culture are more substantially different from those considered here. Third, the similarity in self-concept responses by English and Australian preadolescents does not necessarily mean that their self-concepts are the same. It is possible that different processes are used to formulate self-concepts for the two groups of students, but that these result in similar factor structures and group average scores. Here, as is all fields of research, interpretation of support for the null hypothesis must be made cautiously.

Smith and Marsh (1985) also discussed the suitability of the SDQ for students in England. First, the SDQ was submitted for comment and inspection to counsellors and researchers at the University of Lancaster and to the Lancashire Local Education Authority. These professionals expressed only minor reservations about a few isolated words such as "kid" and "dumb." Second, the children themselves apparently had no trouble understanding these or any of the other words on the SDQ. There were no questions asked when the first author of that study read the individual items out loud to students in each class, nor were difficulties raised in subsequent class and individual discussions of the instrument. These anecdotal results provide further corroboration for the suitability of the SDQ to English children.

The Methodological Issue.

The methodological purpose of the study was to demonstrate the application of CFA to study the invariance of structure and level for responses to the SDQ across different groups. The demonstration of the invariance of the factor structure across the English and Australian students is substantively important, but also provides a justification for the comparison of mean differences in the level of self-concept by the two groups. If the factor structures for the two groups to be compared are not reasonably invariant, then the interpretation of mean differences between the two groups may be dubious.

The CFA-approach to testing for main effects due to group and sex, and the group-by-sex interaction is analogous in many ways to conventional MANOVA procedures. Like MANOVA, the CFA-approach provides a multivariate test of the effect of each independent variable across the entire set of self-concept factors. The unconstrained model (Model 6) provides a test of the statistical significance for the effect of each independent variable on each self-concept factor, and by constraining the covariances between any combination of independent and dependent variables to be zero a wide array of alternative models can be tested. Such comparisons can also be made with MANOVA by defining special contrasts, but MANOVA procedures typically do not provide such flexibility in defining alternative models.

The CFA-approach also has important advantages that are not available with conventional MANOVA procedures. In the present application these advantages have to do with the measurement model used to define the self-concept factors. First, the a priori structure of the self-concept indicators can be established by the researcher, and the CFA procedure provides the optimal weights for the measured variables used to infer each self-concept factor. In this respect, the procedure combines the advantages of factor analysis and MANOVA into a single statistical procedure. Second, so long as each self-concept factor is inferred from multiple indicators, the CFA-approach automatically corrects each self-concept factor for measurement error before relating it to the independent variables.

In summary the CFA-approach demonstrated here combines many of the advantages of conventional factor analysis and MANOVA procedures, but also provides important advantages that are not typically available with either.

REFERENCES

Alwin, D. F., & Jackson, D. J. (1981). Applications of simultaneous factor analysis to issues of factorial invariance. In D. J. Jackson & E. F. Borgotta (Eds.), Factor analysis and measurement in sociological research: A multidimensional perspective. Beverly Hills, Calif.: Sage.

Bagozzi, R. P. (1980). Causal models in marketing. New York: Wiley.

Bentler, P. M. & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. Psychological Bulletin, 88, 588-606.

Boersma, F.J. & Chapman, J.W. (1979). Student's Perception of Ability Scale Manual. University of Alberta. Edmonton, Canada.

Byrne, B. M. (1984). The general/academic self-concept nomological network: A review of construct validation research. Review of Educational Research, 54, 427-456.

Cohen, J. & Cohen, P. (1975). Applied multiple regression/correlation analysis for the Behavioral Sciences. Hillsdale, New Jersey: Lawrence Erlbaum.

Dusek, J. B. & Flaherty, J. F. (1981). The development of self-concept during adolescent years. Monographs of the Society for Research in Child Development, 46, (4,Serial No. 191).

Fleming, J. S. & Courtney, B. E. (1984). The dimensionality of self-esteem: II: Hierarchical facet model for revised measurement scales. Journal of Personality and Social Psychology, 46, 404-421.

Harter, S. (1982). The Perceived Competence Scale for Children, Child Development, 53, 87-97.

Joreskog, K. G. (1981). Analysis of covariance structures. Scandinavian Journal of Statistics, 8, 65-92.

Joreskog, K. G. & Sorbom, D. (1981). LISREL V: Analysis of Linear Structural Relations By the Method of Maximum Likelihood. Chicago: International Educational Services.

Long, K. S. (1983). Confirmatory Factor Analysis. Beverly Hills, CA: Sage.

