DOCUMENT RESUME

ED 256 773 TM 840 773

AUTHOR	Marsh, Herbert W.
TITLE	The Hierarchical Structure of Self-Concept: An Application of Hierarchical Confirmatory Factor Analysis. Revision.
PUB DATE	25 Mar 85
NOTE	39p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Academic Ability; Correlation; *Factor Analysis; *Factor Structure; Goodness of Fit; Interpersonal Competence; *Models; Moral Values; Physical Characteristics; Psychometrics; *Research Methodology; *Self Concept; Self Concept Measures; Self Esteem; Vertical Organization
IDENTIFIERS	Confirmatory Factor Analysis; *Hierarchical Structural Analysis; *Self Description Questionnaire III; Self Report Measures

ABSTRACT

This investigation: (1) tested the ability of an a priori hierarchical structure of self-concept derived from the Shavelson model to explain responses to the Self Description Questionnaire III (SDQ III); and (2) demonstrated the application and problems with the use of hierarchical confirmatory factor analysis. A first-order factor analysis clearly identified the 13 facets of self-concept that the SDQ III is designed to measure. A series of hierarchical models clearly supported the separation of the 13 SDQ III facets of self-concept into academic and nonacademic components, and the academic facets into math/academic and verbal/academic components. However, support for the physical, social, and moral second-order factors was less clear. Third-order hierarchical models resulted in a clearly defined hierarchical general self-concept that was substantially related to general esteem, and to physical, social and emotional components of self-concept, but not to the academic and moral values components. (Author)

The Hierarchical Structure of Self-Concept: An Application of Hierarchical Confirmatory Factor Analysis

Herbert W. Marsh
Department of Education
The University of Sydney

7 October, 1984
Revised: 25 March, 1985

Running Head: Self-concept

Reprint requests should be sent to Dr. Herbert W. Marsh, Department of Education, University of Sydney, NSW 2006, Australia. The author would like to acknowledge the helpful comments by Dennis Hocevar and Raymond Debus on earlier drafts of this article, and also to thank Gary Richards and the Outward Bound program for assistance in the collection of the data.

The Hierarchical Structure of Self-Concept: An Application of Hierarchical Confirmatory Factor Analysis

ABSTRACT

This investigation (a) tested the ability of an a priori hierarchical structure of self-concept derived from the Shavelson model to explain responses to the Self Description Questionnaire III (SDQ III), and (b) demonstrated the application and problems with the use of hierarchical confirmatory factor analysis (HCFA). A first-order factor analysis clearly identified the 13 facets of self-concept that the SDQ III is designed to measure. A series of hierarchical models clearly supported the separation of the 13 SDQ III facets of self-concept into academic and nonacademic components, and the academic facets into math/academic and verbal/academic components. However, support for the physical, social, and moral second-order factors was less clear. Third-order hierarchical models resulted in a clearly defined hierarchical general self-concept that was substantially related to general Esteem, and to physical, social and emotional components of self-concept, but not to the academic and moral values components.

The Hierarchical Structure of Self-Concept: An Application of Hierarchical Confirmatory Factor Analysis

This study has two main purposes. The first, the substantive issue, is to examine the hierarchical structure of multidimensional self-concepts. The second, a methodological demonstration, is to demonstrate the use of confirmatory factor analysis (CFA) and hierarchical confirmatory factor analysis (HCFA), and to examine potential problems with its use and its interpretation.

Self-Concept: A Multifaceted, Hierarchical Construct

Self-concept is widely posited to be a desirable outcome and to explain overt other constructs in many areas of psychology (Burns, 1979; Wells & Marwell, 1976; Wylie, 1974; 1979), but reviews of self-concept research typical identify a lack of theoretical models for defining and interpreting the construct and the poor quality of measurement instruments used to assess it. In an attempt to remedy this situation, Shavelson, Hubner and Stanton (1976) posited a multifaceted hierarchical model of self-concept. In one representation of the model Shavelson proposed a general self-concept defined by academic and nonacademic self-concepts; the academic self-concept was divided into self-concepts in particular academic areas (e.g., English and mathematics); the nonacademic self-concept was divided into social, physical, emotional self-concepts. The social self-concept was further divided into relations with peers and relations with significant others; the physical self-concept was divided into physical ability and physical appearance. (The figure used to illustrate their representation appeared in Shavelson, et al. 1976, and has been reproduced by: Fleming & Courtney, 1984; Shavelson & Bolus, 1982; and Shavelson & Marsh, in press).

Shavelson's model was heuristic and plausible, but in his original presentation there was only modest support for it. His review indicated some support for the separation of self-concept into social, physical and academic components, but he found no one instrument that identified all three components. While exploratory factor analyses of nearly all self-concept instruments indicated that the construct was not unidimensional, factor analyses of the most commonly used instruments typically failed to identify the scales that the instrument had been designed to measure nor were the derived solutions able to be consistently replicated (see Marsh & Smith, 1982; Shavelson, et al., 1976). Some theorists (e.g., Coopersmith, 1967; Marx & Winne, 1978) argued that facets of self-concept were so heavily dominated by a general factor, that separate components could not be readily differentiated. While such a situation would not technically be inconsistent with the Shavelson model, the value of a hierarchical model would be dubious if the hierarchy was so strong that each component could barely be distinguished from a general self-concept.

Through the mid-1970's self-concept instruments typically consisted of a hodge-podge of self-referent items, and little effort was made to develop/refine these instruments in order to measure specific facets of self-concept. Exploratory factor analyses were applied "blindly" to responses from these instruments in the hope that the statistical technique would identify the salient facets; generally it did not. More recently, influenced by increased sophistication in factor analysis and, perhaps, the Shavelson model, a different approach has been adopted. Researchers frequently developed self-concept instruments specifically to measure particular facets of self-concept that are at least loosely based on an explicit theoretical model, and then used factor analysis to test the existence of these a priori facets. This approach has produced instruments in which multiple facets of self-concept are identified (e.g., Boersma & Chapman, 1979; Dusek & Flaherty, 1981; Fleming & Courtney, 1984; Harter, 1982; Marsh, Barnes, Cairns & Tidman, 1984; Marsh, Barnes & Hocevar, in press; Soares & Soares, 1982). Shavelson and Marsh (in press) reviewed research stimulated by the Shavelson model and found strong support for the multidimensionality of self-concept, concluding that self-concept cannot be adequately understood if its multidimensionality is ignored. Perhaps the strongest support for the multidimensionality of self-concept, and particularly for the Shavelson model, came from research with the Self Description Questionnaire (SDQ), a set of three instruments designed to measure self-concepts of preadolescents (SDQ), early-adolescents (SDQ II), and late-adolescents and young adults (SDQ III -- the instrument used in this study).

Marsh and Hocevar (in press; also see Shavelson & Marsh, in press) examined the hierarchical structure of responses to the SDQ for preadolescents. The seven first-order factors that the SDQ is designed to measure (Reading, Math, School, Physical Ability, Physical Appearance, Peer Relationships and Parent Relationships) were identified and a number of different hierarchical models were examined. Neither a single higher-order factor defined by all seven facets, nor two higher-order factors defined by academic and by nonacademic facets was able to fit the data. The best fit was obtained from a model with three higher-order factors: a nonacademic factor, a reading/academic factor, and a math/academic factor. The results of the hierarchical factor analysis, and other research, led Shavelson and Marsh (in press) to propose a revision of the Shavelson model such that self-concepts in specific academic areas define two higher-order academic self-concepts instead of just one as originally proposed by Shavelson. These findings support the contention that self-concept is a multifaceted, hierarchical construct, though they also suggest that the hierarchy may be more complicated than originally proposed by Shavelson.

The term general self-concept has been used in different ways (see Marsh & Shavelson, in press). For example, the apex of the Shavelson model is called general self-concept, while other researchers have constructed scales that are specifically designed to measure a relatively unidimensional construct that is superordinate to particular facets of self-concept and these have also been called general-self. For purposes of the present article these two uses of the term general self-concept are called hierarchical general self-concept and general self-esteem. Items on general self-esteem scales do not refer to particular facets of self-concept, but rather infer a general sense of self-worth or self-competence that could apply to different facets. This approach is used by Rosenberg (1979), Harter (1982), and the General-Self scales on the SDQ instruments. Factor analytic research has shown that these scales can be identified and distinguished from more specific dimensions of self-concept. While both the hierarchical general self-concept and the general self-esteem are superordinant facets, the theoretical relation between them is unclear and the role of general self-esteem is not specified in the Shavelson model. Marsh and Shavelson (in press) indicated the need for research into the hierarchical structure of self-concept that included a general self-esteem scale as well as multiple facets of self-concept. In response to this need, a General-Self scale (a general self-esteem scale) was added to each of the SDQ instruments. Hence, an important aspect of the present investigation will be to examine how a general self-esteem scale fits into the hierarchy of self-concept facets.

<u>The Present Study: A Hierarchical Structure For SDQ III Responses</u>

The present investigation is based on responses to the SDQ III. The SDQ III, the rationale for its construction, its relation to the Shavelson model and the other SDQ instruments, the wording of the items, its psychometric properties, its relation to academic achievement and to self-concept inferred by significant others, is summarized elsewhere (Marsh, Barnes & Hocevar, in press; Marsh & O'Niell, 1984; Marsh, Richards & Barnes, in press). The 13 SDQ III facets can be divided into academic (Math, Verbal, General Academic & Problem Solving), nonacademic (Physical Ability, Appearance, Same Sex Relations, Opposite Sex Relations, Parent Relations, Religion/spiritual values, Honesty, Emotional), and general (Esteem) components. Following the Shavelson model, and its revision, it is proposed that: a) the 4 academic facets define two second-order academic factors -- math/academic and verbal/academic; b) 7 nonacademic facets define three second-order factors -- physical (Physical Ability and Appearance), social (Same Sex Relations, Opposite Sex Relations, and Parent Relations), and moral (Religion/Spiritual Values and Honesty); c) a third-order, hierarchical general self-concept will be defined by Esteem, Emotional, math/academic, verbal/academic, physical, social, and moral self-concepts. This proposed hierarchy differs from that proposed by Shavelson and its subsequent revision is several ways: a) Esteem and its relation to the hierarchical general self, were not included in earlier models, b) Religion/Spiritual Values and Honesty, and their combination into the second-order moral factor were not specified, c) Problem Solving was not identified, nor was its relation to the second-order academic facets, and d) Peer Relations was not divided into Same Sex Relations and Opposite Sex Relations. It should be noted that the figure presented by Shavelson, et al. (1976), and often described as the Shavelson model, was only specified to be one possible representation of a hierarchy of self-concepts. Hence, the hierarchy described here is certainly consistent with the Shavelson model even if it differs from the one 'articular representation of his model.

