DOCUMENT RESUME

ED 319 787 TM 015 040

AUTHOR Marsh, Herbert W.; Byrne, Barbara M.
TITLE Do We See Ourselves as Others Infer: A Comparison of Self-Other Agreement on Multiple Dimensions of Self-Concept from Two Continents.
PUB DATE 21 Feb 90
NOTE 35p.; Filled print throughout; document will not reproduce well.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Comparative Testing; Cross Cultural Studies; Factor Analysis; Foreign Countries; Higher Education; Multidimensional Scaling; Multitrait Multimethod Techniques; *Peer Evaluation; *Psychometrics; *Self Concept; Self Evaluation (Individuals); *Significant Others; *Undergraduate Students
IDENTIFIERS Australia; Canada; Confirmatory Factor Analysis; *Self Description Questionnaire III

ABSTRACT

Self/other agreement between self-concept ratings by the individual and self-concepts inferred by significant others is of theoretical and practical importance, but the review by J. S. Shrauger and T. J. Schoeneman (1979) found no evidence for such agreement. In the present investigation, the Self Description Questionnaire III (SDQIII) was completed by 151 Australian and 941 Canadian university students and by significant others selected by students as the person in the world who knew them best. Self/other agreement on each of the 13 SDQIII scales was very high for Australians and Canadians. Guidelines for evaluating multitrait-multimethod (MTMM) data and a confirmatory factor analysis (CFA) approach strongly support convergent and discriminant validity. An important new extension of the MTMM CFA approach to multistudy data showed that all parameter estimates were reasonably invariant (i.e., equal) for Australians and Canadians. This self/other agreement is much higher than found elsewhere and its replication across two continents dramatically refutes the Shrauger and Schoeneman conclusion. Critical features leading to this high self/other agreement appear to be the use of older subjects, multiple dimensions of self-concept based on instruments with strong psychometric properties, and significant others who know the subject very well. A 47-item list of references is included. Seven data tables and a list of means, standard deviations, and correlations among 26 study variables are provided. (TJH)

Do We See Ourselves As Others Infer: A Comparison of Self-other Agreement On Multiple Dimensions of Self-Concept From Two Continents

Herbert W. Marsh
University of Sydney, Australia

Barbara M. Byrne
University of Ottawa, Canada

21 February, 1990

Running Head: Self-other Agreement

ABSTRACT

Self-other agreement between self-concept ratings and self-concepts inferred by significant others is of theoretical and practical importance, but the Shrauger and Schoeneman (1979) review found no evidence for such agreement. In the present investigation the Self Description Questionnaire III (SDQIII) was completed by Australian (n=151) and Canadian (n=941) university students, and by the significant others selected by students as the person in the world who knew them best. Self-other agreement on each of the 13 SDQIII scales was very high for Australians (mean rs = .568) and Canadians (mean r =.560). Guidelines for evaluating multitrait-multimethod (MTMM) data and a confirmatory factor analysis (CFA) approach both provided strong support for convergent and discriminant validity. An important new extension of the MTMM CFA approach to multistudy data showed that all parameter estimates were reasonably invariant (i.e., equal) for Australians and Canadians. This self-other agreement is much higher than found elsewhere and its replication across two continents dramatically refutes the Shrauger and Schoeneman (1979) conclusion. Critical features leading to this high self-other agreement appear to be the use of: older subjects; multiple dimensions of self-concept based on instruments with strong psychometric properties; and significant others who know the subject very well.

The study of agreement between self-concept ratings and self-concept inferred by others has a long and controversial history (e.g., Baldwin, 1897; Bem & Allen, 1974; Burns, 1979; Cooley, 1902; Duval & Wicklund, 1972; James, 1890; Kinch, 1963; Mead, 1934; Marsh, Barnes & Hocevar, 1985; Shrauger & Schoeneman, 1979; Wells & Marwell, 1976; Wylie, 1974; 1979). Symbolic interactionists argue that: (a) self-concept emerges from social interaction with others, (b) self-concept is based on the ways others respond to the person, and (c) perceptions of how one is perceived by others reflect, in part, actual perceptions by others (Kinch, 1963; Marsh, Barnes & Hocevar, 1985; Shrauger & Schoeneman, 1979). Coombs, Soper, and Courson (1963) argued that ratings by external observers provide a better indicator of self-concept than self-report measures because they are not affected by self-report biases. In contrast, most researchers (e.g., Crandall, 1973; Marsh, Barnes & Hocevar, 1985; Marsh, Smith, Barnes & Butler, 1983; Shavelson, Hubner & Stanton, 1976) argue for the theoretical separation of self-concept that is based on self-report from inferred self-concept that is based on the report of others. Shavelson et al., posited that self-concept is influenced by the evaluations of significant others and that self-other agreement should be stronger for dimensions of self-concept near the base of their hierarchical model that are more closely associated with actual behavior. Shavelson et al., also argued, however, that self-concept measured by self-report is a separate construct from self-concept inferred by external observers and that there need not be any strong relation between the two constructs. Crandall suggested that ratings by others are useful in validating self-report measures, though Marsh, Barnes and Hocevar argued that self-other agreement will be substantial only if the external observer knows the subject well, observes a wide range of behavior, and makes judgments of the same specific characteristics as the subject.

Shrauger and Schoeneman (1979) reviewed studies that correlated self-reports with judgments by others, and concluded that "there is no consistent agreement between people's self-perceptions and how they are actually viewed by others" (p. 549). This far reaching conclusion calls into question the construct validity of self-responses in general and self-concept responses in particular, but methodological characteristics of their study require further consideration.

1) Insufficient detail was given to characteristics of different studies that might influence self-other agreement; the content of the self-reports was quite varied, little concern was given to the psychometric properties of the responses, and no attempt was made to determine if some external observers provided more accurate assessments than others.

2) They did not consider the distinctiveness of different components when multiple characteristics were judged. In the terminology of multitrait-multimethod (MTMM) analyses they consider only convergent validity but not discriminant validity.

3) They did not distinguish studies that asked external observers to record their own perceptions from those where observers inferred the subjects' self-concept ratings. That is, external observers can be asked what they "objectively" think or feel about a person, or to use their observations to infer what that person thinks about him/herself (i.e., inferred self-concept). The first approach might be appropriate to determine how accurately a person views him/herself compared to the perceptions of others. However, self-concept is based upon self-perceptions, whether accurate or not, and so the second approach is used in most self-concept research (see Wells & Marwell, 1976, p. 136-142, for further discussion).

4) They failed to distinguish between self-other agreement based on correlations and agreement inferred from mean level comparisons. Self-other correlations were statistically significant in more than 50% of the correlation studies whereas there were significant differences in mean ratings by self and other for a majority of these studies.
In summary, there is need for further research on self-other agreement in better controlled studies.

Self-other discrepancies are sometimes used to test the frequently hypothesized self-favorability bias -- that self-ratings are systematically higher than they "should" be. In such studies the focus is on mean differences between self-responses and responses by others rather than correlations between responses. Such a self-favorability effect on self-concept may represent the influence of selective perception, memory, and interpretations, the perhaps unrealistic feedback often given particularly to young children, and frame of reference effects (i.e., different standards of comparison), or, alternatively, intentional distortions in self-reports that do not accurately reflect true self-perceptions. In a review of this research, Wylie (1979, p. 681) concluded that: "there appears to be a considerable consistency among the methodologically more adequate studies in showing trends toward self-favorability biases regarding evaluative characteristics." Wylie, however, specifically excluded studies based on "private-self-concept responses" (i.e., the subject is instructed to report how he or she privately sees him or herself, whether or not this agrees with external criteria) and only considered studies in which subjects made "social-self-concept responses" (i.e., regardless of his or her own private

view of him or herself, the subject is to tell how he or she thinks generalized or particular others would characterize him or her) or made self-ratings relative to some objective standard such as school grades. The social-self-concept that she used to evaluate self-favorability biases is not the self-concept considered here, is not the self-concept typically considered in other self-concept research, and is not the self-concept considered by Wylie in other sections of her 1974 and 1979 books. Hence there is need for further research on self-other discrepancies for responses in the form that they are typically used in self-concept research.

Self-other agreement on multidimensional self-concept ratings

Historically, self-concept research has emphasized a general, global or total self-concept, but more recently there has been growing support for the multidimensionality of self-concept. There is particularly good support for the Shavelson model of self-concept (Shavelson, et al., 1976; Marsh, Byrne & Shavelson, 1988; Marsh & Shavelson, 1985) and the Self Description Questionnaire (SDQ) instruments based on this model. According to the Shavelson model, self-concept is a multifaceted, hierarchical construct; general self-concept at the apex of the hierarchy is divided into academic and nonacademic self-concepts which are further divided into more content specific components of self. In her review of self-concept models Byrne (1984, p. 449) concluded that: "Although no one model to date has been sufficiently supported empirically so as to lay sole claim to the within-network structure of the construct, many recent studies, in particular those by Marsh and his colleagues, are providing increasingly stronger support for the hierarchical model." Subsequent reviews of research prompted by this model (Marsh & Shavelson, 1985; Marsh, in press-b) provide further support for the multifaceted structure of self-concept and demonstrate that self-concept cannot be adequately understood if this multidimensionality is ignored.

When multiple dimensions of self-concept are represented by both self-ratings and inferred-ratings, multitrait-multimethod (MTMM) analysis provides an important analytical tool for testing the construct validity of the self-concept facets (see Marsh, 1989; Marsh, Barnes & Hocevar, 1985; Shavelson, et al., 1976 for more general discussions of MTMM analyses). Convergent validity, the traditional focus of self-other agreement studies, is inferred from substantial correlations between self-ratings and inferred-ratings on matching self-concept traits. Discriminant validity provides a test of the distinctiveness of self-other agreement and of the multidimensionality of the self-concept facets; it is inferred from the lack of correlation between nonmatching traits. MTMM studies using the SDQI for preadolescents and the SDQIII for late-adolescents are briefly reviewed below.

