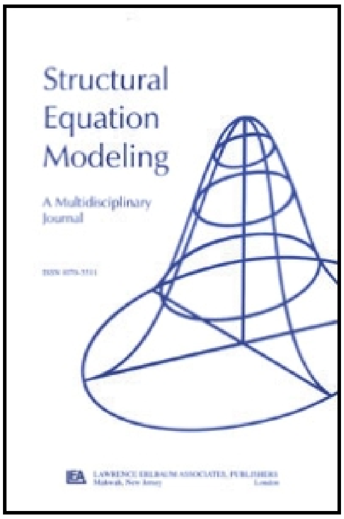

Structural Equation Modeling: A Multidisciplinary Journal
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t775653699

Classical Latent Profile Analysis of Academic Self-Concept Dimensions: Synergy of Person- and Variable-Centered Approaches to Theoretical Models of Self-Concept

Herbert W. Marsh [a]; Oliver Lüdtke [b]; Ulrich Trautwein [b]; Alexandre J. S. Morin [c]
[a] University of Oxford, Oxford, England [b] Max Planck Institute for Human Development, Berlin, Germany [c] University of Sherbrooke, Sherbrooke, Canada

Online Publication Date: 01 April 2009

To cite this Article Marsh, Herbert W., Lüdtke, Oliver, Trautwein, Ulrich and Morin, Alexandre J. S.(2009)'Classical Latent Profile Analysis of Academic Self-Concept Dimensions: Synergy of Person- and Variable-Centered Approaches to Theoretical Models of Self-Concept',Structural Equation Modeling: A Multidisciplinary Journal,16:2,191 — 225

To link to this Article: DOI: 10.1080/10705510902751010

URL: http://dx.doi.org/10.1080/10705510902751010

Structural Equation Modeling, 16:191–225, 2009

ISSN: 1070-5511 print/1532-8007 online
DOI: 10.1080/10705510902751010

Psychology Press
Taylor & Francis Group

Classical Latent Profile Analysis of Academic Self-Concept Dimensions: Synergy of Person- and Variable-Centered Approaches to Theoretical Models of Self-Concept

Herbert W. Marsh
University of Oxford, Oxford, England

Oliver Lüdtke and Ulrich Trautwein
Max Planck Institute for Human Development, Berlin, Germany

Alexandre J. S. Morin
University of Sherbrooke, Sherbrooke, Canada

In this investigation, we used a classic latent profile analysis (LPA), a person-centered approach, to identify groups of students who had similar profiles for multiple dimensions of academic self-concept (ASC) and related these LPA groups to a diverse set of correlates. Consistent with a priori predictions, we identified 5 LPA groups representing a combination of profile level (high vs. low overall ASC) and profile shape (math vs. verbal self-concepts) that complemented results based on a traditional variable-centered approach. Whereas LPA groups were substantially and logically related to the set of 10 correlates, much of the predictive power of individual ASC factors was lost in the formation of groups and the inclusion of the correlates into the LPA distorted the nature of the groups. LPA issues examined include distinctions between quantitative (level) and qualitative (shape) differences in LPA profiles, goodness of fit and the determination of the number of LPA groups, appropriateness of correlates as covariates or auxiliary variables, and alternative approaches to present and interpret the results.

In this investigation, we used classic latent profile analyses (LPAs) to explore a person-centered approach to the study of multiple dimensions of academic self-concept (ASC). Consistent with a person-centered approach, we identified groups of students with similar profiles of ASC factors and validated the LPA groups in relation to a variety of correlates. Traditionally, self-

Correspondence should be addressed to Herbert W. Marsh, Department of Educational, University of Oxford, 15 Norham Gardens, Oxford OX2 6PY, UK. E-mail: herb.marsh@education.ox.ac.uk

concept researchers have focused on a variable-centered approach and examined the structure of self-concept factors across persons, the unique effects of particular self-concept factors, and the relation of each self-concept factor to other variables. However, in our person-centered approach, we identified student types with similar patterns of ASC profiles, examined configurations of ASC factors within the person, and portrayed student self-perceptions in a holistic fashion that complemented and extended traditional variable-centered research (see Magnusson & Stattin, 1998). In contrast to a nomothetic, variable-centered approach that treats each ASC factor as a separate construct, our research takes a more idiographic, person-centered approach that takes into account intraindividual variation in different components of ASC to better represent multiple ASCs of a particular individual as an organized whole (Magnusson, 1990).