Marsh, H. W. (1985a). Age and sex effects in multiple dimensions of preadolescent self-concept. Australian Journal of Psychology, 197-204.

Marsh, H. W. (1985b). The structure of masculinity/femininity: An application of confirmatory factor analysis to higher-order factor structures and factorial invariance. Multivariate Behavioral Research, 20, 427-449.

Marsh, H. W. (1986a). The Self Description Questionnaire (SDQ): A

significance of each effect.

Insert Tables 3 & 4 About Here

A series of a priori models was tested to determine the multivariate effects of sex, group, and their interaction across the 8 SDQ factors (see Table 4). In Model 6, all the effects were freely estimated, and in subsequent models various combinations of effects were fixed to be zero. To the extent that a model in which effects are fixed to be zero does not differ significantly from Model 6, then these effects do not have a multivariate effect on the self-concept factors that is statistically significant. In Model 7, all eight interaction effects were fixed to be zero, and the chi-square for Model 7 did not differ significantly from that of Model 6. In Models 8 and 9, the effects of group or sex were also fixed to be zero along with the set of 8 interaction terms. While the test of no group differences (Model 10) was barely rejected, the test of no sex differences (Model 9) was strongly rejected. Due primarily to the substantial sex effects, setting all 24 effects to be zero (Model 10) was also rejected.

One a posteriori model (Model 11), based on the results of Models 8 - 11, was specified in which 7 of 8 group effects, 2 of 8 sex effects, and all 8 interaction effects were fixed to be zero, and the X^2 for this model was not significantly different from that of Model 6. The parameter estimates for Model 12 (Table 3) indicate that: a) boys had higher self-concepts in Physical Ability, Physical Appearance, Peer Relations and General-self, lower self-concepts in Reading and General-School, and did not differ from girls in Parent Relationships and Math; b) Australians had higher self-concepts in General-School but did not differ from the English students in any other areas of self-concept; c) there were no statistically significant group-by-sex interactions. However, it should be noted that the sizes of all these correlations (standardized factor covariances presented in the footnote of Table 3) are small even though these coefficients have been corrected for unreliability in SDQ responses. In particular, the correlation representing the one statistically significant group effect is only -.10.

In summary the results of this analysis indicate that the self-concepts for samples of Australian and English students are similar for 7 of 8 SDQ factors, and that the difference on the eighth factor is small. While sex effects were found in a majority of the SDQ factors, these effects were similar for the English and Australian samples and similar to those reported in other SDQ studies.

theoretical and empirical basis for the measurement of multiple dimensions of preadolescent self-concept: A test manual and a research monograph. Faculty of Education, University of Sydney, NSW, Australia.

Marsh, H. W. (1986b). Verbal and math self-concepts: An internal/external frame of reference model. American Educational Research Journal, xxx-xxx.

Marsh, H. W., Barnes, J., Cairns, L., & Tidman, M. (1984). The Self Description Questionnaire (SDQ): Age effects in the structure and level of self-concept for 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 & multitrait-multimethod analysis. Journal of Personality and Social Psychology, 49, 1360-1377.

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

Marsh, H. W., & 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. W., & 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, 173-187.

Marsh, H. W. & Richards, G. (1986) The Outward Bound Bridging Course for low achieving high-school males: Effect on academic achievement and multidimensional self-concepts. (In Review). Faculty of Education, University of Sydney, Australia.

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

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. & Barnes, J. (1983). Multitrait-multimethod analyses of the Self Description Questionnaire: Student-teacher agreement on multidimensional ratings of student self-concept.

American Educational Research Journal, 20, 333-357.

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

Marsh, H. W., Smith, I. D. & Barnes, J. (1985). Multidimensional self-concepts: Relationships 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, dimensionality, validity, and the measurement of change. Journal of Educational Psychology, 75, 772-790.

Pedhauzur, E. J. (1982). Multiple regression in behavioral research (2nd ed.). New York: Holt, Rinehart and Winston.

Rosenberg, M. (1965). Society and the adolescent child. Princeton: Princeton University Press.

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

Shavelson, R. J., & Marsh, H. W. (In Press). On the structure of self-concept. In R. Schwarzer (Ed.), Anxiety and cognitions. N. J. Lawrence Earlbaum.

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.

Soares, L. M. and Soares, A. T. (1977). The self-concept: Mini, maxi, multi? Paper presented at the 1977 Annual Meeting of the American Educational Research Association, New York.

Wolfle, L. M. (1981). Causal models with unmeasured variables: An introduction to LISREL. Paper presented at the Annual Meeting of the American Educational Research Association, Los Angeles, 1981.