Preadolescent Studies. Eight MTMM studies using the SDQI (see Marsh, 1988) demonstrated significant agreement between multiple self-concepts inferred by primary school teachers and student responses to the SDQI. Across all 8 studies the average of the 56 convergent validities (self-other agreement on matching scales) was 0.30. Student-teacher agreement was higher in those areas in which relevant behaviors were most observable (math, .37; reading, .37; general school, .33; physical ability, .38; and, peer relations .29), but was lower on Relations With Parents (.17) and somewhat surprisingly Physical Appearance (.16). These studies demonstrate that external observers can infer self-concepts in many areas with at least modest accuracy, thus countering Shrauger and Schoeneman's 1979 claim to the contrary. Whereas support for convergent and discriminant was evident when evaluated by the traditional Campbell-Fiske criteria, the level of self-other agreement was only modest. There are several likely explanations for why self-other agreement is only modest: (a) preadolescents may be more likely than older subjects to base their self-concepts on idiosyncratic criteria that are unobservable or not considered by external observers; (b) teachers may not have an appropriate basis for inferring self-concepts in some areas; and (c) because teachers made ratings of all students in their class, they were only asked to respond to psychometrically weaker single-item scales instead of the multi-item scales completed by students. These suggestions were examined in part in subsequent research based on late-adolescent responses to the SDQIII.

Late-adolescent/young adult studies. A particularly important MTMM study (Marsh, Barnes & Hocevar, 1985) was conducted with SDQIII responses by a small sample 151 of Australian university students. Students were asked to complete the SDQIII and to ask the person in the world who knew them best to complete the SDQIII as if they were the person who had given them the survey. The significant others were typically family members -- most frequently a parent. Separate factor analyses of both self-ratings and responses by significant others identified the 13 dimensions of self-concept which the SDQIII was designed to measure. Self-other agreement was very high (mean r = 0.57), and four of the scales had self-other correlations over .75 (Physical ability, Religion/spiritual values, Parent relations, and Mathematics). An application of the traditional Campbell and Fiske (1959) guidelines for evaluating MTMM matrices provided strong support for the convergent and the discriminant validity of the SDQIII responses. The self-other agreement reported in this study was much higher than previously reported and dramatically refuted the Shrauger and Schoeneman (1979) claim

that there is no systematic self-other agreement. Particularly on the four scales with the highest self-other agreement the results may be the strongest relation between self-reports and an external validity criterion found in personality research.

Marsh, Barnes and Hocevar (1985) also compared the mean responses by subjects and by significant others; there was little systematic difference, or a slight tendency for higher responses by significant others. Although not emphasized in the study, these findings may call into question Wylie's 1979 claim for self-favorability biases. As part of the same study, subjects and significant others also responded to single-item summary ratings like those used in many other studies. Whereas there was still support for the convergent and discriminant validity of these single-item responses, self-other agreement was substantially lower than found on the multi-item scales.

Marsh and Richards (in press) further examined self-other agreement on SDQIII responses for 280 participants in an Outward Bound program. Participants in this 26-day residential program worked primarily in small groups and observed the other members of their group in many different situations. Each subject chose two group members who knew him or her the best, and these external observers were asked to complete single-item summary ratings as the subject "would" complete them and as the subject "should" complete them. MTMM analyses of agreement between the two external observers indicated modest agreement (mean $r = .32$) and support for convergent and discriminant validity of their responses. Similarly, agreement between responses by the two external observers and self-responses by the subject (mean $r = .37$) was moderate. Although the results provided support for convergent and divergent validity, correlations among ratings by external observers to different areas of self-concept were substantially higher than among self-response ratings, suggesting a method/halo effect in external observer responses. External observers were apparently unable to differentiate between "would" and "should" responses; "would" responses by different observers were no more highly correlated than were the "would" and "should" responses by different observers. Although self-other agreement in this study was higher than reported by Shrauger & Schoeneman (1979) it was substantially lower than reported by Marsh, Barnes and Hocevar (1985). The findings should, however, be evaluated in relation to two restrictions; external observers responded to single-item summary scales instead of multi-item scales and had contact, albeit intensive contact, with subjects for only 26 days. Particularly in relation to these limitations, support for the convergent and discriminant validity of the responses was surprisingly good.

Self-concept responses in the Marsh and Richards (in press) study were

consistently lower than those of external observers, suggesting a self-modesty effect instead of a self-favorability bias. This modesty effect was also consistent with the finding that observer "would" ratings were marginally lower than observer "should" ratings. That is, observers said that subjects "would" give themselves lower self-ratings than they "should". In evaluating this modesty effect it is important to note that self-responses and observer responses were actually made by the same individuals. That is, each participant judged him or herself (self-responses) and also made judgments of two other participants (observer responses). On average then, participants indicated that other participants "would" and "should" give themselves higher ratings than participants gave themselves -- again suggesting a modesty effect in the self-responses. Because the stimulus materials and individuals making the judgments were the same for the self-rating and observer tasks, many influences that might differentially affect self and other ratings (e.g., differences in response biases, the constructs being evaluated, frame of reference effects) are less plausible. In evaluating these results, Marsh and Richards suggested the possibility that modesty effects may be more likely in Australian studies than North American studies where self-favorability effects are typically found (Wylie, 1979).

Summary The MTMM studies summarized here clearly refute the Shrauger and Schoeneman (1979) claim that there is no evidence for self-other agreement and may call into question Wylie's (1979) suggestion of a self-favorability bias in self-concept research. Whereas self-other agreement was consistently significant across each of the different studies, this agreement was substantially higher in the Marsh, Barnes and Hocevar (1985) study. Not only was self-other agreement in this study substantially higher than previously found, correlations between self-responses and an external criterion (i.e., responses by significant others) may be among the highest reported in any area of personality research. The authors suggested that self-other agreement was so high because: (a) subjects were older, presumably knowing themselves better and having more realistic self-perceptions, than was the case in preadolescent studies; (b) ratings by significant others were based on responses by a person specifically selected by the subject as knowing them the best rather than responses by a convenient external observer (e.g., teachers or supervisors) used in most research; (c) significant others responded to the same multi-item response scales as completed by the subjects; (d) the traits being evaluated were well-defined and inferred on the basis of instruments with strong psychometric properties. There is indirect support for these suggestions in

that studies that do not have these characteristics consistently show much lower self-other agreement than the Marsh, Barnes and Hocevar study. Stronger support, however, would consist of showing that other research having these characteristics was able to replicate the substantial levels of self-other agreement found by Marsh, Barnes and Hocevar.

Self-reports and ratings in the assessment of personality.

Despite the historical importance of self-other agreement to self-concept research, recently there appears to be more systematic research on this issue in other areas of personality research (e.g., Cheek, 1982; Funder, 1987; Kenny & Albright, 1987; Kenrick & Funder, 1988). McCrae and Costa (1988; also see Conley, 1985; Costa & McCrae, 1988) reviewed theoretical implications and empirical findings in studies correlating personality traits derived from self-responses and from responses by others in research particularly relevant to the present investigation. As in self-concept studies considered here, their research examined self-other agreement on multiple traits to test the construct validity of self-reports and used factor analysis and MTMM analysis as the principal analytic tools. Their research was designed to counter previous claims that correlations between self-reports and ratings by others cannot break the so-called .3 barrier and the widely held beliefs that personality traits based on self-report reflect primarily self-presentation or response biases. Despite these similarities between the personality and the self-concept studies, there are important differences: (a) personality research, self-responses are designed to provide objective indicators of personality traits and thus are more like what Wylie (1979) referred to a public self-concept rather than the private self-concept that is the focus of most self-concept research; (b) observer ratings in the personality research are designed to provide an accurate appraisal of the subject whether or not they reflect the subject's self-perceptions; that is, they are more like the "should" ratings in the Marsh and Richards (in press) study than the inferred self-concept rating (i.e., "would" ratings) used in self-concept research; (c) even though personality studies often consider multiple traits, the major focus is convergent validity rather than divergent validity. Nevertheless, methodological advances in either area are likely to contribute to other area.

McCrae and Costa (1988) summarized 10 recent studies in which there was substantial agreement between self-reports and peer ratings. Whereas self-other agreement was consistently less than that reported by Marsh, Barnes and Hocevar (1985), some studies approached this level of agreement by aggregating responses from as many as 10 different observers. When external observers were spouses, however, self-other correlations on five personality

domains varied between .5 and .6, though correlations based on subscales within each domain were somewhat lower (Costa & McCrae, 1988). In the same study, correlations between self-ratings and peer ratings, and between spouse ratings and peer ratings were lower (median $r = 0.42$), suggesting that spouses knew the subjects better than peers. The McCrae and Costa study based on spouse ratings and the Marsh, Barnes and Hocevar study based on responses by persons selected by subjects as knowing them the best both suggest that knowledge of the subject is a critical variable. McCrae and Costa (1988) concluded their review of self-other agreement by noting that: "When reliable and valid measures are used, the correlations considerably exceed the .3 barrier; they are better characterized as facing a .6 barrier" (p. 5). However, whereas none of the correlations reported by Costa and McCrae was larger than .6, Marsh, et al. reported self-other correlations in excess of .75 for 4 of 13 self-concept scales.