Data for the present investigation come from a previous study of the effect of participation in the Outward Bound program on multidimensional self-concepts (Marsh, Richards & Barnes, in press). A total of 27 groups (N=361 subjects, median age =21, 75% male, 96% single, 60% full-time employed, 33% full-time students) participated in program, and the SDQ III was completed approximately one month before the start of the program (time 1), on the first day of the program (time 2), and again at the end of the 26-day program (time 3). For purposes of the present study, separate analyses were performed on responses from time 1 and from time 2, though the primary focus is on the time 1 data. The psychometric properties of responses to the SDQ III were examined in the original study. The median coefficient alpha estimate of reliability for the 13 scales was .90 for each of the three administrations, and the median stability coefficient between responses at time 1 and time 2 was .87. In exploratory factor analyses of each set of responses the factor loadings of variables designed to measure each factor -- target loadings -- were high (median = .72), while nontarget loadings were low (median = .02). Correlations among the oblique factors were modest, ranging from - .07 to .39 (median = .10). Even though there was a significant increase in SDQ III responses following participation in the Outward Bound program (time 3), the SDQ III factor structure was similar at times 1, 2, and 3.

The Application of Hierarchical Confirmatory Factor Analysis (HCFA).

In order to examine higher-order factor structures researchers have typically used an exploratory factor analysis with an oblique rotation to derive first-order factors (or scores to represent the first-order factors), and then have factor analyzed correlations among first-order factors in order to infer second-order factors (e.g., Fleming & Courtney, 1984; Marsh, Smith, Barnes & Butler, 1983). Such exploratory approaches are heuristic, and they may suggest possible higher-order factor structures. Nevertheless, they are generally unacceptable because: a) researchers have relatively little control over the definition of first-order factors so that the first-order factors used to infer the second-order factors may be ambiguous; b) researchers have little control over the definition of second-order factors so that neither the hypothesized structure nor viable alternatives can be explicitly defined; c) neither first- nor second-order factors are uniquely defined in exploratory factor analysis, so that alternative, mathematically equivalent structures may lead to different interpretations; and d) exploratory factor analysis does not enable to researcher to compare the goodness-of-fit of the hypothesized structure with viable alternatives, or even to compare the hypothesize structure with the empirically obtained structure. In contrast, HCFA allows the researcher to: specifically define and test the structure of first-order factors; formulate and test alternative higher-order models; uniquely estimate parameters to fit each model; test the ability of each model to fit the data; and, compare the goodness-of-fits of the alternative models. (For a more detailed comparison of exploratory and confirmatory approaches to factor analysis see: Bagozzi, 1980; Huba & Bentler, 1982; Marsh & Hocevar, 1983; 1984a; 1984b; in press; Joreskog, 1971; 1980; Joreskog & Sorbom, 1981; Long, 1983a; 1983b; Olson, 1982; Pedhauzur, 1982; Tanaka & Huba, 1984).

In the analyses presented below, a series of CFA and HCFA factor models are described, and their ability to fit the data are tested with the commercially available LISREL program (Joreskog & Sorbom, 1981). Initially a first-order model was proposed to test the a priori structure of 13 SDQ III factors and to examine the correlations among these factors. Then a series of higher-order factor structures were formulated to explain relations among first-order factors, and these models were tested.

The First-Order Factor Model.

The existence of a well-defined first-order factor structure is a prerequisite to testing higher-order structures. This first-order structure should be an a priori structure based on the design of the instrument, preferably one for which each measured variable is allowed to define one and only one first-order factor (see footnote 1). The ability of this model to fit the data is very important, because all subsequent higher-order models are based on the first-order factors as specified in this model. Also, the goodness-of-fit for this model represents the upper-limit for the goodness-of-fit of any higher-order model based on the same first-order factors (see discussion of Target Coefficient below). Hence, the rationale for the first-order factor structure, its ability to fit the data, and the parameter estimates based on this model should be examined carefully in HCFA studies.

Definition of the First-Order Model. In CFA and HCFA performed by LISREL, alternative models are specified by fixing or constraining elements in design matrices (see footnote 2). Three matrices were used to define the first-order factor model in this study that are conceptually similar to matrices resulting from an exploratory factor analysis: the matrix of factor loadings, a factor variance-covariance matrix that represents relations among first-order factors, and a diagonal matrix of error/uniquenesses that are like one minus the communality estimates in exploratory factor analyses. The simple structure model used in this study can be better understood through an examination of the parameter estimates for the first-order model (Model 1A, see Table 1). Three subscales, each the sum of responses to 3 or 4 items, were used to define each of the 13 SDQ III factors -- a total of 39 measured variables. Each measured variable was allowed to define only the factor it was designed to measure, and its loading on all other factors was fixed to be zero. One measured variable for each of the 13 first-order factors was arbitrarily selected to be a reference variable, and its factor loading was fixed to be 1.0. The factor loadings for the other 26 measured variables were estimated as part of the analysis, and their values appear in Table 1. The variance of each of the first-order factors appears in the diagonal of the variance-covariance matrix (see Table 1), and the covariances among the 13 first-order factors appear in the off-diagonal values of this matrix. Finally, the 39 error/uniquenesses -- one for each measured variables -- are presented as a single column in Table 1. In all there were 156 parameters estimated to fit this model; 26 first-order factor loadings, 13 factor variances, 78 factor covariances, and 39 error/uniquenesses.

Insert Table 1 About Here

Goodness-of-Fit. A plethora of indices of a model's goodness-of-fit are used in CFA, but there are no well established guidelines for what minimal conditions constitute an adequate goodness-of-fit. The general approach is to:

1. examine parameters estimated in relation to the substantive, a

priori model (and also for estimates outside the range of permissible values, called Heywood cases, such as negative estimates of factor variances or error/uniquenesses, or factor correlations greater than 1.0);

2. evaluate the overall chi-square value of the model in terms of statistical significance, and to compare this with values obtained from alternative models;

3. evaluate subjective indices of goodness-of-fit that give an indication of the proportion of variance that is explained by the model, and to compare indices from alternative models.

An examination of the factor loadings for Model 1A (Table 1) indicates that every factor loading, and every factor variance, is large and statistically significant. Factor covariances vary from close to zero to moderately positive, but these can be interpreted more easily when this matrix is standardized to correspond to factor correlations (Table 2). The factor correlations vary between -0.02 to 0.71 (median = 0.26) and none approachs 1.0. The same general pattern of results, and nearly the same parameter estimates, were obtained when Model 1A was tested with data collected at time 2 instead of time 1 (the factor correlations appear in Table 2; also see footnote 3). These results provide good support for the a priori model, though it is still important to examine goodness-of-fit indicators.

Insert Table 2 About Here

Goodness-of-fit is evaluated partly by an overall chi-square test. In contrast to traditional significance testing, the researcher may prefer a nonsignificant chi-square value that indicates that the model fits the data. However, for large complex analyses where the sample sizes and the number of measured variables are large, the chi-square test is extremely powerful and will nearly always be statistically significant. Hence, most practical applications of CFA require a subjective evaluation of whether a statistically significant chi-square value is small enough to constitute an adequate fit. Many alternative indices of goodness-of-fit have been developed for this purpose (see Bentler & Bonett, 1981; Cliff, 1983; Fornell, 1983; Joreskog & Sorbom, 1981; Long, 1983a; 1983b; Marsh & Hocevar, 1984a; 1984b; in press; for a general discussion) and 7 such indices are considered here (see Table 3).

Insert Tables 3 & 4 About Here

The chi-square value for the first-order factor model (Model 1A, Table 4) is large and statistically significant. However, the chi-square/df ratio (2.14) and other goodness-of-fit indices suggest that the goodness-of-fit is reasonable. Though not reported, the chi-square/df ratio (2.00) and other indices (e.g., TLI = 0.872) were similar or slightly better when Model 1A was fit to data collected at time 2. This goodness-of-fit could be substantially improved by freeing some of the parameter estimates fixed to be zero; that is by allowing some of the measured variables to have nonzero loadings on factors other than the one they were designed to measure, or by allowing some of the error/uniquenesses to be correlated. In some instances there may be an adequate a priori basis for estimating additional parameters, and thus improving the goodness-of-fit (e.g., Marsh & Hocevar, 1984b). Even when there is no a priori basis, some researchers recommend such procedures (see Tanaka & Huba, 1984), particularly when the goodness-of-fit of the first-order model is poor. Because the original a priori model is judged to adequately fit the data, because there is no a priori justification for freeing additional parameters, because such procedures rarely have a substantial influence on major parameters of interest (if they do, they may not be justified, see Tanaka & Huba, 1984), in order to simplify the interpretation of the first-order factors and maintain a simple structure, and in order to avoid further complication in the description and presentation of the results, these procedures were not used in the present investigation.

First-order Factor Correlations. As described earlier, it is proposed that 11 of 13 SDQ III factors, all but Emotional and Esteem, define five second-order factors. The viability for these proposed factors can be evaluated from an inspection of correlations among the factors derived from the first-order model (Table 2). While the discussion is limited to an inspection of correlations from time 1 data, correlations based on time 2 are similar (see Table 2).

Correlations among the 4 academic factors support the existence of two second-order academic factors. Whereas the Academic and Problem Solving factors are each substantially correlated with Math and Verbal self-concepts, and with each other (r's between 0.48 & 0.62), the Math and Verbal self-concepts are relatively uncorrelated (r = .12) with each other. Correlations between the 4 academic facets and the 8 nonacademic facets are generally much smaller (r's between -.01 & .37).

Correlations among the nonacademic factors provide less clear support for the physical, social, and moral second-order factors. Although Physical Ability and Physical Appearance are substantially correlated (.43), Physical Ability is more highly correlated with Same Sex Relations (.62) and Physical Appearance is as highly correlated with Opposite Sex Relations (.42). While the three social facets are substantially correlated with each other (rs between .32 & .52), they are also substantially correlated with the two physical facets. Although Religion/Spiritual Values is more highly correlated to Honesty than to other factors, the size of the correlation (.27) is modest. This pattern of correlations suggests that support for the higher-order factors defined by nonacademic factors may be more problematic than for those defined by academic facets.

Approaches to Examining Higher-Order Factor Models.

The purpose of HCFA is to explain covariation among the first-order factors with one or more higher-order factors (see footnote 4). It is important to realize that the chi-square goodness-of-fit value for a higher-order model can be no better than the chi-square for the first-order model; the two chi-squares would be equal if the all of the factor covariances could be exactly explained by higher-order factors. Since the number of parameters needed to estimate the higher-order factors is less than for the corresponding first-order factor model, the higher-order model is supported so long as: a) the parameter estimates are defensible in relation to the a priori substantive model that is being tested, b) the goodness-of-fit is reasonable and the chi-square value is not substantially larger than the value observed for the first-order model, and c) technical requirements related to HCFA are met (e.g., the model is identified and there are no Heywood cases). Before examining the higher-order models, several additional points need clarification.

The Comparison of First- and Higher-order Models. The goodness-of-fit of the first-order model in which all factor covariances are estimated (i.e., Model 1A in this study) provides an upper-limit to the goodness-of-fit for a higher-order model -- an optimum or target. Thus, Marsh and Hocevar (in press; also see Table 3) defined the target coefficient (TC) to be the ratio of the chi-squares for a higher-order model and the corresponding first-order model; it varies between 0 and 1.0. If the TC is very high while the Tucker-Lewis index (TLI) is only moderate, then the covariation among the first-order factors is well-explained by the higher-order factors and the lack of fit occurs in the definition of the first-order factors. Hence, TC in conjunction with other indices, allows the researcher to determine whether a lack of fit in a higher-order model occurs with the estimation of first- or second-order factors.