Wylie, R. C. (1979). The self-concept. (Vol. 2) Lincoln: University of Nebraska Press.

TABLE 1

Model 1 Parameter Estimates For Australian (A) and English (E) Samples

		PHYS	APPR	PEER	PRNT	READ	MATH	SCHL	GENL	Error/ Uniqueness
		Factor Loading Matrix (LAMBDA)								
Variables										
Phys1	A	1.00	0	0	0	0	0	0	0	.39*
	E	1.00	0	0	0	0	0	0	0	.37*
Phys2	A	.90*	0	0	0	0	0	0	0	.43*
	E	1.01*	0	0	0	0	0	0	0	.33*
Phys3	A	.99*	0	0	0	0	0	0	0	.31*
	E	1.06*	0	0	0	0	0	0	0	.33*
Phys4	A	.97*	0	0	0	0	0	0	0	.40*
	E	1.04*	0	0	0	0	0	0	0	.37*
Appr1	A	0	1.00	0	0	0	0	0	0	.48*
	E	0	1.00	0	0	0	0	0	0	.50*
Appr2	A	0	.74*	0	0	0	0	0	0	.58*
	E	0	.86*	0	0	0	0	0	0	.57*
Appr3	A	0	1.07*	0	0	0	0	0	0	.38*
	E	0	1.10*	0	0	0	0	0	0	.32*
Appr4	A	0	1.15*	0	0	0	0	0	0	.33*
	E	0	1.23*	0	0	0	0	0	0	.42*
Peer1	A	0	0	1.00	0	0	0	0	0	.50*
	E	0	0	1.00	0	0	0	0	0	.46*
Peer2	A	0	0	1.06*	0	0	0	0	0	.41*
	E	0	0	1.26*	0	0	0	0	0	.30*
Peer3	A	0	0	1.11*	0	0	0	0	0	.40*
	E	0	0	1.35*	0	0	0	0	0	.39*
Peer4	A	0	0	1.20*	0	0	0	0	0	.35*
	E	0	0	1.28*	0	0	0	0	0	.30*
Prnt1	A	0	0	0	1.00	0	0	0	0	.65*
	E	0	0	0	1.00	0	0	0	0	.78*
Prnt2	A	0	0	0	1.03*	0	0	0	0	.67*
	E	0	0	0	.97*	0	0	0	0	.78*
Prnt3	A	0	0	0	1.44*	0	0	0	0	.43*
	E	0	0	0	1.15*	0	0	0	0	.34*
Prnt4	A	0	0	0	1.47*	0	0	0	0	.43*
	E	0	0	0	1.22*	0	0	0	0	.28*
Read1	A	0	0	0	0	1.00	0	0	0	.31*
	E	0	0	0	0	1.00	0	0	0	.28*
Read2	A	0	0	0	0	1.03*	0	0	0	.21*
	E	0	0	0	0	1.15*	0	0	0	.28*
Read3	A	0	0	0	0	.99*	0	0	0	.30*
	E	0	0	0	0	1.08*	0	0	0	.24*
Read4	A	0	0	0	0	.95*	0	0	0	.29*
	E	0	0	0	0	1.17*	0	0	0	.34*
Math1	A	0	0	0	0	0	1.00	0	0	.33*
	E	0	0	0	0	0	1.00	0	0	.27*
Math2	A	0	0	0	0	0	1.06*	0	0	.29*
	E	0	0	0	0	0	1.15*	0	0	.27*

TABLE 1 continued

		PHYS	APPR	PEER	PRNT	READ	MATH	SCHL	GENL	Error/ Uniqueness
Factor Loading Matrix (LAMBDA)										
Math3	A	0	0	0	0	0	1.17*	0	0	.20*
	E	0	0	0	0	0	1.18*	0	0	.16*
Math4	A	0	0	0	0	0	1.17*	0	0	.12*
	E	0	0	0	0	0	1.19*	0	0	.20*
Schl1	A	0	0	0	0	0	0	1.00	0	.45*
	E	0	0	0	0	0	0	1.00	0	.43*
Schl2	A	0	0	0	0	0	0	.90*	0	.49*
	E	0	0	0	0	0	0	.98*	0	.53*
Schl3	A	0	0	0	0	0	0	1.09*	0	.40*
	E	0	0	0	0	0	0	1.09*	0	.36*
Schl4	A	0	0	0	0	0	0	1.26*	0	.25*
	E	0	0	0	0	0	0	1.09*	0	.18*
Genl1	A	0	0	0	0	0	0	0	1.00	.64*
	E	0	0	0	0	0	0	0	1.00	.50*
Genl2	A	0	0	0	0	0	0	0	1.57*	.35*
	E	0	0	0	0	0	0	0	1.57*	.76*
Genl3	A	0	0	0	0	0	0	0	1.13*	.38*
	E	0	0	0	0	0	0	0	1.28*	.37*
Genl4	A	0	0	0	0	0	0	0	1.17*	.43*
	E	0	0	0	0	0	0	0	1.44*	.50*