Summarizing characteristics that seem to enhance self-other agreement, McCrae and Costa (1988) noted the use of: (a) multi-item scales instead of single-item ratings, (b) instruments with superior psychometric properties, (c) factor analytically derived factor scores that maximally distinguish between the multiple dimensions being considered, and (d) responses from observers who know the ratee better in a variety of different contexts. In a proposal similar to that of Shavelson et. al. (1976), they also suggested that agreement may be better on traits that are more closely linked to observable behavior. Hence, characteristics that enhance correlations between self-responses and external observer ratings of personality traits appear to be similar to those that lead to higher correlations between self-concept ratings and inferred self-concept ratings.

The Present Investigation

The present investigation has both substantive and methodological orientations. The substantive orientation is to demonstrate that when critical features of the Marsh, Barnes and Hocevar (1985) study are replicated, self-other agreement on multiple dimensions of self-concept will be higher than found in other research. This study has important implications for the study of self-other agreement in particular and for the construct validity of self-concept responses in general. For these reasons, the critical features of the Australian study were replicated in a large sample of Canadian university students and the results of the two studies are compared here.

The methodological orientation is to extend and refine the application of confirmatory factor analysis (CFA) to MTMM data. Whereas there have been numerous applications of this approach to results of a single MTMM study, an

important extension of this approach is to test the equality of results from two different studies using the same MTMM design. For multigroup comparisons, CFA provides tests of the equality of any one, any set, or all the parameter estimates from two or more groups. Although no previously published application of this analytic procedure to test the replicability of MTMM results in different studies is known to the authors, this approach provides much more powerful tests of the equality of results across two different studies than conventional procedures. Hence, an important contribution will be the demonstration of this multigroup CFA procedure for comparing results from the previously published Australian study (Marsh, Barnes & Hocevar, 1985) with those from a new study of Canadian university students.

Methods

Sample and Procedures.

Australian study. This study is described in detail by Marsh, Barnes and Hocevar (1985) and is summarized here only briefly. The sample consisted of 151 Australian university students from psychology and education classes who volunteered to participate in the study. Students completed the SDQIII and then asked the person in the world who knew them best to complete the companion survey using the same instructions as were subsequently used in the Canadian study. Because students in this study often lived at home, the significant others were typically family members and over half were parents.

Canadian study. Subjects were 941 introductory psychology students in a large Canadian university who volunteered to participate as partial fulfillment of a course requirement. As in the Australian study, students completed the SDQIII and then were asked to chose the person in the world who knew them best to also complete the SDQIII. On the companion survey, the significant others were asked to imagine that they were the person who had given them the survey and to complete the same SDQIII items as if they were that person. Students were explicitly instructed not to discuss the survey with their selected significant other. A pre-addressed envelope was included with the survey that was given to the significant other and they were explicitly instructed to return the survey without discussing their responses with the subject. Although the relation between students and the significant other was not obtained, informal queries suggested that slightly more than half of the significant others were intimate partners (spouse, live-in partners, or boy friend/girl friend) and that most of the remaining significant others were an immediate family member — most frequently a parent. A total of 1081 students completed the SDQIII and 941 pairs of instruments completed by the subject and by the significant other were obtained. Analyses described here are based on the 941 pairs of response where

both self-responses and responses by the significant other were completed.

Psychometric properties of the SDQIII.

Each of the 13 SDQIII scales (Marsh, in press-c; Marsh, Barnes & Hocevar, 1985; Marsh & O'Neill, 1984; Marsh, Richards & Barnes, 1986a; 1986b) is represented by 10 or 12 items, half of which are negatively worded (see Marsh & O'Neill, 1984, for the wording of the items). In the Canadian study the coefficient alpha estimates of reliability for each scale were high for both self-responses (median = .89) and responses by significant others (median = .89), and similar to those found in the Australian study (medians of .88 and .90 respectively). Separate exploratory factor analyses were conducted for self-responses and responses by significant others in the Canadian study and the results were compared to the corresponding two factor analyses in the Australian study (Table 1). In all four factor analyses (see Marsh, Barnes & Hocevar, 1985 for a more detailed description of the analysis), factor loadings for each variable designed to measure a factor (target loadings) were high (medians of .67 to .72), all other factor loadings (nontarget loadings) were consistently low (medians of .02 in all four analyses), and correlations among the factors were low (medians of .07 to .09). Results of these four factor analyses are also very similar to one conducted on the entire sample of 2436 subjects in the normative sample of SDQIII responses described in the test manual (Marsh, in press-c). These psychometric properties of SDQIII responses -- internal consistency and factor structure -- are not emphasized here because they have been previously identified in many different studies (Marsh, in press-c). These properties are, however, an important pre-requisite to subsequent comparisons between self-responses and responses by others that is frequently lacking in other self-concept research.

Insert Table 1 About Here

For purposes of the present investigation, self-responses and responses by others are each represented by 13 factor scores. The use of factor scores instead of simple summated scale scores was recommended by Marsh, Barnes and Hocevar (1985) and also be McCrae and Costa (1988) in their review of related research. The factor scores were derived from the factor analysis of results based on the normative archive of responses by 2436 responses summarized by Marsh (in press-c). Factor scores are a weighted average of responses to the standardized responses to each measured variable in which the weights correspond to factor score coefficients. Measured variables were standardized in relation to results from the normative archive and factor score coefficients were derived from the factor analysis of results from the

normative archive. This is one of the procedures recommended in the test manual (Marsh, in press-c) where all the relevant information needed to compute these scores is presented. Particularly for purposes of the present investigation, this procedure has the important advantage of providing factor scores that are comparable for self-responses and responses by others, and that are comparable in the Canadian and Australian samples. The use of normative results to construct factor scores is justifiable and generally preferable to factor scores based on results from one idiosyncratic sample (Marsh, in press-c). It is interesting to note, however, that the 13x4=72 factor scores representing the 13 SDQIII scales from the four separate factor analyses (i.e., self-responses and responses by others in the two studies) all correlated at least .97 (median .99) with the corresponding 72 factor scores based on the normative results that are used here. This very high agreement between factor scores derived from different factor analyses provides further support of the replicability of the SDQIII factor structure across the different analyses. (Because results previously reported by Marsh, Barnes and Hocevar were based on a factor analysis idiosyncratic to that particular study, some results presented here vary slightly from those previously presented.)

Results: Self-other Agreement on Multiple Dimensions of Self-concept. Traditional approaches to MTMM data.

The major focus of this investigation is on self-other agreement based on the Canadian students and the comparison of these results with those based on the Australian study. For each study, a 26x26 correlation matrix (see appendix) represents correlations between the 13 SDQIII scales based on self-responses and the 13 SDQIII scales based on responses by significant others. Of particular relevance are the 13 correlations between self-responses and responses by others on matching SDQIII scales (i.e., the convergent validities, see Table 2). In the Canadian study these vary from 0.400 to 0.769 (mean = .560) and are very similar to those from the Australian study that vary from .311 to .800 (mean = .568). By visual inspection, the pattern of convergent validities for the different SDQIII scales appears to be similar in the two studies. In order to provide a more objective index of this observation, the two sets of convergent validities were correlated with each other. The large correlation (.91, $p < .001$, df = 11) indicated that the relative size of correlations for different scales was very similar in the two studies. The four SDQIII scales on which self-other agreement was exceptionally high in the Australian study (rs of .74 to 0.80 in Table 2) also had the highest self-other agreement in the Canadian study, although self-other agreement on Parent Relations was smaller in the

Canadian study. In summary, self-other agreement summarized here is remarkably high — much higher than found in other self-concept research. Furthermore, the results for the two different studies are remarkably similar, particularly given the small sample size (and consequently large sampling variability) of the Australian study.

Insert Table 2 About Here

MTMM matrices (see appendix) have traditionally been evaluated with the Campbell and Fiske (1959; see Marsh, 1989) guidelines. Despite important limitations, these guidelines are still useful and recommended as a preliminary analysis even when more sophisticated approaches using CFA are subsequently used (see Marsh, 1989; Marsh, in press-a; Marsh, Barnes & Hocevar, 1985; Marsh & Hocevar, 1983). Applying these guidelines to the MTMM matrix of correlations between self responses and responses by others in the Canadian and Australian studies respectively indicated that:

1) All 13 convergent validities in each study were statistically significant (mean rs of .560 [Canadian] and .568 [Australian]) and substantial.

2) For every one of the 312 possible comparisons between a convergent validity and another correlation in the same row or column of the square (heterotrait-heteromethod) block of coefficients, the validity coefficient (means of .560 and .568) was higher than the comparison coefficient (means of

comparisons).

3) For the 312 possible comparisons between a convergent validity coefficient and other correlations in the same row or column of the two triangular (heterotrait-monomethod) blocks, the validity coefficient (means of

[Canadian] and .108 [Australian]; for 311 comparisons in each studies.

4) For both studies, correlations among ratings by others (means of .112 [Canadian] and .130 [Australian]) and correlations among self-ratings (means of .101 [Canadian] and 0.087 [Australian]) were slightly higher than heterotrait-heteromethod correlations (means of .047 and .039), suggesting a small method/halo effect.

These MTMM findings provide strikingly strong support for both the convergent and divergent validity of responses by the Canadian students to the SDQIII, and the similarity of results in the Canadian and Australian studies. Results described in the next section provide a stronger basis of comparing results from the two studies.

The CFA approach to MTMM data.