Similarly, a first-order model in which all factor covariances are fixed to be zero (Model 1B, Table 4) represents an absolute lower limit for the goodness-of-fit of any higher-order model. If the goodness-of-fit for Model 1B were to approach that of the Model 1A, it would mean that the first-order factors were nearly uncorrelated. Ironically, this situation would result in an acceptable goodness-of-fit for any higher-order model;

since there would be almost no covariance to explain, any model would be able to "fit" this data. For this reason, it is critical that parameter estimates for the higher-order factor models are examined, and that their goodness-of-fits are compared with the goodness-of-fits of first-order factor models such as 1A and 1B.

Variance in First-order Factors Explained By Higher-order Factors. Even when there are moderate to substantial covariances among some first-order factors and a higher-order model provides a reasonable fit to the data, some of the first-order factors may not be well represented by the higher-order factors. For example, assume that there are four first-order factors, that the first three are highly correlated, and that the fourth is relatively independent of the others. A single higher-order factor, defined primarily by the first three factors, will be well-defined and the model may provide a reasonable fit to the data. However, the fourth factor will not be well represented by the higher-order factor. The fit is reasonable because there is little correlation between it and the other three factors that needs to be explained by the higher-order factor; most of the variance in the fourth factor will appear in a residual variance term that is conceptually like the uniqueness for measured variables in the first-order model (see footnote 1). To the extent that this residual variance approaches zero, the lower-order factor is completely explained by higher-order factors. However, when the residual is substantial and approaches the variance in the lower-order factor, then the lower-order factor is not well represented by higher-order factors. In order to evaluate this characteristic, the Explained Variance Ratio (EVR; see Table 5) is defined as one minus the ratio of the residual variance to the factor variance; it varies between 0 (none of the lower-order variance explained by higher-order factors) to 1.0 (all of the lower-order variance explained by higher-order factors).

Identification in HCFA. Identification is a serious problem in CFA and HCFA. Long (1983a, p. 35) indicates that: "Attempts to estimate models that are not identified result in arbitrary estimates of the parameters and meaningless interpretations." It is possible that some parameters are identified while others are not. It is also possible that the model itself is identified in theory, but that in practice the parameter estimates are not identified due to some problem with the data (e.g., linear dependencies among the measured variables or problems related to correlation matrices produced with pair-wise deletion of missing data). Joreskog and Sorbom (1981) describe checks for identification that are performed by LISREL, but necessary and sufficient conditions for determining identification are generally unknown. While LISREL is able to derive parameter estimates when a model is not identified, these may be uninterpretable; Joreskog and Sorbom recommend that additional constraints be placed on the model so that it is identified (e.g., specifying that two parameter estimates are equal, thus reducing the number of parameters that must be estimated).

In HCFA, second-order factors may be unidentified even when the the first-order factors are identified. For example, two lower-order factors are unable to define a single higher-order factor such that the parameters are identified unless further constraints are imposed; it takes 2 df to define a higher-order factor from two lower-order factors, but there is only one df represented by the one covariance between the two first-order factors. An arbitrary, though often reasonable solution is to impose an equality constraint by specifying that the two lower-order factors have the same loading on the higher-order factor. Even when there are 3 lower-order factors, a higher-order factor is "just identified" in that it takes all three df representing the lower-order covariances to estimate the higher-order factor and there are none left to test the model. If there are three first-order factors then the first-order model in which these factors are allowed to be correlated will produce the same goodness-of-fit as a higher-order model in which the three first-order factors are allowed to define a single higher-order factor and no added constraints are imposed. Although the higher-order loadings for such a model are identified, and their substantive interpretation may be useful, it makes no sense to argue that the higher-order model is preferable to the corresponding first-order model, or vice-versa. This problem of identification may impose a serious limitation on the use of HCFA in that a large number of first-order factors are needed to estimate even a moderately complicated hierarchical structure.

The problem of identification may be a limitation in the application of HCFA rather than an inherent flaw in the logic of models that are not identified within a HCFA framework. For example, Shavelson et al. (1976) proposed that a nonacademic and an academic self-concept combined to form a hierarchical general self-concept. While this model is substantively reasonable, it is not identified within a HCFA framework without further constraints (i.e., it takes three lower-order factors to define a higher-order factor). Similarly, Shavelson and Marsh (in press) recognized that a second-order factor model consisting of three second-order factors and the covariances among the second-order factors was equivalent, within a HCFA framework, to a third-order model in which the three second-order factors combined to form a hierarchical general self-concept.

The Law of Parsimony. The law of parsimony states that if two models describe the data equally well, then the conceptually more simple model is preferred. In HCFA this is interpreted to mean that if the goodness-of-fits for two models are comparable, then the model that requires the fewest parameters (i.e., has a larger df) is preferred. In fact, models that require fewer parameters generally are conceptually more simple. However, alternative HCFA models are often specified so that the chi-square for one represents the lowest possible chi-square that could be obtained by a second model (i.e., the models are nested). For example, a higher-order model can never have a smaller chi-square than the corresponding first-order model, and in this sense can never do any better than it. While tests of statistical significance are available, they are typically so powerful that the more parsimonious model will usually be rejected in a strict statistical sense. Instead, researchers must rely on the comparison of alternative models in terms of subjective indicators of goodness-of-fit. Researchers have tended to use the same goodness-of-fit indicators in HCFA as in CFA, though it is likely that new indicators (e.g., the TC described above) and rules of thumb will be developed as HCFA becomes more widely applied.

In order to apply the rule of parsimony in the present investigation, a series of a priori hierarhical models was specified to explain responses to the SDQ III. The logic used to formulate alternative models was to begin with the most parsimonious model, a single higher-order factor defined by all SDQ III factors, and to gradually increase the complexity of the models. The theoretical model postulated earlier to describe responses to the SDQ III (Model 6A in Figure 1) proposes a complicated ordering of self-concepts. If the goodness-of-fit of Model 2A, with only one higher-order factor, approaches that of the more complicated model, then there is evidence against the more complicated model. The formulation of alternative models will depend on substantive issues in the particular application, but at least two types of alternative models will always be appropriate. First, a complicated hierarchical model should always be compared with a model in which only one higher-order factor is hypothesized. Second, whenever four or more lower-order factors define a single higher-order factor, a viable alternative model is one where the lower-order factors are correlated and no higher-order factor is hypothesized. These alternatives are illustrated below.

Tests of A Priori Higher-Order Models to Explain Responses to the SDQ III.

Five alternative HCFA models (see Figure 1) were formulated to explain responses to the SDQ III and were tested with LISREL. For each model, the first-order factors were defined as in Model 1A except that the covariances among first-order factors were explained in terms of higher-order factors. Although the first-order factor loadings for these hierarchical models are not presented, they were similar to those shown in Table 1 for all the models. The alternative models were formulated to test substantive issues described earlier. The first two HCFA models proposed a single higher-order factor (2A), and two higher-order factors defined by academic and nonacademic components (3A). In the next HCFA (4A) model the academic component was divided into the two components suggested by the Shavelson and Marsh revision. In the last two HCFA models, second-order factors were defined by the physical and social components (Model 5A) and by the moral components (Model 6A). For Models 4A, 5A and 6A, where a third-order factor was defined by four or more lower-order factors, alternative models were posited in which there was no higher-order factor and covariances among the lower-order factors were estimated. A discussion of the goodness-of-fit for each model is presented below.

Insert Table 5 About Here

<u>One General Factor.</u> In Model 2A (Figure 1) a single second-order factor was proposed to explain relations among first-order factors. Twelve of 13 first-order factors, all but Religion/Spiritual Values, load significantly on the general factor. For example, the factor loading for Math is .43, the standard error of this parameter estimate is .07, and the t-ratio (.41/.07 = 5.86) is highly significant. The difference in chi-square values for models 2A and 1A (1786 - 1333 = 453) is statistically significant when evaluated against the difference in df (689 - 624 = 65), and the ratio of the two is substantial (453/65 = 6.97). While the TLI (.828) is reasonably high, the TC (.694) suggests that much of the covariation among the first-order factors is unexplained. The EVRs (Table 5) for the 13 first-order factors indicate that the 4 academic and 2 moral facets are not well represented by the general factor. For example, the factor variance for the Math factor is .85 (from Table 1), the residual variance unexplained by the higher-order factor is .73 (Figure 1), and the EVR is .147 (1 - (.73/.85); see Table 5). In summary, though a single higher-order factor is able to explain much of the covariance among first-order factors, the model must be rejected (particularly in comparison with models described below).

<u>Two Higher-order Factors.</u> In Model 3A two second-order factors are proposed, a nonacademic factor and an academic factor. In this model (Figure 1) the Esteem facet is allowed to load on both second-order factors, and the two second-order factors are correlated. Tests of statistical significance and subjective goodness-of-fit indicators indicate that Model 3A does better than 2A but not as well as 1A. Inspection of the EVRs (Table 5) indicates that the 4 academic facets are much better represented in 3A

than in 2A, while there is a small improvement or little change for the nonacademic facets. These results clearly support the separation of academic and nonacademic self-concept, but Model 3A is also inferior to other models described below.

Three Second-Order Factors and One Third-Order Factor. In Model 4A the academic factor from Model 3A is divided into math/academic and verbal/academic factors as proposed in the revision of the Shavelson model. A third-order, hierarchical general self is defined by the three second-order factors, and the first-order Esteem factor. However, preliminary analyses suggested that this model was not identified, and that the problem occurred in the definition of the two second-order academic factors. The model was identified, however, when the Academic and Problem Solving factors were required to load equally on the math/academic and verbal/academic factors (see Figure 2 & footnote 5). The chi-square for Model 4A is significantly smaller than the chi-square for 3A, and the goodness-of-fit indices are better. Inspection of the EVRs for the first-order factors indicates that the Math and Verbal self-concepts are better explained by Model 4A than by Model 3A. However, this is due to variance explained by the two second-order academic factors and not the third-order general factor. Inspection of the factor loadings on the third-order factor (Figure 1) and the EVRs for the second-order academic factors (Table 5) indicate that the academic factors are not well represented by the third-order general factor. Thus, while Model 4A represents a significant improvement over Model 3A, neither the academic nor the moral factors are well represented by the third-order general factor.

Model 4B differs from 4A in that the four lower-order factors in model 4A are not hypothesized to define a hierarchical general self-concept -- the six factor covariances among these lower-order factors are merely estimated in the analysis. The chi-squares for models 4A and 4B differ by only one, and this is not statistically significant. Thus, the six covariances among the lower-order factors are well described by a single factor, the hierarchical general self.

Four Second-order Factors and One Third-order Factor. Model 5A differs from Model 4A in that two new second-order factors are defined; one for the two physical self-concepts and one for the three social self-concepts. The third-order hierarchical general self (see Figure 1) is defined by four second-order factors and four first-order factors. Since the second-order physical self-concept was defined by only two first-order factors, an equality constraint had to be imposed. Because of this equality constraint the df for Models 4A and 5A are the same, and this complicates their comparison. However, the chi-square and goodness-of-fit indicators for Model 5A are marginally better than for 4A even though the differences are modest. The EVRs for the two first-order physical factors and the three first-order social factors are modestly higher for Model 5A than 4A. The factor loadings for the second-order factors on the third-order factor, and also the EVRs for the second-order factors, indicate that physical and social factors are well represented by the general factor while the academic and the moral (second-order) factors are not.