Correlations Among Factors (PSI)

Factors								
PHYS	A	.66*						
	E	.51*						
APPR	A	.33*	.60*					
	E	.26*	.53*					
PEER	A	.44*	33*	.58*				
	E	.25*	.23*	.30*				
PRNT	A	.16*	.11*	.20*	.33*			
	E	.08*	.12*	.14*	.39*			
READ	A	.13*	.13*	.16*	.10*	.72*		
	E	.00	-.06	.04	.06	.58*		
MATH	A	.18*	.16*	.16*	.05	.15*	.64*	
	E	.07	.15*	.12*	.07	.11*	.62*	
SCHL	A	.21*	.16*	.20*	.13*	.32*	.44*	.53*
	E	.06	.11*	.11*	.12*	.21*	.40*	.47*
GENL	A	.43*	.29*	.44*	.24*	.16*	.19*	.23*	.49*
	E	.24*	.20*	.18*	.13*	.06	.13*	.14*	.24*

* p < .01

Note. Parameters with Values of 0 and 1 were fixed and not estimated as part of the analysis, and so no tests of statistical significance were performed for these values. The four measured variables designed to measure each factor are the sums of responses to pairs of positively worded items.

TABLE 2

Summaries of Goodness of Fit Indices for the CFA Models Describing Tests of Factorial Invariance Across the Australian and English Samples

Model Description	X^2	df	X^2/df	BBI	TLI	X^2_d	dfd
0) Null Model	12507.2	992	12.60	.00	.00	----	----
Australian	6535.6	496	13.10	.00	.00	----	----
English	5971.6	496	12.00	.00	.00	----	----
1) No Invariance	1765.6	872	2.02	.86	.91	----	----
Australian	889.4	436	2.04	.86	.91	----	----
English	876.2	436	2.01	.85	.91	----	----
2) Factor Loadings invariant	1797.3	896	2.01	.86	.91	31.7	24
3) Factor Loadings, factor covariances, factor variances invariant	1846.3	932	1.98	.85	.92	80.7	60
4) Factor Loadings, error/ uniquenesses invariant	1859.8	928	2.00	.85	.91	94.2	56
5) Total Invariance	1905.6	964	1.98	.85	.92	140.0	92

Note. The Null model hypothesizes complete independence of all measured variables and provides a measure of the total covariance in the data which is used in computing the Bentler Bonett index (BBI) and the Tucker Lewis index (TLI; see Bentler & Bonett, 1980; Marsh & Hocevar, 1985; for further description of the BBI and TLI).

[a] For Models 2 - 5, the X^2_d and df_d are the differences between the X^2 and df for the model being tested and Model 1 for which no invariance constraints were imposed.

TABLE 3

Model 12 Parameter Estimates For Tests of Group, Sex, and Sex-by-Group Interaction (Inter) Effects in Responses By Australian and English Samples