Introduction. MTMM matrices can be factor analyzed to infer the

underlying dimensions. Factors defined by different measures of the same trait suggest trait effects, whereas factors defined by measures with the same method suggest method/halo effects. With CFA the researcher can define alternative models that posit a priori trait and/or method effects and test the ability of such models to fit the data. Although problems exist in assumptions underlying the model, technical difficulties in estimating parameter estimates, and assessing goodness of fit, this approach to MTMM analysis has been frequently applied and is becoming increasingly popular (see Marsh, in press-a; 1989; Marsh & Hocevar, 1983; Widaman, 1985). For present purposes two general models are examined: one positing only 13 trait factors corresponding to the 13 SDQIII scales and one positing 13 trait factors and two method factors corresponding to the two methods of measurement (self-ratings and ratings by others). In the model with trait and method factors; each measured variable is related to just the one trait factor and the one method factor that it represents; error/uniquenesses for the measured variables are assumed to be uncorrelated; the 13 trait factors are assumed to be correlated whereas the correlations between trait and method factors and between the method factors are assumed to be zero; factor variances are fixed at 1.0 (see Marsh, in press-a; 1989 for further discussion of defining the CFA models used in MTMM studies); and the CFA models were fit with the LISREL V program (Joreskog & Sorbom, 1981).

Goodness of fit is evaluated by: (a) determining that the iterative procedure converges to a proper solution that is well-defined; (b) establishing that parameter estimates are substantively reasonable in relation to the a priori model; and (c) evaluating the X^2 test statistic and various fit indices (e.g., the Tucker-Lewis index (TLI) and the Bentler-Bonett index (BBI); see Marsh, Balla & McDonald, 1988) in relation to rules of thumb and comparing these values for alternative models. In an evaluation of goodness-of-fit indices typically used in CFA, Marsh, Balla and McDonald (1988) and Marsh, McDonald and Balla (1988) determined that the TLI was the only widely used index that was relatively independent of sample size and relatively unaffected by capitalization by chance when additional parameters were added. Because of the large sample size and wide variation in the number of parameters in alternative models in the present investigation, the TLI will be emphasized.

The CFA approach to MTMM data is well known and is not reviewed here (see Marsh, in press-a; 1989; Widaman, 1985). An important extension of this approach, however, is to use the CFA approach to test the equality of parameter estimates (trait and method factor loadings, factor covariances, and uniquenesses) derived from different studies. The major emphasis here is

on these between-study equality constraints in which the equality of parameter estimates from the Canadian and Australian studies is tested. It is also relevant, however, to consider within-study equality constraints in which the equality of parameter estimates for self-ratings and ratings by significant others is tested. When either between-study or within-study equality constraints are imposed, it is important that analyses are conducted on covariance matrices instead of the correlation matrices typically used in MTMM studies and that all measured variables are assessed according to a common metric (Joreskog & Sorbom, 1988) as in the present investigation.

A very large number of different models that impose various between-study and within-study equality constraints are possible and so only a small proportion of these models would actually be tested in any particular study. Whereas the particular set of models to be tested will depend on the substantive issues and comparisons of preliminary models, some general guidelines are appropriate. In the present investigation 13-trait models and 13-trait/2-method models were considered. In preliminary evaluations of these models, alternative models were considered in which: (a) no equality constraints were imposed, (b) all parameters estimates for self-ratings were constrained to be equal to the corresponding parameter estimates for responses by significant others (i.e., total within-study invariance), (c) all parameter estimates in one study were constrained to be equal the matching parameter estimates in the other study (i.e., total between-study invariance), and (d) there was total between-study and within-study invariance.

Comparison of fit indices in alternative models. Whereas the four 13-trait models (models 1a-1d in Table 3) provide a reasonable fit to the data (TLIs of .793 to .830), the corresponding 13-trait/2-method models (models 2a-2d) fit the data substantially better (TLIs of .933 to .958). This suggests that the measured variables reflect primarily the effects of traits but also are influenced to a small extent by method effects. Consistent with this interpretation, trait factors account for 8 to 10 times as much variance as the method factors for both self-responses and responses by significant others (see variance components in Table 4). These conclusions are also consistent with earlier evaluations of the MTMM matrices. Based on these preliminary findings, the focus of subsequent discussion will be on the evaluation of the 13-trait/2-method models.

Insert Table 3 About Here

A comparison of the first four 13-trait/2-method models (models 2a-2d) provides a clear pattern of results. All four models provide reasonable fits to the data according to traditional guidelines (e.g., all TLIs are greater

than .90 and X^2/df ratios are less than 2.0). The imposition of between-study equality constraints appears to have little effect on goodness of fit, independent of whether within-study equality constraints are imposed. In fact, based on the TLI and the X^2/df ratios, models imposing total between-study invariance (models 2c and 2d) fit the data slightly better than the corresponding models imposing no between-study invariance constraints (models 2a and 2b). On the other hand, within-study equality constraints result in a poorer goodness of fit independent of whether between-study equality constraints are imposed. Based on the focus of this investigation on between-study constraints and the results of these preliminary models, a more detailed evaluation of models imposing between-study constraints but no within-study constraints is considered.

Insert Tables 4 and 5 About Here

Solutions for the Canadian and Australian studies each consist of 156 parameter estimates (factor loadings, factor covariances, uniquenesses; see Tables 4 and 5). The global test of invariance in which all 156 parameter estimates were constrained to be equal in the two studies (model 2c) provided a reasonable fit and one that did not differ substantially from the corresponding model in which no equality constraints were imposed (model 2a). Because such a large number of parameter estimates were constrained to be equal, however, it is possible that invariance constraints for a few parameters made a substantial difference but that this difference was lost when averaged across all 156 constraints. Marsh and Hocevar (1985) describe a systematic approach to testing invariance in confirmatory factor analysis. Based on this approach, alternative models were formulated in which various combinations of trait-factor loadings, method-factor loadings, factor covariances, and uniqueness were constrained to be equal (models 3a-3d in Table 2). Fit indices for these alternative models, however, are all very similar, and the TLIs and X^2/df ratios for none of these models is any better than model 2c that imposes total between-study invariance. Alternatively, it is possible that between-study invariance constraints have substantially different effects on parameter estimates associated with self-responses and responses by others (models 4a and 4b). The fit indices associated with these two models, however, are similar and the TLIs and X^2/df ratios are slightly poorer than in model 2c that imposes total between-study invariance. In summary, these additional models provide strong support for the total invariance of results in the Canadian and Australian studies.

Preliminary analyses indicated that the imposition of within-study invariance constraints resulted in substantially poorer fits. That is, requiring parameter estimates based on self-responses to be the same as

those based on responses by others was not supported. There are, however, three logically distinct sets of within-study invariance constraints: trait-factor loadings, method-factor loadings, and uniquenesses. Various combinations of these constraints were tested in models 5a-5c. All three models, however, resulted in poorer fit indices than did the corresponding model (2c) in which no within-study equality constraints were imposed. It should be noted, however, that even the model imposing total within-study and total between-study invariance (model 2d) provides a reasonable fit (TLI =.940 and chi-square/df ratio of 1.874) when evaluated by traditional guidelines.

Parameter estimates for the selected model. Results summarized in the last section suggested that model 2c (total between-study invariance and no within-study invariance) provided the best fit to the data. Because of the between-study invariance constraints, parameter estimates (see Tables 3 and 4) are necessarily the same for the Canadian and Australian studies. For both self-responses and responses by others, the variance components for trait-factors (.530 and .637) are substantially higher than the variance components for method-factors (.065 and .063) and the variance components for uniquenesses (.157 and .555). Interestingly, trait-variance in the responses by others is slightly higher than trait-variance in self-responses, even though trait-variance accounts for a higher proportion of the total variance in self-responses than in response by others. The explanation of this apparent anomaly is that the variance in the responses by others is somewhat higher (see appendix).

Results: Self-other Discrepancies on Multiple Dimensions of Self-concept.

Mean differences in responses by students and by the significant others were tested in both the Canadian and Australian studies. For present purposes, a 2 (Country -- Canada vs. Australia) x 2 (methods -- self-response vs. ratings by others) x 13 (traits -- the 13 SDQIII scales) ANOVA was conducted in which country was a between-subjects factor and traits and methods were within-subject factors (Table 6). Mean responses and supplemental analyses for each separate trait are presented in Table 7. Across all 13 SDQIII scales there were no significant differences between Canadian and Australian responses. The effect of country did, however, vary somewhat with the particular trait. In separate analyses of each trait (Table 7) Canadian responses were significantly higher for 2 traits, significantly lower for 1 trait, and did not differ significantly from Australian responses for the remaining 10 traits.

Insert Tables 6 and 7 About Here

Across all 13 SDQIII traits, responses by significant others were higher

than self-responses, though this main effect varied somewhat depending on the trait and the country. In separate analyses of each trait, responses by others were significantly higher for 9 traits, significantly lower for 2 traits, and did not differ significantly from self-responses on 2 traits. Whereas responses by significant others were higher for both the Canadian and the Australian studies, this difference was somewhat larger in the Australian study. In the separate analyses of each trait, however, this interaction was significant for only 1 of 13 traits.

In summary, the comparisons of mean responses in the two studies revealed very small differences. Whereas some differences were statistically significant due to the very large sample size, none of the effects involving country, method, or their interactions with other variables accounted for as much as 1/2 of 1% of the variance. Whereas responses by significant others were higher than self-responses on most of the SDQIII scales in both studies, these differences were not large. Because self-other discrepancies are in the direction of higher responses by significant others in both studies, these results argue against a self-favorability bias in the self-responses and suggest a modesty effect. Because the modesty effect is slightly larger in the Canadian study than in the Australian study, the results offer no support for the Marsh and Richards (in press) suggestion that such an effect may be idiosyncratic to responses by Australians.

Discussion and Implications.