The hierarchical general self in Model 5A is defined by eight lower-order factors. In Model 5B, no third-order factor is posited, and the 28 covariances among the lower-order factors were estimated. The difference in chi-squares (61, df_{diff} = 20) is statistically significant, and moderately large. This suggests that a significant proportion of the covariation among the lower-order factors is not represented by the hierarchical general self. This apparently occurs because the modest correlation (r = .27) between Religion/Spiritual Values and Honesty, and the large correlation between the second-order physical and social factors (r = .94), were not adequately explained in terms of the hierarchical general factor.

Five Second-order Factors and One Third-Order factor. Model 6A is closest to the structure originally proposed to explain responses to the SDQ III. It differs from Model 5A in that one new second-order factor is defined by the two moral value self-concepts. Since the moral value self-concept is defined by only two first-order factors, it was necessary to impose another equality constraint. The third-order general factor is defined by five second-order factors, and by the Emotional and Esteem (first-order) factors. Even though the df for Model 6A are the same as for Models 4A and 5A, the chi-square and goodness-of-fit indices are marginally better (Table 4). The EVRs for the two first-order moral values factors, though better in Model 6A than the other models, are still only modest. The factor loadings on the third-order factor, and the EVRs for the second-order factors, again show that the third-order factor represents primarily physical, social, and Esteem factors, but not the academic and moral factors. Nevertheless, Model 6A appears to represent a modest improvement over the other models.

The hierarchical general self in Model 6A is defined by 7 lower-order factors. In Model 6B, no third-order factor was posited, and the 21 covariances among the lower-order factors were estimated. The difference in chi-squares (34, df_{diff} = 14) is statistically significant, but only modest. The addition of the second-order moral factor was able to account for the correlation between Religion/Spiritual Values and Honesty (see discussion of the comparison of Models 5A & 5B), even though this second-order factor was not strongly represented in the hierarchical general self-concept. Even though the second-order physical and social components are well represented by the hierarhical general self-concept, the extremely high correlation among the two second-order factors could not be explained in terms of the third-order factor.

Tests of Models From Time 2. Each of the models described above was also tested with data from time 2 (see Note in Table 4). Though not a major focus of the study, a brief discussion of these findings is informative. Parameter estimates for time 2 data were similar and the goodness-of-fit tests were slightly better. There was, however, one major difference in the results based on data from time 2. The chi-squares and goodness-of-fit indicators for Models 4A, 5A and 6A were nearly identical. Again, the pattern of equality constraints meant that each of these models had the same df and this makes their comparison difficult. Nevertheless, perhaps even more than with data from time 1, data from time 2 suggest that these three models are equally able to explain the data.

An Additional A Posteriori Model. The formulation of the models described above, except perhaps the equality constraints needed to define the second-order academic factors, was a priori. However, the previous discussion and empirical findings (see footnote 5) suggest additional alternatives. The very high correlation observed between the second-order physical and social factors could not be explained adequately in terms of the third-order factor. Also allowing the Problem Solving to load on the hierarchical general self-concept, in addition to the second-order academic factors, improved the fit of the hierarchical models. In order to test these possibilites, Model 6A was altered so that the 3 social and 2 physical facets defined only one second-order factor instead of two, and this produced a smaller chi-square (1501, df= 686). Then, in a subsequent analysis, the Problem Solving facet was allowed to contribute directly to the hierarchical general self-concept, and this further improved the chi-square (1477, df = 685). The a posteriori model based on both of these alterations is presented as Model 7A in Table 4 and Figure 1. The a posteriori nature of Model 7A suggests that it must be interpreted cautiously. However, these alterations also produced comparable improvements in the chi-square values for time 2 data, thus providing additional support for Model 7A.

Summary and Implications

The Substantive Issue -- A Multifaceted, Hierarchical Self-Concept.

The Shavelson model posits that self-concept is multidimensional, and the results of the present investigation, as has been the case with all SDQ research, provides strong support for this proposal. The SDQ III is designed to measure 13 facets of self-concept; these were clearly identified in the first-order factor model, and even the simple structure imposed in that model provided a reasonable fit to the data from time 1 and from time 2.

The Shavelson model posits that self-concept is hierarchically ordered, and this proposal was supported. Shavelson et al. (1976) also proposed one possible representation of what the hierarchy of self-concept facets might be, and the hierarchy for academic self-concept was revised by Shavelson and Marsh (in press). The hierarchy proposed to describe responses to the SDQ III, and additional constraints necessary to actually test the proposal with HCFA, differed from that revision in a number of ways. However, the most important differences were the inclusion of Religion/Spiritual Values and Honesty facets, and their representation as moral values, and the inclusion of the general esteem factor.

The proposed hierarchy of academic facets was supported in that two second-order academic facets -- math/academic and verbal/academic -- were required instead of just one as originally proposed by Shavelson. As before, this was necessary because of the relative lack of correlation between Math and Verbal self-concepts. These findings provide strong support to the revision proposed by Shavelson and Marsh (1984), particularly since the study took place in a nonacademic environment and the sample was comprised of young adults who were primarily nonstudents. The inclusion of the Problem Solving factor on the SDQ III provided some complications for the proposed hierarchy of academic self-concepts; first, because it had not previously been considered in proposed hierarchies; and second, because it apparently contributes directly to a hierarchical general self-concept beyond its contribution through the second-order academic self-concepts. Nevertheless, its representation in Model 7A is reasonable, and is also supported in the analysis of data from time 2.

Support for the second-order facets proposed to explain the physical and social facets is more tenuous. The examination of correlations among the first-order facets clearly foreshadowed some of the difficulties. Correlations between some physical and some social facets were higher than correlations among the physical and among the social factors (see footnote 6). This problem was also illustrated by the extremely high correlation between the second-order physical and social factors; it was .94 for both Models 5B and 6B. There is also an intuitive consistency to the pattern of correlations among the physical and social factors that may be inconsistent with the logic of the Shavelson model. In particular, it is reasonable that Physical Appearance is more strongly related to Opposite Sex Relations than to Physical Abilities. The support for Model 7A, where the second-order physical and social factors were collapsed into a single second-order factor, provides further support for these observations. It should also be noted that for the best fitting model proposed by Shavelson and Marsh (in press), two physical factors and two social factors were incorporated into a single second-order factor and no attempt was made to test for separate physical and social second-order factors. In this respect, their model was more like Model 7A than 6A. Although further research is clearly warranted and alternative formulations may exist, these findings suggest that the social and physical facets combine to form a single second-order component.

Support for the second-order moral factor defined by Religion/Spiritual Values and by Honesty, must also be interpreted cautiously. Although neither of these facets, nor their incorporation into a second-order factor, was proposed by Shavelson, et al. (1976), such facets have frequently been posited by other self-concept theorists starting with William James (1890). Furthermore, on an earlier version of the SDQ III that did not contain these facets, respondents indicated these areas as ones that were important to how they felt about themselves that had not been included. The inclusion of these facets appears to be justified in terms of theory and empirical results, but their combination into a second-order factor may be more problematic. Although these factors tend to be more highly correlated to each other than to other factors, the size of the correlation (.27 and .28 for times 1 and 2) is modest. Nevertheless, only the hierarchical models that incorporated this second-order factor were able to explain the correlation between the two facets (see discussion of Model 4A). In this respect their incorporation into a second-order factor is supported.

In each of the hierarchical models there was strong support for a hierarchical ordering of the SDQ III facets; most of the covariation among the first-order facets could be explained in terms of the hierarchy, and the hierarchical general self-concept was well defined. However, even this finding must be interpreted cautiously. In particular, this finding should <u>not</u> be interpreted to mean that each of the first-order factors was well represented by the hierarchical general self-concept; this was definitely not the case. In each model, the hierarchical general self-concept was defined primarily in terms of the physical and social factors, Emotional self-concept, and general Esteem. The covariation among the four academic factors was well represented by the two second-order academic factors, even though these second-order factors contributed only modestly to the hierarchical general factor. Similarly, the modest covariation among the two moral factors was well represented by a second-order factor, even though it contributed little to the hierarchical general self. This caution does not, however, undermine support for the hierarchy -- quite the contrary. If every first-order factor was well-represented by the hierarchical general self-concept, than support for the a single higher-order factor in Model 2A would have been much stronger and there would have been no need to consider a more complicated hierarchy. Thus, it the inability of a single higher-order factor to represent each of the first-order factors that dictates the need for a more complicated hierarchy (see footnote 7).

The hierarchy described here also differs from those previously examined, or even proposed, in that it includes both general Esteem and a hierarchical general self-concept. For purposes of this study it was hypothesized that the general Esteem factor contributes directly to the hierarchical general self-concept. Most of the covariation between Esteem and other first-order factors could be explained by this formulation. Esteem, along with Emotional self-concept, and the second-order physical and social factors, were the primary determinants of the hierarchical general self. In each of the hierarchical models the hierarchical general self correlated about .90 with Esteem. These findings support the proposed model, and demonstrate that general Esteem and the hierarchical general self-concept are highly correlated.

The Application of HCFA.

In addition to the substantive issues, the demonstration of an application of HCFA represents an important contribution of this study. The use of HCFA, instead of exploratory factor analysis, is clearly preferred. Since researchers are rarely provided with concrete guidelines for conducting HCFA, the following recommendations are proposed:

1. HCFA should begin with a clearly articulated theoretical model of the proposed hierarchy, a well-defined set of first-order factors that form the basis of the hierarchy, and a sufficient number of first-order factors to adequately test the hierarchy. In addition to the hypothesized model, a series of a priori alternative models should be formulated to test substantive issues and to test more parsimonious hierarchies.

2. Support for the first-order factor structure underlying responses to the measured variables is a prerequisite for testing higher-order structures. If the first-order factors are not well-defined, or if they are unable to fit the data, then tests of hierarchical structures are moot. An examination of correlations among the first-order factors provides insight into the hierarchical structure. Finally, the fit of first-order models

provides an important basis of comparison for testing the hierarchical structures. Researchers should avoid using total scores as the measured variables in HCFA. When each first-order factor is well defined by many indicators, as in this study, the formation of three or more subscales to represent each first-order factor appears to be reasonable.

3. There are no absolute criteria for evaluating goodness-of-fit. The evaluation of the goodness-of-fit of CFA and HCFA should begin with an examination of parameter estimates in terms of the substantive model. If factors proposed in the theoretical model are not well-defined, then the ability of the model to fit the data may be irrelevant. Tests of statistical significance are relevant, but their power -- particularly when there are many measured variables and the sample size is large -- is so great, that researchers generally must use subjective criteria to determine whether a statistically significant lack of fit is substantial enough to be practically important. Subjective goodness-of-fit indicators are also relevant, but they are more useful in comparing alternative models than in determining absolute guidelines of goodness-of-fit.