Robins, John, Caspi, Moffitt, and Stouthamer-Loeber (1996) noted that person-centered, typological approaches have a long history dating back to early work and theories such as that by Allport (1937), the Myers–Briggs types based on Jungian theory (Myers & McCaulley, 1985), typologies hypothesized in many psychological theories, and even the typological approaches posited by ancient Greek philosophers. Robins et al. (1996) emphasized that, "Trait dimensions and person typologies are complementary, not competing systems" (p. 170). Loken and Molenaar (2008) also emphasized that despite the long history of distinctions between person types and variable dimensions, both these latent variable modeling approaches share many similarities, and much can be learned by juxtaposing the two.

Importantly, our study is a substantive-methodological synergy (Marsh & Hau, 2007) in which we apply classical LPA to identify self-concept typologies with substantively important theoretical issues in self-concept research. Methodologically, the study provides a demonstration of an evolving methodological approach and attention to some unresolved issues—guidance on how they should be addressed when there are no clear guidelines or "golden rules" (or to challenge existing golden rules), and directions for further research in how to make the statistical procedure more useful.

ACADEMIC SELF-CONCEPT (ASC): STRUCTURE, DIMENSIONALITY AND RELATION TO ACHIEVEMENT

Shavelson, Hubner, and Stanton (1976) proposed a multidimensional, hierarchical model of self-concept that combined important aspects of a unidimensional perspective and a multidimensional perspective (see historical review by Marsh & Hattie, 1996). Broadly speaking, Shavelson et al. posited a single global component of self-concept at the apex of their hierarchy, which was then divided into academic (e.g., math, English, science, etc.) and nonacademic self-concepts (e.g., social, physical, emotional) at the next level. For the first of those, Shavelson et al. proposed that multiple dimensions of ASC—like the corresponding measures of achievement—would be substantially correlated and well represented by a single component of global ASC. However, subsequent empirical research (e.g., Marsh, 1986, 1989b, 2007b; Marsh, Byrne, & Shavelson, 1988) demonstrated that math and verbal self-concepts (MathSC and VerbSC) were nearly uncorrelated. Mathematics achievement was substantially correlated with MathSC but not VerbSC, whereas verbal achievement was substantially correlated with VerbSC but not MathSC.

Complications such as these led to the Marsh-Shavelson revision (Marsh, Byrne & Shavelson, 1988; Marsh & Shavelson, 1985) of the original Shavelson et al. model. To explain this extreme separation between math and verbal components of ASC, Marsh (1986) proposed the internal/external frame of reference (I/E) model. According to the I/E model, ASC in a particular school subject is formed in relation to two comparison processes. The first is the *external* (normative or nomographic) process in which students compare their self-perceived performances in a particular school subject with the perceived performances of other students in the same school subject, as well as other normative indexes of actual achievement (e.g., school grades in relation to implicit or explicit grade distributions). The external comparison process should result in substantial positive correlations between MathSC and VerbSC because math and verbal achievements are substantially positively correlated. The second is an *internal* (ipsative or idiographic) process in which students compare their own performance in one particular school subject with their own performances in other school subjects. Hence this internal comparison posits a within-person or person-centered process based on the configuration of math and verbal performances that is idiosyncratic to each individual student. The internal (ipsative) comparison process should result in a negative correlation between MathSC and VerbSC because as one domain is seen to be stronger, the other will seem comparatively weaker. The combination of the two processes, depending on how they are weighted, is consistent with the near-zero correlation that led to the Marsh–Shavelson revision of the original Shavelson et al. (1976) model. There is a large body of support for I/E predictions (see Marsh, 1986, 1990, 1992a, 1993, 2007b; Marsh, Byrne & Shavelson, 1988; Marsh & Yeung, 2001) that (a) math and verbal achievements are highly correlated, but MathSC and VerbSC are nearly uncorrelated; (b) math achievement has positive effects on MathSC, but negative effects on VerbSC; and (c) verbal achievement has positive effects on VerbSC, but negative effects on MathSC. Marsh and Hau (2004) used multigroup confirmatory factor analysis and structural equation modeling (SEM) to provide cross-cultural support for these predictions for nationally representative samples of 15-year-olds from 26 countries—the Organization for Economic Cooperation and Development Programme for international student assessment (OECD-PISA) database.

APPLICATION OF CLASSICAL LATENT PROFILE ANALYSIS

LPA is a person-centered approach, but is also related to factor analysis, in which the covariation of observed variables is explained by latent continuous variables. SEM and regression analyses take a variable-centered approach in which the aim is to predict outcomes, relate independent and dependent variables, or assess intervention effects. Person-centered approaches like cluster analysis, latent class analysis, and LPA focus on relations among individuals with the goal to sort individuals into groups of individuals who are similar to each other and different from those in other groups (Lubke & Muthén, 2005; Muthén, 2001; B. Muthén & Muthén, 2000; also see Pastor, Barron, Miller, & Davis, 2007, for a didactical application). Thus, Bauer and Curran (2004) emphasized that the key difference between factor analysis and LPA is that "the common factor model decomposes the covariances to highlight relationships among the variables, whereas the latent profile model decomposes the covariances to highlight relationships among individuals" (p. 6). In other words, factor models cluster indicator variables, whereas LPAs cluster individuals (Lubke & Muthén, 2005).