Substantive Issues

The results summarized here have important implications for further theoretical work in self-concept. The most obvious implication is for the study of ratings-by-others. However, interpreting the self-other agreement found here in terms of theory and previous research is difficult because of the ambiguous use of ratings-by-others and ambiguities in the theory underlying this research. The present findings apparently are consistent with the Shavelson et al. (1976) model, but Shavelson et al. intentionally de-emphasized the use of inferred self-concept responses and argued that they should not necessarily have any close correspondence with the preferred self-report measures of self-concept. Crandall (1973) suggested the pragmatic use of ratings-by-others as a means of validating self-concept measures, though the suggestions were apparently not based upon any specific theoretical position. The present findings do offer support for this application of inferred self-concepts. The results apparently support the symbolic interactionist perspective, but several qualifications are necessary. First, the prediction of self-other agreement in the symbolic interactionist theory is not clear-cut (see Kinch, 1963; Marsh, Barnes & Hocevar, 1985; Shibutani,

1961; Wells & Marwell, 1976). In particular, subject's perceptions of how he or she is viewed by others, and how these are related to self-perceptions were not examined here. The symbolic interactionist perspective posits that perceptions of others "cause" self-concept, but the direction of causality was not tested here. Results summarized here do, however, unambiguously demonstrate that significant others are able to accurately infer multiple self-concepts of a person who they know well, and this empirical relationship has practical and theoretical implications.

MTMM studies based on the SDQI found significant self-other agreement between self-concept ratings by school children and self-concepts inferred by their teachers. Student-teacher agreement tended to be stronger in areas where teachers were able to make relevant observations (academic facets, physical ability, and peer relations), but was poorer for Parent Relations and Physical Appearance. The poor agreement on Parent Relations was not surprising, since this is the area where teachers and peers are least likely to have an adequate basis for accurately inferring self-concepts. In the present investigation where many of the significant others were parents — particularly in the Australian study — self-other agreement on the Parents scales was very high. The lack of agreement on Physical Appearance found in the preadolescent research was unexpected. Perhaps, the standards used by teachers to infer Physical Appearance are different from those used by students, but even student-peer agreement on this factor was poor (Marsh, Smith & Barnes, 1984). This suggests that students may be using idiosyncratic standards in forming their own self-concepts of Physical Appearance and that these standards may not generalize to those that they employ in making ratings about one of their classmates. For responses by university students considered here, self-other agreement on Physical appearance was much better, but still below the average for all scales. Perhaps, by this age, respondents are using internal standards that are more similar to those used by significant others, but this is clearly an area where further research is needed.

Shavelson, et al. (1976) predicted that self-other agreement would be lower on general dimensions of self-concept near the apex of their hierarchy than dimensions closer to the base of their hierarchy that are more directly related to observable behavior. Based on their review of self-other agreement in personality research, McCrae and Costa (1988) also suggested that agreement would be higher on traits that are more observable. In terms of the Shavelson, et al. predictions, the two most general SDQIII scales are the General-Self and General-Academic scales. Averaged across the Australian and

Canadian studies (Table 2) self-other agreement is lower for these two scales than any other SDQIII scale. Only two other scales have self-other correlations nearly as low as the two general scales, but they are also less reliable than other SDQIII scales. Hence, particularly after correction for unreliability, self-other agreement is clearly lowest for the two general scales. Because the General-self scale was added to the SDQI in its last revision, self-other correlations are available for only two studies (Marsh, 1988, Table 14). Based on these two studies, however, self-other agreement on the General-self scale is the lowest of all the SDQI scales. Because self-other studies using the SDQI are based on agreement between students and teachers, it is not surprising the that self-other agreement is higher on the general academic scale than some of the non-academic scales, but self-other agreement on the general academic scale is lower than for the either the Mathematics or the Verbal scale. Thus, this prediction from the Shavelson model is supported by results of the present investigation.

The poor self-other agreement on the General-Self and General-Academic scales supports theoretical predictions, but the practical implications of these results may be even more important. Historically, self-concept researchers relied almost solely on general or global scales and this practice is still prevalent. Marsh and Shavelson (1985), however, argued that self-concept cannot be adequately understood if its multidimensionality is ignored. In support of this contention the present investigation found an average correlation of only .094 among the 13 SDQIII scales (see Appendix), so that a single global dimension cannot adequately account for the specific scales. Marsh and Shavelson argued if the role of self-concept is to better understand the complexity of self in different contexts, to predict a variety of behaviors, to provide outcome measures for diverse interventions, and to relate self-concept to other constructs, then multiple specific dimensions of self-concept are more useful than than a single general scale. In support of these contentions Marsh (in press-b) demonstrated that general self-concept tends to be less stable over time than other self facets, is less related to other external constructs than the specific dimensions most logically related to the construct, and is less sensitive to interventions designed to enhance self-concept than the specific dimensions that most closely match the intended outcomes of the intervention. More recently, focusing specifically on academic self-concept, Marsh, Byrne, and Shavelson (1988) made a similar distinction between general general academic self-concept and more content specific dimensions such as Verbal and Math self-concept. In support of this contention, the two studies presented here found that Math and Verbal scales were <u>negatively</u> correlated for self-responses (-

CFA model (-.28, Table 5) that was based on both studies. If Math and Verbal self-concept are uncorrelated, or even negatively correlated, than they cannot be adequately explained by a single dimension of academic self-concept (Marsh, 1986). Thus, research on self-other agreement contributes to a growing body of research calling into question the usefulness of the General-Self and General-Academic scales.

Self-other agreement found here is substantially higher than found in other self-concept research, and so it is informative to evaluate why. The apparent reasons are that: a) subjects were older (e.g., subjects knew themselves better or based their self-responses on more objective, observable criteria); b) both subjects and significant others made their responses on the same well developed instrument consisting of multi-item scales; c) self-other agreement was for specific characteristics rather than for broad, ambiguous characteristics or an overall self-concept; and d) the significant others knew the subjects better and in a wider range of contexts than the observers in most research. Marsh, Barnes and Hocevar (1985) demonstrated support for "b" by showing the self-other agreement and support for divergent validity were weaker when based on single-item ratings than on multi-item scales. Results from both the Australian and Canadian studies support "c" in that self-other agreement is weaker for General Self and General Academic scales. Support for "d" comes from the comparison of the present results and Marsh, Barnes and Hocevar with Marsh and Richards where the others did not know the subjects as well. Limited support for "a" comes from the comparison of the present results with those of the preadolescent (SDQI) studies, though there are so many differences that this comparison must be viewed cautiously. In related personality research, Funder (1987), Kenrick and Funder (1988), and particularly McCrae and Costa (1988) suggested that a similar set of characteristics would enhance self-other agreement on personality traits.

Whereas there is at least reasonable support for the suggestions as to why self-other agreement on self-concept is better here than found elsewhere -- except, perhaps, for age differences in self-other agreement -- there is also need for further research. Because the significant others in both studies were selected as the person who knew subjects the best, it might be argued that the nature of the relationship (e.g., romantic partners, siblings, parent-child) was of secondary importance. Nevertheless, it would be relevant determine if there were systematic differences in self-other agreement depending on the nature of the relationship. Alternatively, asking each subject to select two others who differed in terms of how well they

knew the subject would provide a more direct test of the prediction that knowing the subject better leads to better self-other agreement. There are substantial and consistent differences in the level of self-other agreement in different SDQIII scales. The poorer agreement for the general scales is consistent with predictions, but it would be useful to ask subjects and others to rate the scales in terms of specificity/generality, publicly observable/private, the consistency/variability of the subject on the trait being considered, and their confidence in their ratings.

Though not a focus of the present investigation, the distinction between personality and self-concept needs further consideration. Personality characteristics are apparently a subset of the characteristics that individuals use in forming their self-concepts. Self-concept ratings, however, are designed to reflect (private) self-perceptions whereas personality traits are designed to reflect (public) objective traits. From this distinction it follows that external observers should evaluate how subjects feel about themselves in order to validate self-concept ratings but should evaluate objective characteristics in order to validate personality ratings. There is, however, apparently no empirical evidence that observers can distinguish between how subjects feel about themselves vs. their objective characteristics. It also follows from this distinction that some influences that bias personality ratings may represent a valid influence on self-concept ratings. Nevertheless, personality and self-concept instruments both ask people to describe themselves and may be operationally indistinguishable. Thus, it may be that these distinctions are an artifact of the distinct historical developments of these two constructs.

Methodological Implications

The major focus of the present investigation was on substantive issues in the study of self-concept, but the research also extended previous applications of the the CFA approach to MTMM data. Whereas the CFA approach is apparently becoming increasingly popular, this appears to be the first application in which formal tests of factorial invariance were used to test simultaneously the equivalence of results based on different methods within the same study and across two parallel MTMM studies. This new multigroup CFA approach offers a substantially stronger basis for testing the generality of findings across different studies — particularly when, as in the present investigation, there is reasonable support for the equality all parameter estimates in the different studies. In this respect, the present investigation provides an important demonstration of a new analytic tool that may have broad applicability in personality and social psychological research.

Baldwin, J. M. (1897). Social and ethical interpretations in mental development: A study of social psychology. New York: McMillian.

Bem, D. J., & Allen, A. (1974). On predicting some of the people some of the time: The search for cross-situational consistencies in behavior. Psychological Review, 81, 88-104.

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

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

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

Cheek, J. M. (1982). Aggregation, moderator variables, and the validity of personality tests: A peer-rating study. Journal of Personality and Social Psychology, 43, 1254-1269.

Cooley, C. H. (1902). Human nature and social order. New York: Scribner.

Coombs, A., Soper, D., and Courson, C. (1963). The measurement of self-concept and self report. Educational and Psychological Measurement, 23, 493-500.

Conley, J. J. (1985). Longitudinal stability of personality traits A multitrait-multimethod-multioccasion analysis. Journal of Personality and Social Psychology, 49, 1266-1282.