This set of recommendations emphasizes the use of HCFA as the means for testing an a prior hierarchy that is based on a theoretical model with measured variables specifically designed to test the hierarchy. The use of HCFA may also be justified when no a priori hierarchy exists, and the purpose is to explore possible hierarchies. In such an application, Tanaka and Huba (1984) suggest the use of exploratory factor analysis to form a basis for HCFA models to be tested with one set of data and then to be replicated with another set of data. The use of empirical results from one set of data, instead of theory, to formulate HCFA models to be tested with another set of data may be reasonable if the set of measured variables provide a clearly defined set of first-order factors. However, the use of hierarchical factor analysis, whether exploratory or confirmatory, to infer a structure from an ill-defined collection of first-order factors is likely to encounter the same problems as "blind" factor analyses. For this reason, it is recommended that theory be the basis of the design/selection of measured variables, for the formulation of hierarchical models, and for the evaluation of results in HCFA studies.

Footnotes

1 -- This recommendation is based on the assumption that each first-order factor is inferred from multiple indicators that are intended to measure one specific factor as is the case with the SDQ III. If there is an a priori basis for hypothesizing that a measured variable should contribute to the definition of more than one variable, then the first-order factors should be defined accordingly. Even when the basis is a posteriori, the practice is justifiable when the original first-order factor structure does not fit the data very well. In such instances, particularly when many changes in the original structure are needed to obtain an acceptable fit to the data, then the proposed changes should be tested with new data. The best strategy is to use an instrument that has a well-defined factor structure.

2 -- The three LISREL design matrices used to fit the first-order model (model 1 in Table 1) were LAMBDA Y (factor loadings), PSI (factor variance/covariance matrix), and THETA EPSILON (error/uniquenesses). For this study THETA EPSILON was always constrained to be a diagonal matrix (i.e., error/uniquenesses are uncorrelated). In the higher-order models, the second-order factor loadings were estimated in BETA and factor residuals were estimated in the diagonal of the PSI matrix. For the higher-order models each diagonal element in PSI is a factor variance when that factor is not incorporated into a higher-order factor, otherwise it is a factor residual -- the unique variance in that factor that is not explained by higher-order factors. When a lower-order factor is incorporated into a higher-order factor, each of the off-diagonal elements in PSI that involve that lower-order factor is set to 0. Thus, for models 4A, 5A, and 6A, PSI is a diagonal matrix in which all the elements except the factor variance estimate of the hierarchical general self factor are factor residual estimates (see Marsh & Hocevar, in press, for further discussion about how the LISREL design matrices were used to specify the first and second-order factor models).

3 -- Tests for each of the models was performed on data from time 1 and from time 2, and the two sets of results were compared. These do not constitute tests of factorial invariance -- tests where parameters are estimated for each data set subject to the constraint that some or all of the parameter estimates are the same. However, a series of tests of the invariance of parameter estimates for the first-order model (Model 1A) was performed on covariance matrices for responses from times 1 and 2. The chi-square and df values were: a) 2564 and 1248 when no invariance constraints were imposed; b) 2796 and 1404 when all 156 estimated parameters were specified to be invariant; c) 2654 and 1365 when factor loadings, factor variances and

factor covariances were invariant; and d) 2609 and 1274 when just factor loadings were invariant. Although the difference between tests of total invariance and of no invariance is statistically significant, the chi-square/df ratio ((2796 - 2564) / (1404 - 1248) = 1.48) is small. Furthermore, the difference between tests of no invariance and the invariance of factor loadings, factor variances and factor covariances does not even reach statistical significance. These tests provide strong support for the invariance the factor structure, factor variances and factor covariances across the two sets of responses (see Marsh & Hocevar, in press, for further discussion of the interpretation of factorial invariance), and also suggest that there would be reasonable support for invariance of parameter estimates for each of the higher-order models if these tests had been conducted.

4 -- In some applications of HCFA (e.g., Tanaka & Huba, 1984) researchers begin with first-order factors that are defined by a single score; an average of responses to items designed to define a scale, or factors that are defined by a single item. In such applications the first-order factors are really measured variables rather than latent constructs that are inferred from multiple indicators, and what is called the first-order model in the present application is completely eliminated. This alternative approach may be reasonable in preliminary studies, or studies where researchers do not have access to the original data, but not when multiple indicators of each of the first-order factors are available. It is imperative that tests of higher-order structures are based on first-order factors that are well defined. Thus rigorous tests of the hypothesized first-order structure are essential, and these tests require multiple indicators of each first-order factor. It is ironic that researchers would apply powerful HCFA techniques to first-order factors that are inferred from single indicators, or inferred from multiple indicators that are averaged in accordance with a hypothesized factor structure that is not tested.

5 -- Models that imposed no equality constraints, or that imposed only one equality constraint, on the academic factors in Model 3A did not converge when tested with data from time 1 or time 2; the residual variance estimates for the Verbal or Math factors were negative. The requirement that the Academic and Problem Solving factors contribute equally to the two second-order academic factors (see Figure 1) is pragmatic, but it is also intuitively logical and can be justified empirically. LISREL computes a modification index (see Joreskog & Sorbom, 1981; Marsh & Hocevar, in press) for every parameter, particularly those that are fixed at a constant value or constrained, that is an estimate of how much the chi-square value would

change if a the parameter were free to be estimated -- in this case if the equality were not imposed. If a modification index is less than five, then freeing that parameter will have almost no impact on the goodness-of-fit of the solution. The modification indices of the constrained factor loadings on the second-order academic factors never exceeded 1.3 for Model 3A, for any of the subsequent models that used this constraint, or for solutions for any of these models with data from time 2, thus supporting the use of these equality constraints.

6 -- This observation, that correlations between some physical and social factors are as high or higher than correlations among facets within each of the categories, is also consistent with other SDQ research using the SDQ III (Marsh & Hocevar, in press; Marsh & O'Niell, 1984), the SDQ II (with high school students; Marsh, Parker & Barnes, in press), and the SDQ (with preadolescents; Marsh, 1984a).

7 -- The implications of the inability of the academic and moral factors to be accounted for by the hierarchical general self depends on the ordering of the hierarchy in the hypothesized model. For purposes of the theoretical discussion in this study I assumed that the hierarchical general self can be explained in terms of lower order factors, and thus avoided specifying a causal ordering between lower- and higher-order facets. This inability apparently has more negative implications for a model that makes the stronger assumption that the hierarchical general self "causes" the lower-order facets than one that proposes that the hierarchical self is caused by the lower-order facets or that no causal ordering exists. If a hierarchical general self is posited to "cause" a lower-order factor but the two are nearly uncorrelated, than the postulated causal relation is not supported. It should also be noted that while HCFA models are often depicted in path analytic terms such that a higher-order factor causes a lower-order factor, this is not inherent in the application of factor analysis and it may not be consistent with the substantive nature of the particular application.

REFERENCES

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

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

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

Burns, R. B. (1979). The self-concept: Theory, measurement, development and behaviour. London: Longman.

Cliff, N. (1983). Some cautions concerning the application of causal modeling methods. Multivariate Behavioral Research, 18, 115-126.

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

Dusek, J. B., & Flaherty, J. F. (1981). The development of self-concept during the 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.

Fornell, C. (1983). Issues in the application of covariance structure analysis. Journal of Consumer Research, 9, 443-448.

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

Huba, G. J., & Bentler, P. M. (1982). On the usefulness of latent trait modeling in testing theories of naturally occurring events. Journal of Personality and Social Psychology, 43, 604-611.

James, W. (1890). The principles of psychology. New York: Holt, Rinehart & Winston.

Joreskog, K. G. (1980). Structural analysis of covariance and correlation matrices. Psychometrika, 43, 443-477.

Joreskog, K. G. (1971). Simultaneous factor analysis in several populations. Psychometrika, 36, 409-426.

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. (1983a) Confirmatory factor analysis: A preface to LISREL. Beverly Hills, CA: Sage.

Long, K. S. (1983b) Covariance structure models: An introduction to LISREL. Beverly Hills, CA: Sage.

Marsh, H. W. (1984a). Self Description Questionnaire (SDQ): An instrument for measuring multiple dimensions of preadolescent self-concept. Department of Education, University of Sydney, Australia.

Marsh, H. W. (1984b). Verbal and math self-concepts: An internal/external frame of reference model. (In Review) Department of Education, The University of Sydney, Australia.

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, 75, 940-956.

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

Marsh, H. W., & Hocevar, D. (1983). Confirmatory factor analysis of multitrait-multimethod matrices. Journal of Educational Measurement, 20, 231-248.

Marsh, H. W., & Hocevar, D. (1984a). The comparison of factors from different groups with exploratory and confirmatory factor analysis. (In Review).

Marsh, H. W. & Hocevar, D. (1984b). The factorial invariance of students' evaluations of college teaching. American Educational Research Journal.

Marsh, H. W. & Hocevar, D. (in press). 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.

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

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

Marsh, H. W., Richards, G., & Barnes, J. (in press). Multidimensional self-concepts: The effect of participation in an Outward program. Journal of Personality and Social Psychology.

Marsh, H. W. and Shavelson, R. J. (in press) Self-Concept: Its multifaceted, hierarchical structure. Educational Psychologist.

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. & Butler, S. (1983). Self-concept: Reliability, dimensionality, validity, and the measurement of change.

Journal of Educational Psychology, 75, 772-790.

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

Olson, G. H. (1983). Covariance structure analysis involving measured variables assumed to have higher-order factor structures. Paper presented at the Annual Meeting of the American Education Research Association, Division D, Montreal, Canada.

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

Rosenberg, M. (1979). Conceiving the self. New York: Basic Books.

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). 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. Erlbaum.

Soares, L. M., & Soares, A. T. (1982). Convergence and discrimination in academic self-concepts. Paper presented at the 20th Congress of the International Association of Applied Psychology, Edinburgh, Scotland, July, 1982.

Tanaka, J. S., & Huba, G. J. (1984). Confirmatory hierarchical factor analyses of psychological distress measures. Journal of Personality and Social Psychology, 46, 621-635.

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

Wylie, R. C. (1974). The self-concept (Rev. ed., Vol. 1) Lincoln: University of Nebraska Press.