Data appropriate for LPA are typically assumed to consist of unobservable subgroups of individuals with different probability distributions. Although cluster analysis techniques are commonly used to divide persons into homogeneous subgroups, LPA offers several advantages over this approach (see Pastor et al., 2007). Whereas cluster analysis is an exploratory technique, LPA is a model-based procedure that allows for more flexible model specification. Indeed, traditional cluster analysis is equivalent to a very restricted specification of a latent profile model (Pastor et al., 2007; Vermunt & Magidson, 2002). Moreover, the fit indexes provided in LPA enable different models to be compared and informed decisions to be made regarding the number of underlying classes.

In this investigation, we demonstrate the application of classical LPA to identify groups of individuals with qualitatively distinct patterns of ASC profiles, and to validate these groups in relation to a diverse set of correlates (gender, achievement test scores, coursework selection, and school grades in different subjects). Multiple dimensions of ASC are ideally suited to demonstrate the strengths of LPA and should identify profiles of scores that differ qualitatively (profile shape) as well as quantitatively (profile level). This is important, because many applications of LPA in other domains result in profiles in which groups merely differ mainly in terms of quantitative level (e.g., uniformly high, medium, or low scores across all components of the profile). However, research in support of the I/E model (described earlier) suggests that an LPA of multiple dimensions of ASC should result in groups of students with qualitatively different profiles: groups with opposite levels of VerbSC and MathSC (high VerbSC and low MathSC or vice versa), as well as groups that differ in terms of overall levels of ASC (i.e., uniformly high, medium, or low ASCs). In particular, according to the I/E model described earlier, there is an idiographic process in the formation of different ASCs that suggests that a person-centered approach is needed. More specifically the LPA is designed to provide a more idiographic, holistic representation of specific configurations of ASC factors within each person and how these configurations relate to various correlates. Specific issues that we address are covered next.

Number of Groups and Goodness of Fit

To what extent are there "golden rules" that allow researchers to determine the number of groups that should be used to represent the ASCs? Typically, the first step in an LPA is to determine the number of groups with well-defined, differentiated profiles (e.g., Lubke & Muthén, 2005; Muthén, 2001; B. Muthén & Muthén, 2000; Pastor et al., 2007). Although there is diverse opinion on how to choose the correct number of groups, it is useful to explore solutions with varying numbers of groups and to select one that makes most sense in relation to theory, previous research, the nature of the groups, and interpretation of the results—as well as alternative goodness-of-fit indexes and tests of statistical significance. More recently, researchers have emphasized the selection of the number of groups based on goodness-of-fit indexes. When the models being compared are nested, one approach is to test the statistical significance to determine whether a more complex model is able to fit the data significantly better than a more parsimonious model. Although the difference in the log-likelihood ratio tests (LRT) for two such models does not have a chi-square distribution, the fit of the two models can be compared with a test developed by Lo, Mendell, and Rubin (LMR; 2001). Alternatively, bootstrapping can be used as operationalized in bootstrapped LRT (BLRT; McLachlan & Peel, 2000). More generally, researchers have relied on information criterion indexes—particularly the Baysian Information Criterion (BIC) and Akaike's Information Criterion (e.g., Collins, Fidler, Wugalter, & Long, 1993; Henson, Reise, & Kim, 2007; Magidson & Vermunt, 2004) or the sample-size-adjusted BIC (SSA–BIC; Yang, 2006). Tofighi and Enders's (2007) simulation study concluded that the SSA–BIC and LMR test of significance were the best in determining the appropriate number of groups. However, Nylund, Asparouhov, and Muthén's (2007) simulation study argued that although the SSA–BIC outperforms the other information indexes, the BLRT consistently performed the best. Based on the extensive literature on goodness-of-fit in CFA research, Marsh, Hau, and Wen (2004; Marsh, Hau, & Grayson, 2005) warned against the common practice of using goodness-of-fit indexes as "golden rules" that obviate the need for the researcher to make subjective evaluations of models based on parameter estimates in relation to substantive theory as well as indexes of fit. From this perspective, we recommend that researchers start with models that make sense in relation to theory, a priori predictions, and substantive findings, as well as goodness-of-fit indexes.