Costa, P. M. & McCrae, R.R. (1988). Personality in adulthood: A six year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. Journal of Personality and Social Psychology, 54, 853-863.

Crandall, R. (1973). The measurement of self-esteem and related constructs. In J. Robinson & P. Shaver (Eds.), Measures of Social Psychological Attitudes (pp. 45-158). Ann Arbor, MI: Institute for Social Research.

Duval, S., & Wicklund, R. A. (1972). A theory of objective self-awareness. New York: Academic Press.

Funder, D. C. (1987). Errors and mistakes: Evaluating the accuracy of social judgment. Psychological Bulletin, 101, 75-90.

James, W. (1890/1963). The principles of psychology. New York: Holt, Rinehart & Winston.

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

Joreskog, K. G. & Sorbom, D. (1988). LISREL 7: A guide to the program and applications. Chicago: SPSS Inc.

Kenrick, D. T. & Funder, D. C. (1988). Profiting from controversy: Lessons

from the person-situation debate. American Psychologist, 43, 23-34.

Kenny, D. A., & Albright, L. (1987). Accuracy in interpersonal perception: A social relations analysis. Psychological Bulletin, 102, 390-402.

Kinch, J. W. (1963). A formalized theory of self-concept. American Journal of Sociology, 68, 481-486.

Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. American Educational Research Journal, 23, 129-149.

Marsh, H. W. (1988). Self Description Questionnaire: A Theoretical and empirical basis for the Measurement of multiple dimensions of preadolescent self-concept: A test manual and a research monograph. San Antonio, TX: The Psychological Corporation.

Marsh, H. W. (1989). Multitrait-multimethod analyses. In J. P. Keeves (Ed.), Educational research methodology, measurement and evaluation: An international handbook. Oxford, Pergamon Press.

Marsh, H. W. (In Press-a). Confirmatory factor analyses of multitrait-multimethod data: Many problems and a few solutions. Applied Psychological Measurement, .

Marsh, H. W. (in press-b). A multidimensional, hierarchical self-concept: Theoretical and empirical justification. Educational Psychology Review.,

Marsh, H. W. (in press-c). Self Description Questionnaire (SDQ) III: A Theoretical and empirical basis for the Measurement of multiple dimensions of late adolescent self-concept: An interim test manual and a research monograph. San Antonio, TX: The Psychological Corporation.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indices in confirmatory factor analysis: The effect of sample size. Psychological Bulletin, 103, 391-410.

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

Marsh, H. W., Byrne, B. M., & Shavelson, R. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. Journal of Educational Psychology, 80, 366-380.

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

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., McDonald, R. P., & Balla, J. R. (1988). Goodness of fit in

confirmatory factor analysis: The effect of sample size and of over-fitting or under-fitting models. In review [also see ERIC TM010581]

Marsh, H. W., & O'Neill, 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. & Richards, G. (in press). Self-other agreement and self-other differences on multidimensional self-concept ratings. Journal of Personality,.

Marsh, H. W., Richards, G., & Barnes, J. (1986a). Multidimensional self-concepts: A longterm followup of the effect of participation in an Outward Bound program. Personality and Social Psychology Bulletin, 12, 475-492.

Marsh, H. W., Richards, G. & Barnes, J. (1986b). Multidimensional self-concepts: The effect of participation in an Outward Bound program. Journal of Personality and Social Psychology, 45, 173-187.

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

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. & Butler, S. (1983). Self-concept: Reliability, dimensionality, validity, and the measurement of change. Journal of Educational Psychology, 75, 772-790.

Mead, G. H. (1934). Mind, self and society. Chicago: University of Chicago Press.

McCrae, R. R., & Costa, P. T. (1988). Different points of view: Self-reports and ratings in the assessment of personality. Paper presented at the XXIV International Congress of Psychology, Sydney, Australia.

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

Shibutani, T. (1961). Society and personality. Englewood Cliffs, NJ: Prentice-Hall.

Shrauger, J. S., & Schoeneman, T. J. (1979). Symbolic interactionist view of self-concept: Through the glass darkly. Psychological Bulletin, 86, 549-573.

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.

Widaman, K. F. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. Applied Psychological Measurement, 9, 1-26.

Table 1

Summary of Factor Analyses of SDQIII Responses For the Total Normative Archive (Norm Group; Marsh, in press-c), Australian self-responses, Australian responses by others, Canadian self-responses, and Canadian responses by others.

	Norm Group (n=2436)	Aust Self (n=151)	Aust Others (n=151)	Canadian Self (n=941)	Canadian Others (n=941)
Target Loadings					
No. of Coefficients	68	68	68	68	68
Highest	.94	.92	.91	.93	.92
Lowest	.44	.23	.40	.40	.43
Median	.71	.67	.72	.69	.69
% > .30	100.0%	96.0%	100.0%	100.0%	100.0%
Non-Target Loadings					
No. of Coefficients	716	716	716	716	716
Highest	.25	.43	.36	.33	.30
Lowest	-.17	-.28	-.21	-.11	-.11
Median	.02	.02	.02	.02	.02
% > .30	0.0%	3.0%	0.1%	0.1%	0.0%
Factor Correlations					
No. of Coefficients	78	78	78	78	78
Highest	.36	.32	.33	.35	.36
Lowest	-.06	-.19	-.14	.11	-.06
Median	.10	.07	.08	.09	.08
% > .30	5.1%	1.2%	1.2%	3.8%	2.6%

Note: As described in the test manual (Marsh, in press-c) the 136 SDQIII items are used to form 68 item pairs (the first two items within each scale form one pair, the next two items within each scale a second item-pair, and so forth) that are the basis of factor analyses. Target loadings are the factor loadings of the 68 item-pairs on the factor each is designed to measure, while all other factor loadings are Nontarget loadings. Factor Correlations are the factor pattern correlations among the 13 oblique correlations identified in each analysis. The norm group includes the Australian self-responses but not any of the other responses considered here. Factor scores considered in subsequent analyses were derived from the factor analysis of the norm group as described in the test manual.

Table 2

Internal Consistency and Self-other Agreement in studies on Australian University Students (N=151) and Canadian Australian Students (N=941).

	Internal Consistency				Self-other Correlations	
	Australian		Canadian			
SDQIII Scales	Self	Other	Self	Other	Australian	Canadian
1) Physical Ability	.96	.97	.96	.96	.780	.754
2) Appearance	.86	.89	.90	.90	.461	.454
3) Opposite Sex Peers	.90	.90	.91	.90	.520	.587
4) Same Sex Peers	.86	.90	.85	.84	.457	.473
5) Parents	.91	.93	.89	.89	.759	.659
6) Honesty	.74	.81	.80	.82	.394	.437
7) Religion	.95	.93	.94	.94	.800	.745
8) Emotional	.91	.93	.89	.89	.624	.482
9) General	.94	.93	.94	.92	.403	.422
10) Mathematics	.95	.95	.95	.95	.741	.769
11) Verbal	.84	.86	.86	.86	.638	.614
12) Academic	.86	.88	.87	.89	.311	.440
13) Problem Solving	.79	.82	.80	.81	.490	.400
Mean	.88	.90	.89	.89	.568	.560

NOTE: Internal consistency refers to coefficient alpha estimates of reliability. Self-other agreement in both studies is based on factor scores (see Table 1). All correlations in both studies are statistically significant. When the set of 13 self-other agreemsnt coefficients one study were correlated with those in the other study, the correlation was $r = .908$ ($p < .001$, $df = 11$).

Table 3

Goodness of Fit for CFA Models of MTMM Data From the Australian and Canadian Studies

Model	Within Study Constraints	Between Study Constraints	X^2	df	X^2/df Ratio	TLI	BBI
Null Model							
0	————	————	10135	351	15.593	——	——
13-Traits/0-Methods Total Invariance							
1a[a]	None	None	1778	442	4.022	.793	.825
2b	All (T,U)	None	1922	494	3.892	.802	.810
1c	None	All (T,C,U)	1989	572	3.478	.830	.804
1d	All (T,U)	All (T,C,U)	2079	598	3.476	.830	.795
13-Traits/2-Methods Total Invariance							
2a[a]	None	None	632	390	1.620	.958	.938
2b	All (T,M,U)	None	926	468	1.979	.933	.909
2c	None	All (T,M,C,U)	870	546	1.593	.959	.862
2d	All (T,M,U)	All (T,M,C,U)	1096	585	1.874	.940	.892
13-Traits/2-Methods Between-Study Invariance							
3a	None	T	688	416	1.654	.955	.932
3b	None	T,M	725	442	1.641	.956	.928
3c	None	T,M,U	748	468	1.598	.959	.926
3d	None	T,M,C	846	520	1.626	.957	.917
4a[b]	None	All Self	818	507	1.614	.958	.919
4b[b]	None	All Other	827	507	1.632	.957	.918
13-Traits/2-Methods Within- and Between-Study Invariance							
5a	T	All (T,M,C,U)	942	559	1.686	.953	.907
5b	M	All (T,M,C,U)	1011	559	1.809	.945	.900
5c	None	T,M,U	1042	572	1.823	.944	.897

Note. The null model (Model 0) is of no substantive interest, but is used in the definition of the Bentler Bonett Index (BBI) and the Tucker Lewis Index (TLI). Models 1a-1d posit only trait factors whereas models 2a-5c posit trait and method factors. For within-study constraints, total invariance consists of constraining trait-factor loadings (T), method-factor loadings (M) and error/uniquenesses (U) to be the same for self responses and the corresponding responses by others. For between-study constraints, total invariance consists of constraining trait-factor loadings (T), method-factor loadings (M), factor covariances (C) and error/uniquenesses (U) to be the same in the Canadian and Australian studies.

a — For just these two models, the solutions were improper in that uniquenesses in the Australian data were slightly (nonsignificantly) negative. In order to obtain proper solutions the offending parameters were fixed to a small positive value and the remaining parameters were re-estimated. Solutions for the Canadian data with these models and solutions for both Canadian and Australian data for all other models were proper.

b — For these two models, total between-study invariance was imposed on only the entire set of parameter estimates based on self-responses (model 4a) or on only those based on responses by significant others (model 4b).