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

Table 1
CFA Model 1: 13 First-Order Factors With No Higher-Order Factors

	Factor Loadings													Error/
	Math	Verb	Acad	Prob	Pabl	Appr	Ssex	Osex	Prnt	Relg	Hnst	Emot	Estm	Uniqueness
1	.96	0	0	0	0	0	0	0	0	0	0	0	0	.21
2	.98	0	0	0	0	0	0	0	0	0	0	0	0	.18
3	1.0	0	0	0	0	0	0	0	0	0	0	0	0	.15
4	0	.85	0	0	0	0	0	0	0	0	0	0	0	.47
5	0	.86	0	0	0	0	0	0	0	0	0	0	0	.46
6	0	1.0	0	0	0	0	0	0	0	0	0	0	0	.27
7	0	0	.86	0	0	0	0	0	0	0	0	0	0	.40
8	0	0	1.04	0	0	0	0	0	0	0	0	0	0	.12
9	0	0	1.0	0	0	0	0	0	0	0	0	0	0	.19
10	0	0	0	1.08	0	0	0	0	0	0	0	0	0	.40
11	0	0	0	1.11	0	0	0	0	0	0	0	0	0	.36
12	0	0	0	1.0	0	0	0	0	0	0	0	0	0	.48
13	0	0	0	0	1.07	0	0	0	0	0	0	0	0	.22
14	0	0	0	0	1.14	0	0	0	0	0	0	0	0	.12
15	0	0	0	0	1.0	0	0	0	0	0	0	0	0	.32
16	0	0	0	0	0	.98	0	0	0	0	0	0	0	.32
17	0	0	0	0	0	.96	0	0	0	0	0	0	0	.35
18	0	0	0	0	0	1.0	0	0	0	0	0	0	0	.29
19	0	0	0	0	0	0	.80	0	0	0	0	0	0	.50
20	0	0	0	0	0	0	1.02	0	0	0	0	0	0	.19
21	0	0	0	0	0	0	1.0	0	0	0	0	0	0	.23
22	0	0	0	0	0	0	0	.89	0	0	0	0	0	.34
23	0	0	0	0	0	0	0	.95	0	0	0	0	0	.24
24	0	0	0	0	0	0	0	1.0	0	0	0	0	0	.16
25	0	0	0	0	0	0	0	0	.95	0	0	0	0	.29
26	0	0	0	0	0	0	0	0	.93	0	0	0	0	.32
27	0	0	0	0	0	0	0	0	1.0	0	0	0	0	.20
28	0	0	0	0	0	0	0	0	0	1.09	0	0	0	.11
29	0	0	0	0	0	0	0	0	0	1.05	0	0	0	.18
30	0	0	0	0	0	0	0	0	0	1.0	0	0	0	.25
31	0	0	0	0	0	0	0	0	0	0	1.49	0	0	.40
32	0	0	0	0	0	0	0	0	0	0	1.16	0	0	.64
33	0	0	0	0	0	0	0	0	0	0	1.0	0	0	.73
34	0	0	0	0	0	0	0	0	0	0	0	.97	0	.24
35	0	0	0	0	0	0	0	0	0	0	0	.93	0	.31
36	0	0	0	0	0	0	0	0	0	0	0	1.0	0	.17
37	0	0	0	0	0	0	0	0	0	0	0	0	.99	.18
38	0	0	0	0	0	0	0	0	0	0	0	0	1.05	.08
39	0	0	0	0	0	0	0	0	0	0	0	0	1.0	.17

Table 1 (continued)

CFA Model 1: 13 First-Order Factors With No Higher-Order Factors

Factor Variance/Covariances

	Math	Verb	Acad	Prob	Pabl	Appr	Ssex	Osex	Prnt	Relg	Hnst	Emot	Estm
Math	.85												
Verb	.10	.73											
Acad	.52	.41	.81										
Prob	.32	.33	.36	.52									
Pabl	.20	.15	.19	.17	.68								
Appr	.14	.23	.15	.19	.30	.71							
Ssex	.12	.20	.14	.17	.45	.27	.78						
Osex	.11	.29	.19	.21	.26	.32	.42	.84					
Prnt	.12	.14	.19	.16	.28	.23	.38	.26	.80				
Relg	.00	-.01	.05	.02	.06	.06	.05	-.01	.01	.76			
Hnst	.10	.08	.10	.09	.08	.08	.04	.05	.09	.12	.27		
Emot	.20	.21	.20	.22	.31	.32	.41	.32	.38	-.03	.11	.81	
Estm	.21	.27	.25	.35	.35	.42	.45	.47	.39	.09	.14	.59	.83

Note: Parameters with the values of 0 and 1.0 were fixed in order to define the model. Standard errors for all estimated parameters vary from .02 to .08. Thus, all parameters, with the exception of some of the factor covariances, differ substantially from zero. Math=Math; Verb=Verbal; Acad = Academic; Prob=Problem Solving; Appr=Physical Appearance; Ssex=Same Sex Relationships; Osex=Opposite Sex Relationships; Prnt=Parent Relationships; Relg=Religion/Spiritual Values; Hnst=Honesty; Emot=Emotional Stability; Estm=General Self Esteem.

BEST COPY AVAILABLE

Table 2

Correlations Among First-Order Factors in Model 1A For: Responses From Time 1 (below diagonal) and Time 2 (above diagonal)

Factors

	Math	Verb	Acad	Prob	Pabl	Appr	Ssex	Osex	Prnt	Relg	Hnst	Emot	Estm
Math	1.00	.14	.63	.43	.30	.20	.15	.11	.17	.01	.21	.24	.19
Verb	.12	1.00	.55	.53	.28	.34	.30	.35	.19	-.03	.18	.30	.32
Acad	.62	.53	1.00	.51	.30	.26	.18	.22	.28	.09	.29	.27	.25
Prob	.48	.54	.55	1.00	.34	.34	.29	.36	.22	-.02	.30	.35	.47
Pabl	.26	.21	.26	.28	1.00	.46	.57	.39	.42	.05	.30	.40	.46
Appr	.19	.32	.19	.32	.43	1.00	.31	.48	.30	.08	.24	.46	.61
Ssex	.15	.26	.17	.27	.62	.36	1.00	.55	.37	.05	.21	.50	.55
Osex	.14	.37	.23	.32	.35	.42	.52	1.00	.27	-.04	.13	.40	.56
Prnt	.15	.18	.23	.25	.38	.32	.48	.32	1.00	-.01	.25	.43	.45
Relg	-.01	-.02	.07	.03	.08	.08	.07	-.02	.01	1.00	.19	-.04	.03
Hnst	.22	.18	.22	.24	.18	.19	.08	.11	.20	.27	1.00	.28	.25
Emot	.24	.27	.25	.34	.42	.42	.52	.39	.47	-.04	.24	1.00	.69
Estm	.25	.35	.30	.54	.47	.55	.56	.56	.47	.11	.29	.72	1.00

Note: Correlations represent the results of the standardized solution for Model 1A for time 1 (below the diagonal) and for time 2 (above the diagonal). See Note in Table 1 for the abbreviations of the factor names.

BEST COPY AVAILABLE

Table 3

Description of Seven Goodness-of-fit Indicators

Indicator	Description
X^2/df	The ratio of the chi-square to the degrees-of-freedom (df)
GFI	The Goodness of Fit Index is a measure of the relative amount of variances and covariances jointly accounted for the model and varies between 0 and 1 (Joreskog & Sorbom, 1981, I.40-41).
AGFI	The Adjusted Goodness of Fit Index is like the GFI, but it is adjusted for the number of parameters that are estimated (Joreskog & Sorbom, 1981, p. i.40-41).
RMS	The Root Mean Square residual is a measure of the average of residual variances and covariances for the original measured variables (Joreskog & Sorbom, 1981, p. I.41).
TLI	The Tucker-Lewis index is: $1 - [(X^2/df_n)/(X^2/df_m)]$. The X^2/df_n is the chi-square/df ratio for the null model, while X^2/df_m is the corresponding ratio for the model being tested (see Bentler & Bonett, 1980).
BBI	The Bentler-Bonett Index is: $1 - [(X^2_m) / (X^2_n)]$. The X^2_n and X^2_m are the chi-square values for the null and tested models (see Bentler & Bonett, 1980).
TC	The Target Coefficient, a measure of a higher-order model's ability to explain the covariation among first-order factors, is: $1 - [(X^2_{fo}) / (X^2_{ho})]$ where the X^2_{fo} and X^2_{ho} stand for the chi-square values for the first-order model (Model 1A) and the higher-order model being tested (see Marsh & Hocevar, in press).

BEST COPY AVAILABLE

Table 4

Goodness of Fit Indicatorsa for All CFA Models

Model	χ^2 (df)	χ^2/df	GFI	AGFI	RMS	TLI	BBI	TC
0	11183(741)	15.09	.235	.195	.285	.000	.000	---
1A	1333(624)	2.14	.634	.542	.042	.858	.881	1.000
1B	2811(702)	4.00	.505	.450	.235	.734	.749	---
2A	1786(689)	2.59	.608	.556	.083	.828	.840	.747
3A	1613(687)	2.35	.617	.566	.065	.844	.862	.826
4A	1539(686)	2.24	.627	.569	.061	.852	.862	.867
4B	1538(684)	2.25	.621	.568	.061	.851	.862	.867
5A	1527(686)	2.23	.622	.571	.061	.852	.863	.873
5B	1466(666)	2.20	.625	.561	.054	.854	.869	.909
6A	1515(686)	2.21	.623	.572	.061	.854	.865	.880
6B	1481(672)	2.20	.623	.563	.057	.854	.868	.900
7A	1476(685)	2.15	.624	.572	.057	.857	.868	.903

Model Descriptions (also see Table 1 & Figure 1)

Model 0 = Null model (resulting in a diagonal reproduced matrix)

Model 1A = First-order factors only (see Table 1)

Model 1B = Same as 1A except that all factor are uncorrelated.

Model 2A = 1 higher-order factor defined by all 13 first-order factors

Model 3A = 2 higher-order correlated factors

Model 4A = 3 second-order factors & 1 third-order factors

Model 4B = same as 4A except that there is no third-order factor, and the 4 lower order factors that define it are correlated.

Model 5A = 4 second-order factors & 1 higher-order factor

Model 5B = same as 5B except that there is no third-order factor, and the 8 lower-order factors that define it are correlated.

Model 6A = 5 second-order factors & 1 third-order factor;

Model 6B = same as 6A except that there is no third-order factor, and the 7 lower-order factors that define it are correlated.

Model 7A = same as 6A except that the physical and social second-order factors are combined to form a single second-order factor, and the Problem Solving factor contributes directly to the general hierarchical self-concept.

a -- the goodness-of-fit indicators are defined in Table 3

Note: The first-order factor structure (see Table 1) was the same for all models. The structure of higher-order factors in Models 2A, 3A, 4A, 5A, 6A and 7A are shown in Figure 1. Models 4A - 7A all had two equality constraints used to define the second-order academic factors (see Figure 1). Models 5A-6B had one equality constraint used to define the second-order physical factor. Models 6A, 6B and 7A had one equality constraint used to define the second-order moral factor. When the 12 models described above were tested with the data from time 2 the observed chi-square values were: 11620; 1251; 2726; 1698; 1536; 1458; 1458; 1408; 1459; 1461; 1426; 1438. See Note in Table 3 for the abbreviations of the goodness-of-fit indicators.

Table 5

Percentage of Variance in First-order and Second-order Factors Explained by the Higher-order factor Models Shown in Figure 1

First-Order Factors	Variance estimate	% of Variance Explained By Model 2A	3A	4A	5A	6A
Math	.856	14.7	41.6	90.7	90.7	90.7
Verb	.731	16.6	69.8	86.3	86.7	86.3
Acad	.816	22.8	65.7	64.3	64.3	64.3
Prob	.517	26.5	65.2	55.5	55.5	55.5
Pabl	.685	37.2	38.7	38.7	44.5	44.5
Appr	.708	36.4	36.4	36.4	44.4	44.4
Ssex	.778	46.0	49.9	49.9	56.3	56.3
Osex	.841	38.2	37.0	38.2	42.9	42.9
Prnt	.795	30.8	33.3	32.1	35.8	34.6
Relg	.755	0.6	0.6	0.6	0.6	12.6
Hnst	.251	4.4	4.4	4.4	4.4	56.2
Emot	.805	54.0	57.8	57.8	59.0	59.0
Estm	.835	68.9	75.9	77.2	78.4	78.4
Second-order Factors						
Math/Acd	.769	--	--	10.3	10.3	10.3
Read/Acd	.629	--	--	20.5	18.9	18.9
Physical	.306	--	--	--	80.4	80.4
Social	.471	--	--	--	80.9	83.0
Moral	.120	--	--	--	--	16.7

Note: The percentage of variance explained is defined as: [(1 - (Residual Variance/Variance Estimate)) x 100%]. Variance estimates for first-order factors were obtained from Model 1A (see Table 1) while those for the second-order factors came from Model 6A (see Figure 1). Residual variance estimates for each model appear in Figure 1. A "--" indicates that the second-order factor was not estimated in a particular model. See Note in Table 1 for the abbreviations of the factor names.