Inclusion of the Correlates

How should relations between grouping variables (based on the ASCs) and the correlates be represented? More specifically, should the correlates be included in the LPA model as covariates, predictors posited to be causally related to ASC groups? To what extent does inclusion of correlates as covariates influence the resulting groups and their relation with the correlates? Particularly recent research by Lubke and Muthén (2007) suggests that inclusion of covariates in the LPA can improve parameter coverage and classification accuracy. However, failure to include covariates should not result in a misspecified model, as the covariates are assumed to affect only the class probabilities. The marginal class proportions and profiles of the classes should therefore be unbiased by omission of the class predictors, even though classifications that include covariates should be more accurate. However, the classes should not qualitatively change due to the inclusion of covariates. Hence, if there is a qualitative change in the nature of the groups, then the assumption that the covariates affect only the latent class probabilities is violated and their inclusion might distort the interpretation of the groups. Thus, Lubke and Muthén (2005) cautioned that when considering models with covariates, "it is important to realize that the assigned class membership of a participant is model dependent and not an innate quality of the participant" (p. 37). This ambiguity in labeling groups is similar to problems identified by Howell, Breivik, and Wilcox (2007) in characterizing formative factors, and more generally what Anderson and Gerbing (1988) referred to as interpretational confounding when the interpretation of a latent variable is altered when the specification of the model is changed. Although we argue that the treatment of covariates should be based on the purpose of the study, we do contend that it is important to determine the extent to which groups are qualitatively altered by the inclusion of covariates in the model and explicate ways to evaluate this key assumption underlying LPA.

It is also important to emphasize that the inclusion of covariates into the LPA is based on the strong assumption that the causal ordering is from the covariates to the latent groups—that the arrows go from covariates to latent groups. In this sense, the covariates are assumed to be strictly antecedent variables—not concurrent or distal outcomes that are influenced by the groups or grouping variables. This is clearly reasonable for many applications in which covariates are preexisting background variables that necessarily precede indicators of the latent groups. In this investigation, however, there is ample evidence (e.g., Marsh, 2007b; Marsh & Craven, 2006) for reciprocal causal ordering of ASC and achievement—prior ASC (indicators of the groups in this investigation) influences subsequent achievement and prior achievement influences subsequent ASC. In this sense, it is appropriate to consider the correlates as auxiliary variables and to use them to validate the latent classes by examining the class-specific means and variances for the correlates without directly including them in the model (B. Muthén, personal communication, March 2, 2008) and this is facilitated by the auxiliary variable function recently introduced in M*plus* (version 5) as demonstrated here. Because the term *covariate* in M*plus* is given a special meaning as an antecedent variable (i.e., one that precedes the latent groups), we use the more general term of *correlate* to mean variables that are posited to be correlated with the latent groups without specifying a causal ordering. In this application it is reasonable to argue that inclusion of the correlates into the LPA might violate a critical assumption of the model (the causal ordering of the correlates and the groups), might alter the nature of the groups, and might confound relations between groups and the correlates that are intended to test the construct validity of the latent groups. Here we describe diagnostic approaches to test these suggestions and evaluate their implications.

In LPA it is typical to use multinomial logistic regression to relate the latent groups to the correlates. In this investigation, we also use discriminant function analysis and canonical correlation to relate the grouping variables to each other, to the ASC indicators on which they are based, and to the correlates used to validate the groups. Of particular relevance, the application of canonical correlation provides an index of variance explained that can be used to compare the traditional variable-centered and person-centered approaches. More specifically, we use canonical correlation to determine whether (and how much) information is lost in the reduction of the ASC scores into latent classes in terms of being able to explain variance in the set of correlates. In this way, we are able to compare the results across different analyses in terms of visual plots and an index of the total variance explained in ways that supplement and extend interpretations based on traditional applications of logistic regression.

METHOD

Sample

Data considered here are part of a large, ongoing German study, "Transformation of the Upper Secondary School System and Academic Careers" (Köller, Watermann, Trautwein, & Lüdtke, 2004). Students were from 149, randomly selected, upper secondary schools in one German state and were representative of upper secondary schools in this state. In each school, up to 40 students were randomly selected to participate in the study. Students in the sample used here ($N = 4{,}475$; 45% male) were in their final year of upper secondary schools (typically aged 17–19). Participation rate was 99% at the school level, and 80.2% at the student level.

Materials

SDQIII instrument. The SDQIII (Marsh, 1992b; Marsh & O'Neill, 1984; see Appendix) is a multidimensional self-concept instrument for late adolescents and young adults based on the Shavelson et al. (1976; Marsh, Byrne, & Shavelson, 1988) model. Previous research with the SDQ instrument (see Marsh, 1990, 1992b, 1993; Marsh & Craven, 1997) and reviews (see Boyle, 1994; Byrne, 1996; Hattie, 1992; Wylie, 1989) suggest that it is one of the strongest multidimensional self-concept instruments for this age group. This investigation is based on a German adaptation of the SDQIII (Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2006; Schwanzer, Trautwein, Lüdtke, & Sydow, 2005). In this short version of the SDQIII, four items per scale were chosen such that (a) items, particularly for the academic and performance-oriented scales, focused on competency; and (b) factor analyses in pilot research showed that the selected items had the highest factor loadings on their respective factor. This German adaptation includes five scales that are not part of the original instrument, but are deemed relevant to the transition from school to university or working life: intellectual (which replaced the general ASC scale on the SDQIII), artistic, political, technical, and computer self-concepts (see Appendix).