Table 4

Trait-factor loadings, Method-factor loadings, and Error/Uniquenesses for Model 2c (Total Between-study invariance, No Within-study invariance).

	Self-responses			Responses by others		
SDQIII Scales	Trait	Method	Error	Trait	Method	Error
1) Physical Ability	1.083	0.111	0.502	1.219	0.221	0.873
2) Appearance	0.589	0.354	0.384	0.647	0.122	0.549
3) Opposite Sex Peers	0.580	0.358	0.451	0.849	0.135	0.921
4) Same Sex Peers	0.571	0.328	0.463	0.748	0.224	0.793
5) Parents	0.755	0.215	0.445	0.953	0.140	0.908
6) Honesty	0.572	0.235	0.529	0.746	0.307	0.638
7) Religion	0.797	0.009	0.281	0.843	0.030	0.869
8) Emotional	0.633	0.354	0.433	0.800	0.071	0.712
9) General	0.529	0.537	0.367	0.683	0.172	0.638
10) Mathematics	1.081	0.021	0.227	0.942	0.205	0.732
11) Verbal	0.805	0.159	0.199	0.642	0.425	0.744
12) Academic	0.561	0.221	0.243	0.466	0.344	0.482
13) Problem Solving	0.605	0.142	0.362	0.547	0.437	0.669
Variance Component	0.530	0.065	0.157	0.637	0.063	0.555

Note: The variance component is the mean of the squared coefficients in each column of the table.

Table 5

Trait-factor Variances and Covariances

SDQIII Scales	1	2	3	4	5	6	7	8	9	10	11	12	13
1) Physical Ability	1.00												
2) Appearance	.38	1.00											
3) Opposite Sex Peers	.16	.38	1.00										
4) Same Sex Peers	.19	.12	.28	1.00									
5) Parents	.03	.01	-.04	.23	1.00								
6) Honesty	-.14	-.07	-.15	.03	.22	1.00							
7) Religion	.00	.04	-.07	.08	.25	.22	1.00						
8) Emotional	.22	.13	.22	.36	.24	.02	.04	1.00					
9) General	.13	.45	.35	.37	.17	.10	.12	.56	1.00				
10) Mathematics	.06	-.09	-.12	-.07	.03	-.09	-.03	.06	-.02	1.00			
11) Verbal	-.16	.02	.11	.06	-.04	.24	-.08	.05	.17	-.28	1.00		
12) Academic	-.07	.01	.01	-.04	-.01	.29	.06	-.01	.22	.16	.56	1.00	
13) Problem Solving	.09	.15	.14	-.08	-.17	-.03	-.03	.12	.29	.19	.35	.36	1.00

NOTE: Because factor covariances were all fixed to have variances equal to 1.0 in order to fix the scale of the solution, factor covariances presented here are the same as factor correlations.

Table 6

ANOVA Comparing Responses to the 13 SDQIII scales (traits) in the Canadian and Australian Studies

Source	SS	df	F-Ratio	p	Variance Explained
Country (Canada vs. Australia)	0.1	1	0.02	.893	0.00%
Subjects (error)	3431.2	1090	----	----	11.73%
Method (self vs. other ratings)	69.4	1	69.44	.000	0.23%
Country x method	4.1	1	3.95	.047	0.01%
Subjects x methods (error)	1135.7	1090	----	----	3.88%
Traits (13 SDQIII scales)	1700.0	12	101.75	.000	5.81%
Country x traits	68.4	12	4.09	.000	0.23%
Subjects x traits (error)	18210.3	13080	----	----	62.26%
Method x traits	100.2	12	24.12	.000	0.34%
Country x method x traits	3.5	12	0.85	.617	0.01%
Subjects x method x traits (error)	4527.82	13032	----	----	15.48%
Total	29250.7	28391	----	----	100.00%

Note. Variance explained is the ratio of the sums of squares (SS) due to each effect over the Total SS multiplied by 100%.

Table 7

Mean responses and the effects of Method (self vs. other), Country (Australia vs. Canada) and their interaction on responses to each of the 13 SDQIII scales (traits)

Score	Self-ratings Aust	Canada	Other ratings Aust	Canada	Country	Method	Interaction
Physical	-.41	-.28	-.45	-.35	ns	**	ns
Appearance	-.01	.11	.26	.32	ns	***	ns
Opposite Sex	.31	.15	.31	.25	ns	**	ns
Same Sex	.03	.15	.09	.29	*	***	ns
Parents	-.27	-.04	-.49	-.17	***	***	ns
Honesty	-.01	-.16	.02	.02	ns	***	ns
Religion	.16	-.11	.11	-.10	**	ns	ns
Emotion	-.31	-.51	-.25	-.34	ns	***	ns
General	.12	.22	.15	.29	ns	*	ns
Math	-.38	-.57	-.37	-.38	ns	***	**
Verbal	.21	.08	.44	.35	ns	***	ns
Academic	.24	.12	.21	.14	ns	ns	ns
Problem Solve	-.25	-.24	-.05	.03	ns	***	ns
Mean	-.04	-.08	-.01	.03			

Note. Standard deviations corresponding to each of the mean responses is presented in the appendix.

* p <.05, ** p < .01, *** p < .001.

Appendix

Means, Standard Deviations, and Correlations Among the 26 Variables (13 SDOIII factors for self-responses and responses by others) from the Australian and the Canadian Studies.

Canadian Study

	SELF-CONCEPT SCALES (SELF-RESPONSES)													INFERRED SELF-CONCEPT SCALES (RESPONSES BY OTHERS)											
	PHYS	APPR	OSEX	SSEX	PRNT	HONS	RELG	EMOT	GEN	MATH	VERB	ACAD	PROB	PHYS	APPR	OSEX	SSEX	PRNT	HONS	RELG	EMOT	GEN	MATH	VERB	ACAD	PROB
Self-responses																										
PHYS	1.00	.27	.09	.15	.03	-.11	-.02	.15	.12	.07	-.05	.00	.12	.75	.19	.11	.14	.00	-.13	-.04	.14	.05	.04	-.16	-.08	.02
APPR	.27	1.00	.35	.10	.05	.01	.02	.14	.42	-.06	.08	.11	.16	.26	.45	.23	.00	-.03	-.04	.03	.06	.20	-.06	.03	.03	.06
OSEX	.09	.35	1.00	.23	.07	.01	-.09	.18	.31	-.10	.17	.10	.13	.09	.19	.59	.05	-.02	-.07	-.04	.01	.17	-.11	.07	.03	.06
SSEX	.15	.10	.23	1.00	.18	.08	.02	.24	.24	-.07	.11	.09	.03	.12	.04	.15	.47	.11	-.01	.05	.15	.16	-.07	.01	-.02	-.04
PRNT	.03	.05	.07	.18	1.00	.19	.20	.18	.16	-.03	.02	.06	-.05	.04	.03	-.01	.11	.66	.10	.19	.09	.06	-.01	-.00	.05	-.05
HONS	-.11	.01	.01	.08	.19	1.00	.13	.11	.12	-.06	.18	.22	.04	-.09	-.07	-.06	-.03	.10	.44	.11	-.06	.04	-.06	.16	.16	.03
RELG	-.02	.02	-.09	.02	.20	.13	1.00	-.01	.09	-.05	-.04	.05	.01	-.03	.00	-.07	.05	.19	.10	.74	.02	.06	-.03	-.03	.04	-.01
EMOT	.15	.14	.18	.24	.18	.11	-.01	1.00	.37	.04	.11	.12	.12	.11	.05	.13	.17	.12	.00	.03	.48	.26	.02	.02	.04	.03
GEN	.12	.42	.31	.24	.16	.12	.09	.37	1.00	.00	.16	.26	.23	.11	.19	.20	.11	.07	.01	.09	.19	.42	-.03	.06	.08	.11
MATH	.07	-.06	-.10	-.07	-.03	-.06	-.05	.04	.00	1.00	-.26	.07	.13	.08	-.06	-.10	-.05	.01	-.03	-.04	.07	-.01	.77	-.16	.12	.15
VERB	-.05	.08	.17	.11	.02	.18	-.04	.11	.16	-.26	1.00	.41	.28	-.07	.02	.13	.04	-.05	.15	-.06	.02	.15	-.19	.61	.27	.19
ACAD	.00	.11	.10	.09	.06	.22	.05	.12	.26	.07	.41	1.00	.22	-.04	.00	.05	-.02	.00	.15	.03	-.01	.13	.08	.32	.44	.20
PROB	.12	.16	.13	.03	-.05	.04	.01	.12	.23	.13	.28	.22	1.00	.04	.07	.08	-.05	-.09	-.05	-.05	.07	.15	.07	.10	.14	.40
Significant Other Responses																										
PHYS	.75	.26	.09	.12	.04	-.09	-.03	.11	.11	.08	-.07	-.04	.04	1.00	.29	.17	.19	.03	-.06	-.02	.18	.10	.07	-.10	.01	.10
APPR	.19	.45	.19	.04	.03	-.07	.00	.05	.19	-.06	.02	.00	.07	.29	1.00	.30	.14	.07	.06	.06	.11	.26	-.06	.07	.06	.09
OSEX	.11	.23	.59	.15	-.01	-.06	-.07	.13	.20	-.10	.13	.05	.08	.17	.30	1.00	.24	-.04	-.09	-.04	.12	.27	-.08	.13	.08	.14
SSEX	.14	.00	.05	.47	.11	-.03	.05	.17	.11	-.05	.04	-.02	-.05	.19	.14	.24	1.00	.17	.10	.09	.21	.28	-.02	.13	.08	.07
PRNT	.00	-.03	-.02	.11	.66	.10	.19	.12	.07	.01	-.05	.00	-.09	.03	.07	-.04	.17	1.00	.23	.23	.17	.13	.06	.03	.09	-.01
HONS	-.13	-.04	-.07	-.01	.10	.44	.10	.00	.01	-.03	.15	.15	-.05	-.06	.00	-.09	.10	.23	1.00	.12	.02	.14	.00	.29	.25	.11
RELG	-.04	.03	-.04	.05	.19	.11	.74	.03	.09	-.04	-.06	.03	-.05	-.02	.06	-.04	.09	.23	.12	1.00	.06	.09	-.03	-.04	.03	.01
EMOT	.14	.06	.01	.15	.09	-.06	.02	.48	.19	.07	.02	-.01	.07	.18	.11	.12	.21	.17	.02	.06	1.00	.39	.04	.03	.05	.09
GEN	.05	.20	.17	.16	.06	.04	.06	.26	.42	-.01	.15	.13	.15	.10	.26	.27	.28	.13	.14	.09	.39	1.00	.01	.19	.20	.22
MATH	.04	-.06	-.11	-.07	-.01	-.06	-.03	.02	-.03	.77	-.19	.08	.07	.07	-.06	-.08	-.02	.06	.00	-.03	.04	.01	1.00	-.06	.22	.25
VERB	-.16	.03	.07	.01	-.00	.16	-.03	.02	.06	-.16	.61	.32	.10	-.10	.07	.15	.13	.03	.29	-.04	.03	.19	-.06	1.00	.46	.38
ACAD	-.08	.03	.03	-.02	.05	.16	-.04	.04	.08	.12	.27	.44	.14	.01	.06	.08	.08	.09	.25	.03	.05	.20	.22	.46	1.00	.35
PROB	.02	.06	.06	-.04	-.05	.03	-.01	.03	.11	.15	.19	.20	.40	.10	.09	.14	.07	-.01	.11	.01	.09	.22	.25	.38	.35	1.00
Means	-.28	.11	.15	.15	-.04	-.16	-.11	-.51	.22	-.57	.08	.12	-.24	-.36	.32	.25	.29	-.17	.02	-.10	-.34	.29	-.38	.35	.14	.03
SDs	1.30	.93	.97	.93	1.01	.96	.94	.95	.97	1.19	.93	.79	.86	1.31	.90	.89	.89	1.00	1.03	.90	.97	.87	1.14	.89	.81	.84