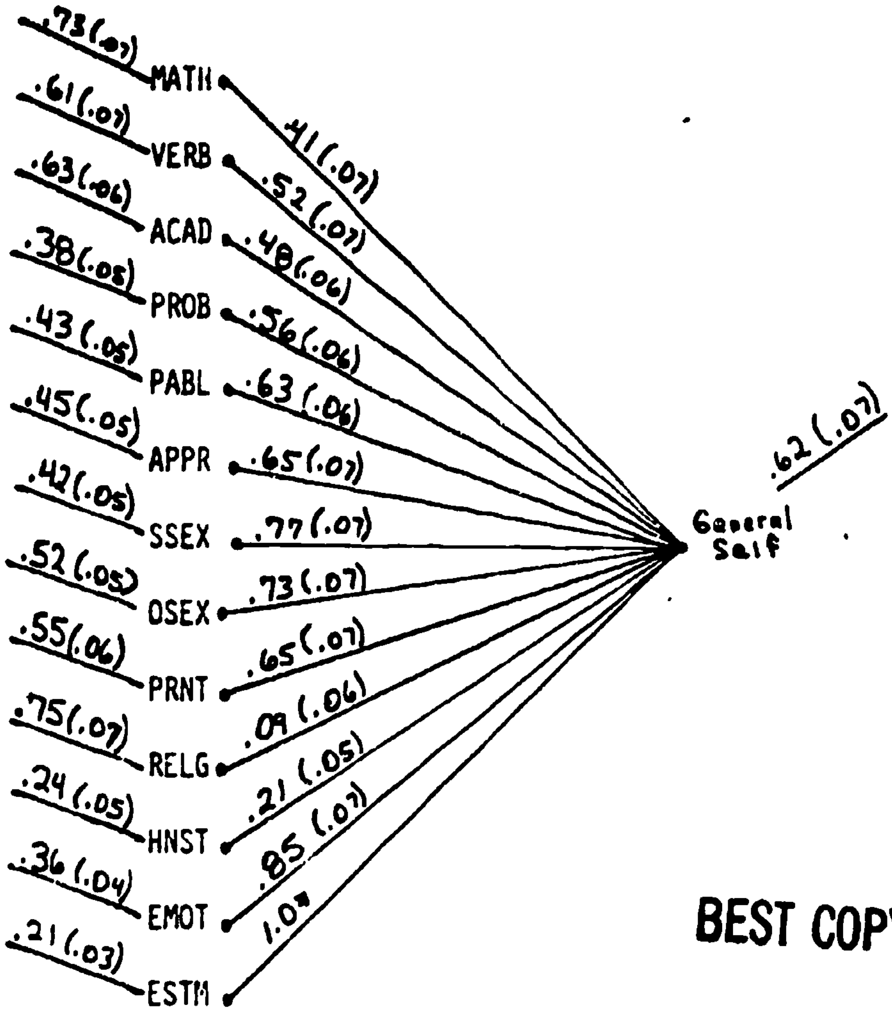
BEST COPY AVAILABLE

FIGURE 1 -- Seven Hierarchical Models and Parameter Estimates

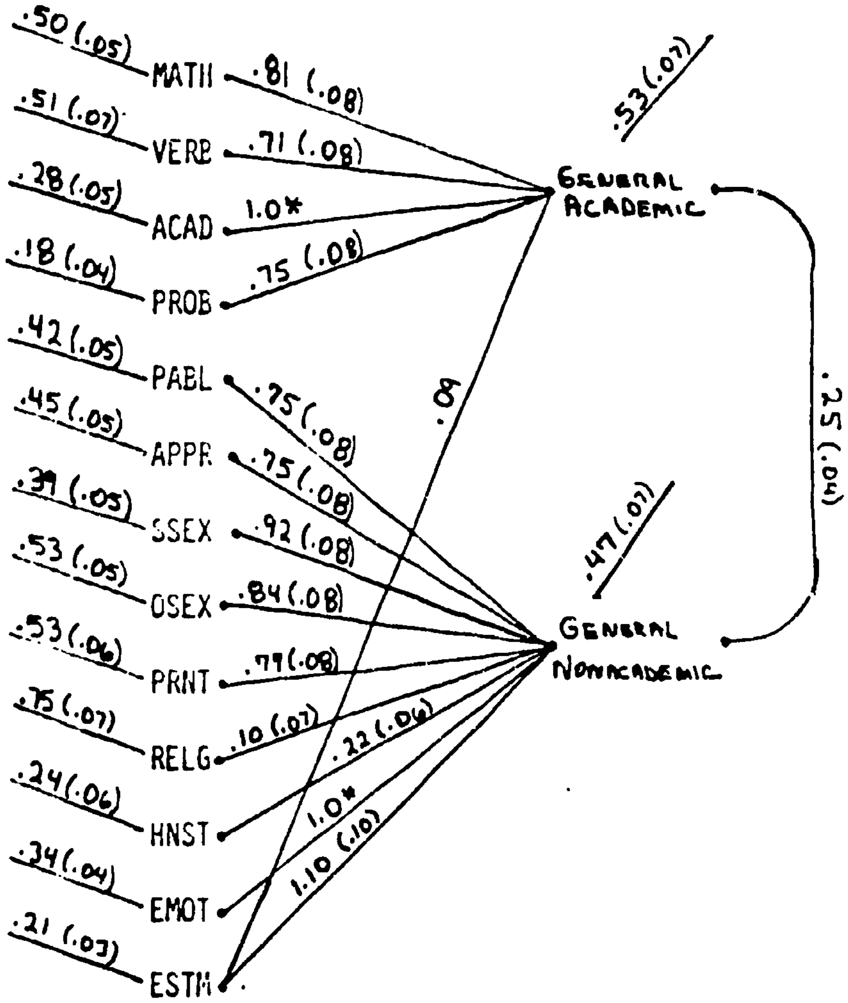
*Indicates loadings fixed at 1.0

a. The Academic Factor was constrained to load equally on the Math/Academic and Verbal/Academic Factors

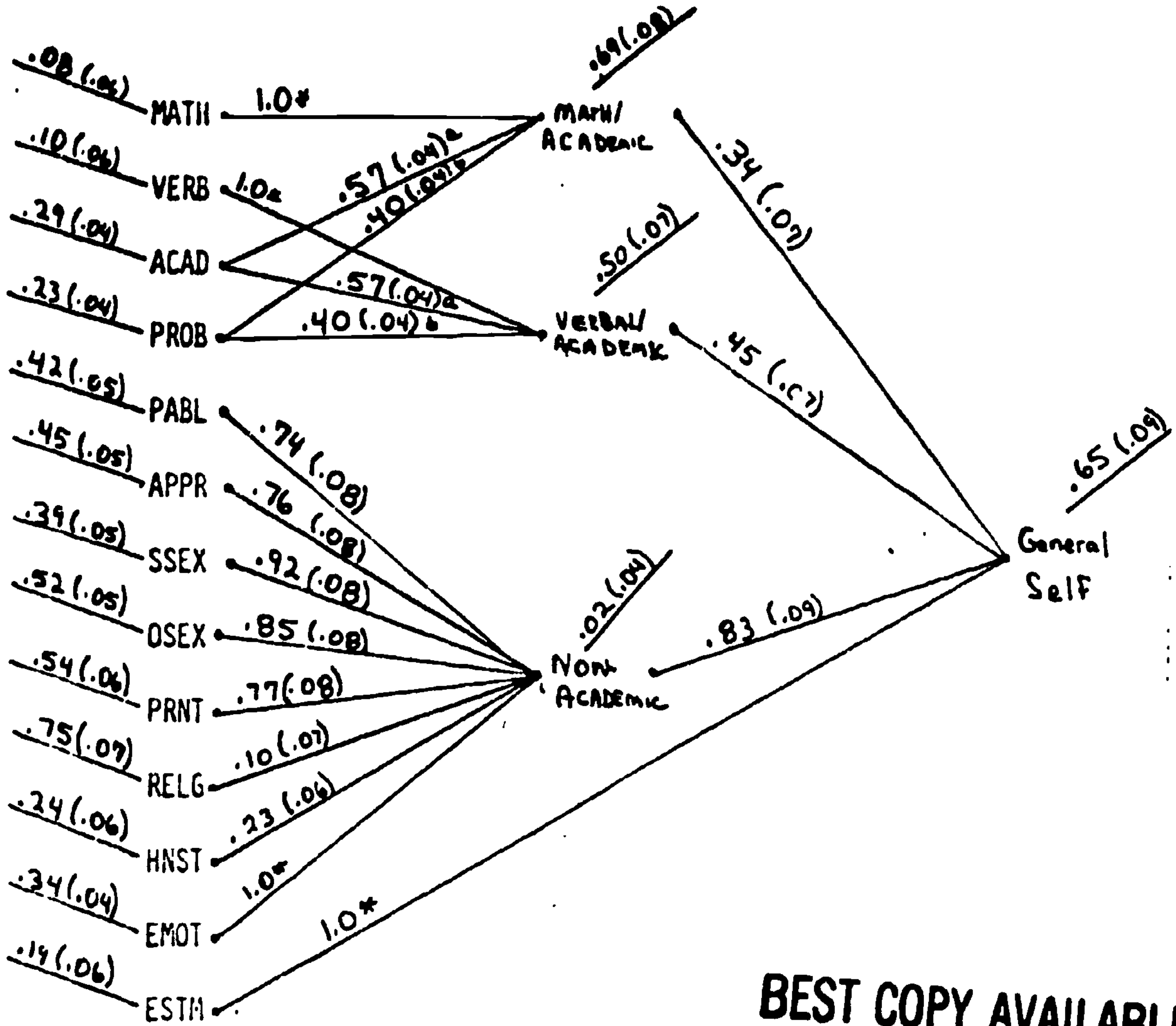
b. The Problem Solving factor was constrained to load equally on the Math/Academic and Verbal/Academic Factors