CFA of responses to these items on the German SDQIII, based on the data considered here, demonstrate that the a priori 17-factor solution fits the data well; all factor loadings were substantial; and the reliability of each of the scales is good (also see Marsh et al., 2006). In relation to the original Shavelson et al. (1976) model, these 17 self-concept factors can be divided broadly into a set of nine nonacademic factors from the SDQIII instrument and eight academic factors considered as part of this investigation—those from the SDQIII (verbal, math, problem solving), and those added in the German adaptation (intellectual, artistic, political, technical, computers) that place more emphasis on generic life skills (see Appendix).

Academic outcomes and other correlates. The mathematics achievement test consisted of original items from the Third International Mathematics and Science Study (e.g., Baumert, Bos, & Lehmann, 2000). The English achievement test was based on the Test of English as a Foreign Language, which was constructed by the Educational Testing Service (ETS, 1997). Additionally, as a third achievement indicator, the *Abiturgesamtnote* (an overall grade point average [GPA]) was obtained from school records and students reported their school grades in mathematics, English, and German from their most recent report card. Finally, students were asked whether they took an advanced course in mathematics, German, and English. In the upper secondary schools, all students had to select two advanced courses from a broad array of different course offerings. Although there were restrictions with regard to specific combinations of advanced courses, all students had the opportunity to select at least one of these three courses. Köller, Daniels, Schnabel, and Baumert (2000) demonstrated that course selection is based on a combination of academic achievement, self-concept, and interest in different school subjects.

Statistical Analyses

LPAs were conducted with M*plus* (version 5.0; for further discussion see Lubke & Muthén, 2007; Muthén, 2001; B. Muthén & Muthén, 2000; Pastor et al., 2007). Based on preliminary (exploratory) analyses described in more detail as part of the results, we chose a five-group solution that provided the best representation of our data in relation to a priori predictions that observed groups would represent a combination of level (i.e., high, medium, and low ASC averaged across the different scales) and shape (i.e., high scores on math-related ASCs and low score on verbal-related ASCs, or vice versa). We also considered a range of different goodness-of-fit indexes and tests of statistical significance that were used to aid in the determination of the number of groups: the difference in the LRT developed by Lo, Mendell, and Rubin (2001) and its alternative based on BLRT (McLachlan & Peel, 2000) and information criterion indexes—the Baysian Information Criterion (BIC) and Akaike's Information Criterion (e.g., Collins et al., 1993; Henson et al., 2007; Magidson & Vermunt, 2004), or the SSA-BIC (Yang, 2006).

Two different approaches to LPA were used. In LPA Model 1 (see Figure 1a) there were eight ASCs (dependent, *y*-variables) and five latent groups (i.e., classes = c(5) in M*plus* notation) based on the M*plus* "type = mixture" procedure. Because "c" (the grouping variable) is a latent categorical variable, the paths from "c" to the ASCs meant that the means of the ASCs are free to vary across the classes of "c." Within-class means and variances of the observed variables were estimated, but residual covariances between the indicators were fixed to zero—consistent with the assumption of local independence in the classical LPA. The default estimator is maximum likelihood with robust standard errors. As part of the analysis, we saved the five-category latent grouping variable and the corresponding set of five probabilities (the model estimated the probability that each individual falls into a given group) for each individual that were the basis of subsequent analyses. In supplemental analyses based on Model 1, we included the set of 10 correlates using the auxiliary variable approach in M*plus.* This approach allowed us to test the equality of the means of correlates across latent classes using a Wald chi-square test of statistical significance based on pseudo-class draws from posterior probabilities, a strategy that combines some of the advantages of relating the auxiliary variables to group dichotomies and group probabilities. In LPA Model 2 (see Figure 1b), we added the 10 correlates in the model. This set of correlates was regressed on the latent categorical grouping variable representing the