	Phys	Appr	Peer	Prnt	Read	Math	Schl	Genl	Sex	Group	Inter	
	Factor Loading Matrix (LAMBDA)											Error/
Variables												Uniqueness
Phys1	1.00	0	0	0	0	0	0	0	0	0	0	.37*
Phys2	.98*	0	0	0	0	0	0	0	0	0	0	.40*
Phys3	1.01*	0	0	0	0	0	0	0	0	0	0	.35*
Phys4	.98*	0	0	0	0	0	0	0	0	0	0	.40*
Appr1	0	1.00	0	0	0	0	0	0	0	0	0	.47*
Appr2	0	.87*	0	0	0	0	0	0	0	0	0	.60*
Appr3	0	1.08*	0	0	0	0	0	0	0	0	0	.38*
Appr4	0	1.15*	0	0	0	0	0	0	0	0	0	.30*
Peer1	0	0	1.00	0	0	0	0	0	0	0	0	.54*
Peer2	0	0	1.16*	0	0	0	0	0	0	0	0	.38*
Peer3	0	0	1.80*	0	0	0	0	0	0	0	0	.36*
Peer4	0	0	1.19*	0	0	0	0	0	0	0	0	.35*
Prnt1	0	0	0	1.00	0	0	0	0	0	0	0	.66*
Prnt2	0	0	0	.98*	0	0	0	0	0	0	0	.67*
Prnt3	0	0	0	1.31*	0	0	0	0	0	0	0	.41*
Prnt4	0	0	0	1.36*	0	0	0	0	0	0	0	.37*
Read1	0	0	0	0	1.00	0	0	0	0	0	0	.34*
Read2	0	0	0	0	1.08*	0	0	0	0	0	0	.22*
Read3	0	0	0	0	1.03*	0	0	0	0	0	0	.30*
Read4	0	0	0	0	1.05*	0	0	0	0	0	0	.28*
Math1	0	0	0	0	0	1.00	0	0	0	0	0	.36*
Math2	0	0	0	0	0	1.08*	0	0	0	0	0	.25*
Math3	0	0	0	0	0	1.14*	0	0	0	0	0	.17*
Math4	0	0	0	0	0	1.15*	0	0	0	0	0	.15*
Schl1	0	0	0	0	0	0	1.00	0	0	0	0	.47*
Schl2	0	0	0	0	0	0	.95*	0	0	0	0	.52*
Schl3	0	0	0	0	0	0	1.08*	0	0	0	0	.38*
Schl4	0	0	0	0	0	0	1.17*	0	0	0	0	.27*
Genl1	0	0	0	0	0	0	0	1.00	0	0	0	.63*
Genl2	0	0	0	0	0	0	0	1.20*	0	0	0	.46*
Genl3	0	0	0	0	0	0	0	1.17*	0	0	0	.49*
Genl4	0	0	0	0	0	0	0	1.21*	0	0	0	.46*
Sex	0	0	0	0	0	0	0	0	1.0	0	0	.00
Group	0	0	0	0	0	0	0	0	0	1.0	0	.00
Inter	0	0	0	0	0	0	0	0	0	0	1.0	.00

Table 3 continued

	Phys	Appr	Peer	Prnt	Read	Math	Schl	Genl	Sex	Group	Inter
Correlations Among Factors (PSI)											
Factors											
PHYS	.64*										
APPR	.30*	.53*									
PEER	.37*	.28*	.46*								
PRNT	.13*	.12*	.18*	.34*							
READ	.06*	.02	.10*	.08*	.66*						
MATH	.13*	.15*	.14*	.06	.13*	.64*					
SCHL	.14*	.13*	.16*	.12*	.28*	.44*	.53*				
GENL	.35*	.26*	.31*	.18*	.11*	.16*	.19*	.38*			
Sex[a]	-.15*	-.21*	-.08*	0	.11*	0	.12*	-.12*	1.0		
Group[a]	0	0	0	0	0	0	-.08*	0	0	1.0	
Inter	0	0	0	0	0	0	0	0	0	0	1.0

* $p < .05$

Note. Parameters with Values of 0 and 1 were fixed ar , not estimated as part of the analysis, and so no tests of statistical significance were performed for these values. The X^2 for this model did not differ significantly from the X^2 of the model in which all sex, group, and interactions effects on self-concept were estimated (see Table 4).

[a] Sex is scored 1=male, 2=female so that positive covariances indicate higher self-concepts for girls. Groups is scored 1=Australian 2=English so that positive covariances indicate higher self-concepts by English students. The standardized covariances, that is correlations, representing the 6 sex effects and 1 group effect are -.19, -.28, -.11, .13, .16, -.19, and -.10 respectively.

TABLE 4

Summaries of Goodness of Fit Indices for the CFA Models Describing Tests of Group, Sex, and Sex-by-Group Interaction (Inter) Effects in Responses By Australian and English Samples

Model Description	X^2	df	X^2/df	BBI	TLI	X^2_d	dfd
0) Null Model	12019.2	595	20.20	.00	.00	----	----
6) No Constraints	1282.4	514	2.49	.89	.92	----	----
7) Inter effects fixed to be zero	1295.8	522	2.48	.89	.92	13.4	8
8) Inter & Sex effects fixed to be zero	1393.6	530	2.63	.88	.92	111.2	16
9) Inter & Groups effects fixed to be zero.	1316.7	530	2.48	.89	.92	34.3	16
10) Inter, Sex & Group effects fit to be zero	1414.6	538	2.62	.88	.92	132.2	24
11) Inter, selected Sex & Group effects fixed to be zero	1306.3	531	2.46	.89	.92	23.9	17

Note. See notes in Tables 2 and 3.