Model 2A $X^2(689) = 1785$

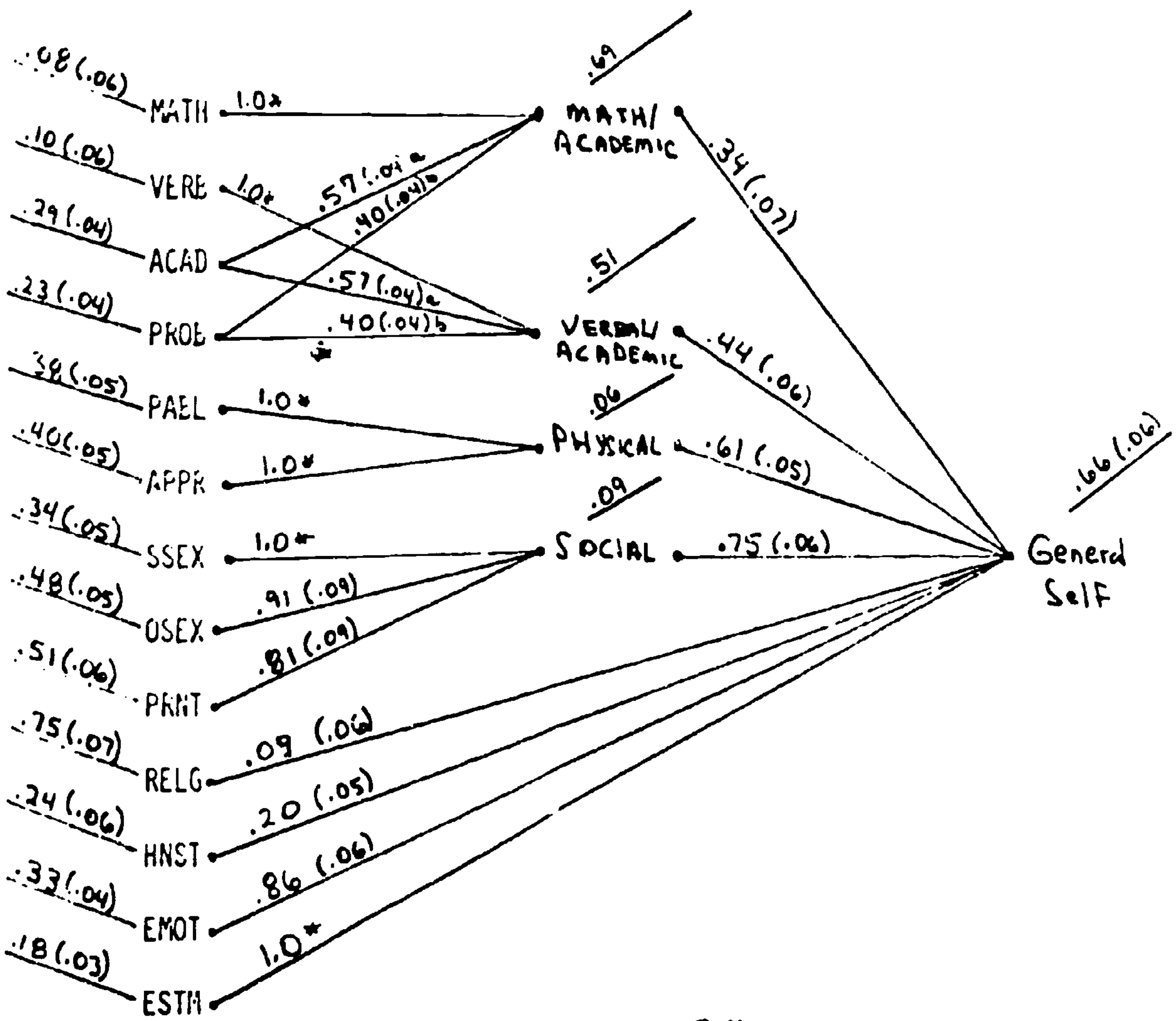


Model 3A $X^2(687) = 1613$

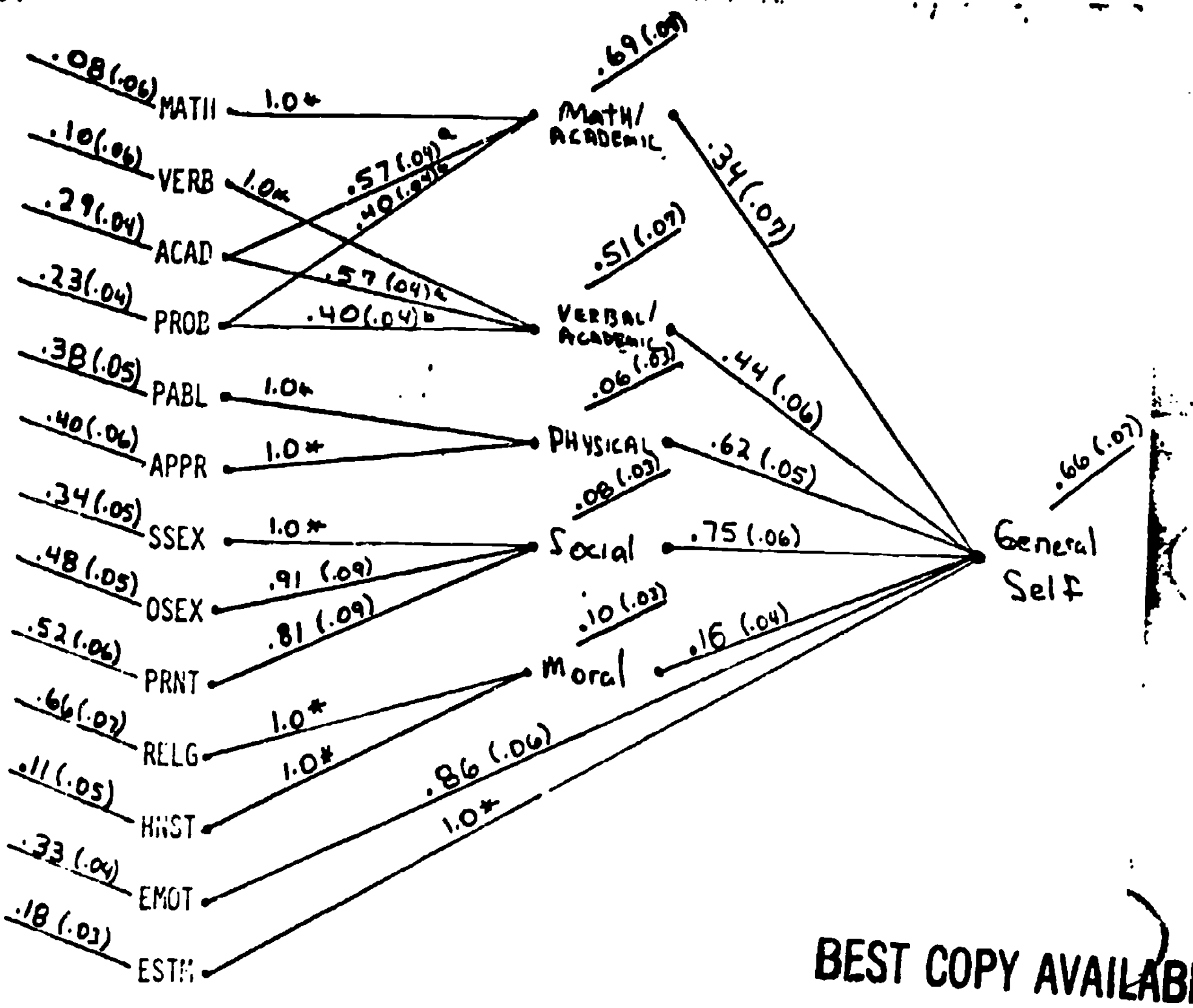


BEST COPY AVAILABLE

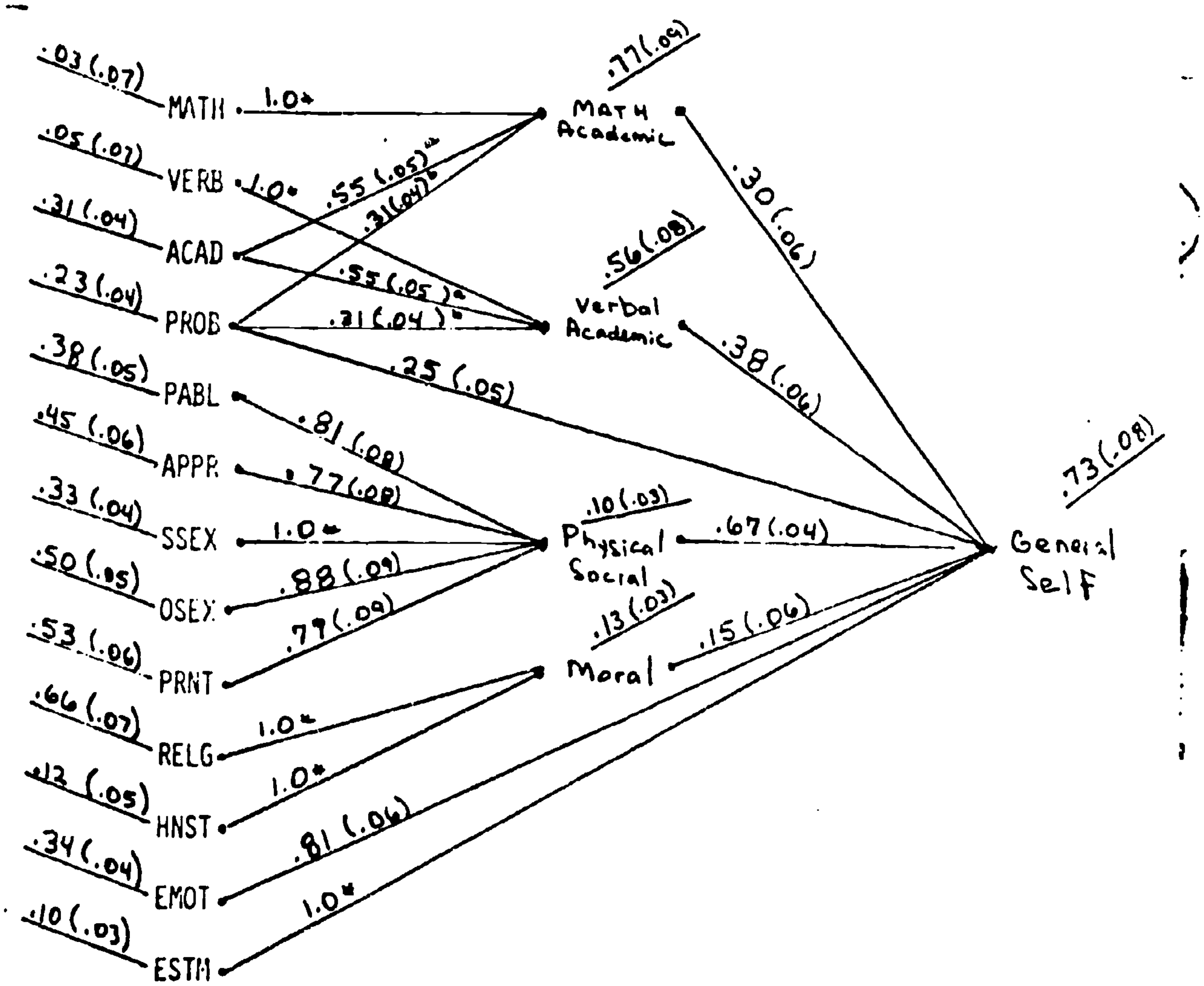
MODEL 4A $\chi^2(684) = 1539$



MODEL 5A $\chi^2(686) = 1527$



MODEL 6A $\chi^2(686) = 1515$



MODEL 7A $\chi^2(685) = 1476$

mates