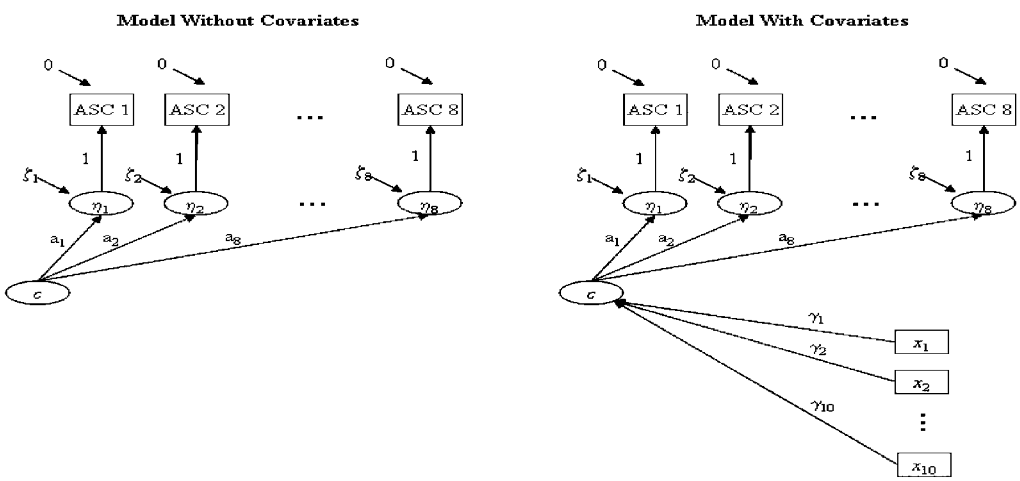

FIGURE 1 Schematic representation of the latent profile analysis models without covariates and with covariates. ASC 1 to ASC 8 denote the eight academic self-concepts and x_1 to x_{10} the covariates. η_1 to η_8 denote the latent constructs that are assessed by single indicators. The latent categorical variable c classifies persons in terms of their profile on the ASC 1 to ASC 8. Residual errors within classes are denoted by ζ_1 to ζ_8. The effects of the covariates on latent class membership are indicated by γ_1 to γ_8.

five groups. The paths from each correlate to the grouping variables represent a multinomial logistic regression.

To address the potential problem of local maxima, we instructed M*plus* to use 1,000 random sets of starting values for both Models 1 and 2. After 20 iterations—which are recommended for a thorough investigation (L. Muthén & Muthén, 2008)—the 100 best sets of starting values identified by the highest likelihood values were selected for final optimization. Although Models 1 and 2 converged to proper solutions, to ensure that this solution did not reflect a local maximum, both final models were reproduced by increasing to 5,000 the number of random starts, to 100 the number of iterations, and to 500 the number of final-stage optimizations (Hipp & Bauer, 2006). In both models, the obtained solution was clearly replicated and the log likelihood value was also replicated many times. We are thus reasonably confident that these solutions represent the best fitting solution, but it should be noted that this potential problem of local maxima represents a potentially serious problem for LPAs and mixture models more generally (see Hipp & Bauer, 2006).

In the classic LPA model as initially proposed by Lazarfeld and Henry (1968; also see Gibson, 1959) there is an assumption of local independence such that indicators within groups are uncorrelated and associations among indicators are explained in terms of the grouping variables (e.g., Bartholomew, 1987; McLachlan & Peel, 2000; Muthén, 2001; Uebersax, 1999). As stated by Muthén (2001):

> This model has features in common with factor analysis in that it assumes that latent variables explain the association between outcomes. This is also referred to as the conditional independence assumption, with the idea that if a sufficient number of classes is introduced, the independence is more and more likely to hold. (p. 3)

As emphasized by Muthén (2001), this assumption of conditional independence is the fundamental difference between LPA and more general mixture models that relax this constraint. Hence, although it is possible to relax this assumption, the solution is no longer a classic LPA solution. By analogy, in CFA, correlations among indicators are assumed to be explained in terms of latent factors. Although it is possible to relax this assumption of conditional independence by the inclusion of correlated uniquenesses (correlations among indicators not explained by factors), best practice (e.g., Marsh, 2007a) is not to do so except in special circumstances that are posited a priori (e.g., multiwave longitudinal data). Indeed, correlated uniquenesses actually represent apparently minor factors based on supposedly substantively unimportant features of the data that are “hidden” from further consideration when the factors are embedded in larger models and related to other constructs. Similarly, relaxation of the assumption of local independence in the classic LPA model can result in a smaller number of groups. We argue, however, that (a) such a model would no longer be a classical LPA solution; (b) this should only be done in special circumstances based on a priori assumptions; and (c) groups based on the assumption of local independence should first be validated in relation to appropriate correlates before researchers explore post-hoc relaxation of these constraints in the classic LPA model.

Although LPA using M*plus* was the primary analytic technique, we also used SPSS (version 14) to produce box plots (graphs showing 25th, 50th, and 75th percentiles for each of the LPA groups for selected variables), discriminant function analysis, and canonical correlation.