Austrian Study

	SELF-CONCEPT SCALES (SELF-RESPONSES)													INFERRED SELF-CONCEPT SCALES (RESPONSES BY OTHERS)											
	PHYS	APPR	OSEX	SSEX	PRNT	HONS	RELG	EMOT	GEN	MATH	VERB	ACAD	PROB	PHYS	APPR	OSEX	SSEX	PRNT	HONS	RELG	EMOT	GEN	MATH	VERB	ACAD	PROB
Self-Responses																										
PHYS	1.00	.12	-.01	.04	-.08	.00	.08	.16	.01	-.04	-.22	-.16	-.02	.78	.07	.10	-.01	.02	.04	.10	.15	.09	-.12	-.20	-.01	.01
APPR	.12	1.00	.11	.12	-.10	.02	-.08	-.02	.26	.08	.14	.19	.12	.09	.46	.02	-.00	-.10	-.00	-.12	-.05	.13	-.07	.10	.12	.16
OSEX	-.01	.11	1.00	.27	.19	.18	-.03	.34	.33	.10	.16	-.02	.02	-.14	.12	.52	.16	.09	-.02	-.08	.26	.21	-.00	.02	-.15	-.06
SSEX	.04	.12	.27	1.00	.25	.16	.11	.31	.42	.18	.13	.09	-.01	-.05	.06	.16	.46	.22	-.07	.03	.18	.21	.06	-.05	-.02	-.03
PRNT	-.08	-.10	.19	.25	1.00	.09	.13	.16	.12	.12	.10	-.06	-.22	.06	.11	.10	.23	.76	.08	.06	.25	.19	.14	.15	-.05	-.11
HONS	.00	.02	.18	.16	.09	1.00	.13	.11	.13	.01	.12	.14	-.01	-.12	.12	-.04	.03	.02	.39	.17	.04	-.02	-.10	-.08	-.02	-.00
RELG	.08	-.08	-.03	.11	.13	.13	1.00	-.05	.02	.04	-.07	-.03	-.10	.13	.05	-.01	.09	.19	.29	.80	.07	.10	.09	-.03	.14	-.10
EMOT	.16	-.02	.34	.31	.16	.11	-.05	1.00	.43	.06	.08	-.01	-.03	.17	.06	.36	.27	.17	.05	-.05	.62	.27	-.04	.02	-.16	.13
GEN	.01	.26	.33	.42	.12	.13	.02	.43	1.00	.12	.07	.22	.12	.00	.21	.19	.19	.12	.02	-.00	.19	.40	-.05	-.02	-.03	.08
MATH	-.04	.08	.10	.18	.12	.01	.04	.06	.12	1.00	-.11	.09	.07	.01	.04	.00	.06	.12	-.03	.08	.04	.02	.74	-.09	.22	.09
VERB	-.22	.14	.16	.13	.10	.12	-.07	.08	.07	-.11	1.00	.41	.23	-.20	-.06	-.09	.02	-.02	.03	-.10	.08	.05	-.11	.64	.06	.17
ACAD	-.16	.19	-.02	.09	-.06	.14	-.03	-.01	.22	.09	.41	1.00	.26	-.10	-.10	-.19	-.19	-.14	.03	.00	-.21	-.01	.08	.24	.31	.11
PROB	-.02	.12	.02	-.01	-.22	-.01	-.10	-.03	.12	.07	.23	.26	1.00	-.03	.10	.09	-.16	-.18	-.02	.02	-.03	.13	.65	.12	.06	.49
Significant Other Responses																										
PHYS	.78	.09	-.14	-.05	.06	-.12	.13	.17	.00	.01	-.20	-.10	-.03	1.00	.19	.14	.02	.13	.12	.16	.17	.15	.01	-.04	.13	.13
APPR	.07	.46	.12	.06	.11	.12	.05	.06	.21	.04	-.06	-.10	.10	.19	1.00	.33	.24	.07	.16	.04	.19	.32	.02	-.02	-.06	.21
OSEX	.10	.02	.52	.16	.10	-.04	-.01	.36	.19	.00	-.09	-.19	.09	.14	.33	1.00	.38	.09	-.01	-.04	.46	.40	.01	-.03	-.06	.20
SSEX	-.01	-.00	.16	.46	.23	.03	.09	.27	.19	.06	.02	-.19	-.16	.02	.24	.38	1.00	.36	.15	.01	.38	.37	.03	.07	-.03	.01
PRNT	.02	-.10	.09	.22	.76	.02	.19	.17	.12	.12	-.02	-.14	-.18	.13	.07	.09	.36	1.00	.19	.14	.28	.28	.17	.08	.04	-.08
HONS	.04	-.00	-.02	-.07	.08	.39	.29	.05	.02	-.03	.03	.03	-.02	.12	.16	-.01	.15	.19	1.00	.33	.10	.15	.01	.18	.25	.20
RELG	.10	-.12	-.08	.03	.06	.17	.80	-.05	-.01	.08	-.10	.00	.02	.16	.04	-.04	.01	.14	.33	1.00	.03	.07	.09	-.10	.11	.02
EMOT	.15	-.05	.26	.18	.25	.04	.07	.62	.19	.04	.08	-.21	-.03	.17	.19	.46	.38	.28	.10	.03	1.00	.40	-.03	.05	-.05	.14
GEN	.09	.13	.21	.21	.19	-.02	.10	.27	.40	.02	.05	-.01	.13	.15	.32	.40	.37	.28	.15	.07	.40	1.00	-.03	.16	.11	.29
MATH	-.12	-.07	-.00	.06	.14	-.10	.09	-.04	-.05	.74	-.11	.08	.05	.01	.02	.01	.03	.17	.01	.09	-.03	-.03	1.00	-.04	.19	.15
VERB	-.20	.10	.02	-.05	.15	-.08	-.03	.02	-.02	-.09	.64	.29	.12	-.04	-.02	-.03	.07	.08	.18	-.10	.05	.16	-.04	1.00	.24	.36
ACAD	-.01	.12	-.15	-.02	-.05	-.02	.14	-.16	-.03	.22	.06	.31	.06	.13	-.06	-.06	-.03	.04	.25	.11	-.05	.11	.19	.24	1.00	.11
PROB	.01	.16	-.06	-.03	-.11	-.00	-.10	.13	.08	.09	.17	.11	.49	.13	.21	.20	.01	-.08	.20	.02	.14	.29	.15	.36	.11	1.00
Means	-.41	-.01	.31	.03	-.27	-.01	.16	-.31	.12	-.38	.21	.24	-.25	-.45	.26	.31	.09	-.49	.02	.11	-.25	.15	-.37	.44	.21	-.10
SDs	1.28	.86	.90	.94	1.14	.89	1.05	1.03	.90	1.10	.95	.73	.92	1.39	.91	.88	1.04	1.23	1.07	1.09	1.07	.84	1.15	.89	.78	.86