Discriminant function analysis (e.g., Tabachnick & Fidell, 2001) is typically used to determine which variables discriminate between two or more previously identified groups and how accurately individuals can be classified into groups on the basis of selected variables. In the analysis phase, a classification rule is developed using cases for which group membership is known. The grouping variable must be categorical, and the independent (predictor) variables must be interval or dichotomous. In the classification phase, the rule is used to classify cases for which group membership is not known. This can be done with a jackknife cross-classification scheme in which each case is classified by functions derived from all cases other than that case and then evaluated in relation to the known group membership. Here we used discriminant function analysis to evaluate the construct validity of the LPA groups. When applied to the ASCs on which the groups are based, the purpose of discriminant function analysis was to describe differences between the latent classes, not to validate the groups in relation to external criteria. When applied to the external validity criteria, the results provide an evaluation of the validity of the groups in relation to these external criteria. In each case, class sizes for the discriminant function analysis were the class sizes based on the LPA. However, to evaluate multivariate relations based on grouping variables represented by probabilities (e.g., the probability of each case of being member of each profile)—the main focus of our study—we used canonical correlation.

Canonical correlation analysis (e.g., Tabachnick & Fidell, 2001) is designed to investigate the relationship between any two sets of variables (e.g., ASCs and correlates, LPA grouping variables and correlates). Pairs of canonical variates are formed, one based on the first set of variables and one based on the second, such that the relation between the pair of canonical variates (the canonical correlation) is maximized. The maximum number of pairs of canonical variates is equal to the number of variables in the smallest sets (e.g., with eight ASCs and 10 correlates, the maximum number of canonical variate pairs is eight). However, the number of canonical variates can also be determined in terms of statistical significance or variance explained. For each pair of canonical variates there is an eigenvalue that reflects the size of the canonical correlation and the proportion of variance in the canonical variate explained by the canonical correlation. However, this does not say how much variance in the original sets of variables is explained. The redundancy analysis measures the percentage of variance each canonical variate is able to explain, in both the set of original variables on which it was based and also in the other set. Summing across each pair of canonical variates provides a measure of the effectiveness of the set of canonical variates in explaining variance in each set of original variables.

RESULTS: RELATIONS BETWEEN ASCS AND CORRELATES—A VARIABLE-CENTERED APPROACH

We began with an initial variable-centered evaluation of the relations between eight ASCs and the 10 correlates (Table 1). Consistent with our substantive orientation and a priori predictions, we also included two additional marker variables—the mean of the eight ASCs (representing ASC level) and the difference between MathSC and VerbSC (representing an a priori “shape” distinction).

TABLE 1
Relations Between the Original Eight Academic Self-Concept Scales and the 10 Correlates

	8 Academic Self-Concept (ASC) Scores								*10 Correlates (CORR)*										*Mean of*		*MATHSC − VERBSC*
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*Mult R^2*	*8 ASCs*	
8 academic self-concepts																					
1 MATHSC	1.0																		.60	.37	.79
2 VERBSC	−.24	1.0																	.41	.42	−.79
3 PROBSC	.13	.25	1.0																.04	.66	−.07
4 INTLSC	.29	.30	.45	1.0															.13	.69	−.01
5 ARTSSC	−.20	.29	.14	.04	1.0														.12	.31	−.31
6 POLTSC	.00	.28	.29	.30	.04	1.0													.21	.54	−.18
7 TECHSC	.25	−.14	.23	.19	.06	.05	1.0												.11	.49	.25
8 COMPSC	.20	−.10	.08	.12	−.16	.15	.29	1.0											.23	.41	.19
10 correlates																					
9 GENDER	−.14	.16	−.10	−.16	.23	−.31	−.25	−.44	1.0										.34	−.26	−.19
10 MATHACH	.56	−.13	.08	.24	−.16	.07	.21	.23	−.30	1.0									.34	.28	.44
11 ENGLACH	.01	.33	.07	.24	.05	.19	.01	.11	−.14	.35	1.0								.16	.26	−.20
12 MATHCRS	.49	−.33	−.01	.06	−.21	−.06	.17	.19	−.18	.56	−.09	1.0							.30	.08	.52
13 ENGLCRS	−.26	.19	−.05	−.01	.09	.01	−.10	−.04	.04	−.17	.39	−.43	1.0						.09	−.04	−.29
14 GERMCRS	−.25	.27	.04	−.00	.13	.08	−.15	−.15	.20	−.25	−.07	−.33	−.17	1.0					.12	−.01	−.33
15 MATHGRD	.66	−.08	.02	.19	−.12	−.02	.06	.05	.02	.45	.13	.23	−.11	−.14	1.0				.46	.19	.47
16 GERMGRD	.06	.46	.12	.20	.17	.19	−.10	−.17	.21	.16	.31	−.07	.06	.15	.30	1.0			.28	.24	−.25
17 ENGSGRD	.10	.40	.08	.25	.06	.16	−.09	−.02	.04	.27	.57	−.03	.13	.02	.35	.54	1.0		.22	.24	−.19
18 TOTGRD	.42	.21	.09	.27	.00	.17	−.03	−.06	.07	.48	.40	.16	−.06	−.04	.67	.63	.67	1.0	.32	.27	.13

Note. See Appendix for a description of the variables. Correlations greater than $r = |.03|$ are statistically significant ($p < .05$). Mult R^2 values are based on the extent to which each academic self-concept (ASC) can be predicted by the set of 10 correlates (CORR), or each CORR can be predicted by the set of 8 ASCs. MATHSC − VERBSC = the difference between math and verbal self-concept scores.

Relations Among ASCs

In evaluating these correlations, it is important to emphasize that all the scales are positively oriented (i.e., more positive responses reflect the high ASC end of the response scale). Correlations among the eight ASCs are modest—ranging from −.24 to .45—demonstrating the ASCs are very distinct (see Appendix for a description of the ASCs and the acronyms used here). There is a clear logical pattern of correlations. The most negative correlation was between MathSC and VerbSC, although VerbSC is also negatively correlated with CompSC and TechSC, and MathSC is also negatively correlated with ArtsSC. The most global scale is IntlSc, so it is logical that it is most highly correlated with other ASCs—particularly ProbSC, MathSC, VerbSC, and PoltSC—as well as the mean ASC. The other most positive correlations are between conceptually related ASCs (e.g., CompSC with TechSC, VerbSC with ArtsSC). A substantively important focus of this study is on the distinctiveness of the MathSC and VerbSC ($r = -.24$). The correlations between each ASC and the difference between the two ASCs (MathSC − VerbSC) provides a summary of how each ASC is related to this math–verbal continuum, ranging from MathSC (.79) to VerbSC (−.79). Two ASCs are more math-like (TechSC, .25; CompSC, .19). Two ASCs are more verbal-like (ArtsSC, −.31; PolitSC, −.18). Two ASCs fall near the middle of this continuum (ProbSC, −.07; Intel, −.01). In summary, these results provide strong preliminary support for the multidimensional perspective of self-concept, and the distinctiveness of the multiple dimensions of ASC—particularly MathSC and VerbSC.

Gender Differences in ASCs

Consistent with previous research (e.g., Marsh, 1989a, 1989b; Marsh & Yeung, 1998), girls typically had somewhat lower self-concepts than boys (Table 1). Thus, for example, girls had modestly lower scores on IntlSC (−.16), the most global of the ASCs, and for the mean of the ASCs (−.26). However, this overall value is deceptive, as correlations between gender and specific ASCs varied substantially and in a logical pattern. Also consistent with previous research showing stereotypical gender differences in specific components of self-concept (e.g., Marsh, 1989a), girls, compared to boys, had significantly higher VerbSCs and significantly lower MathSCs. Interestingly however, the largest gender differences are in the four new ASCs that were added to this German version of the SDQIII—one favoring girls (ArtsSC) and three favoring boys (PoltSC, TechSC, and particularly CompSC). Again, these contrasting gender differences support the multidimensional perspective of self-concept and the distinctiveness of the multiple dimensions of ASCs.

Univariate Relations Among ASCs and Achievement Indicators

There are large and systematic patterns of relations between the eight ASCs and 10 correlates (Table 1). Particularly dramatic are MathSC and VerbSC, which are explicitly referenced to school subjects. Thus, MathSC was substantially and positively related to math school grades (.66), math standardized achievement test scores (.56), taking advanced math courses (.49), and total GPA (.42) but was not highly related to English standardized test scores (.01) or school grades in German (.06) and English (.10), and was negatively correlated with taking advanced courses in English (−.26) and German (−.25). In marked contrast, VerbSC (German) was substantially and positively related to school grades in German (.46) and English (.40), English standardized achievement test scores (.33), taking advanced courses in English (.19) and German (.27), and total GPA (.21), but was negatively correlated with math standardized test scores (−.13), math school grades (−.08) and, particularly, taking advanced math courses (−.33). The pattern of relations with MathSC was also evident to a lesser extent for TechSC and CompSC, whereas the pattern of relations with VerbSC was also evident in ArtsSC and, perhaps, PoltSC. IntlSC and, to a lesser extent, ProbSC, were more global ASC scales and had moderately positive correlations with all four measures of school grades and both test scores. Interestingly, because of the ipsative nature of taking advanced courses (i.e., taking advanced courses in one area is negatively related to taking them in another area), IntlSC and ProbSC were not substantially related to any of these three course-selection variables.

Multivariate Relations Among ASCs and Correlates

We also addressed the issue of how much of the variance in each set of variables—the eight ASCs and the 10 correlates—could be explained by the other. Overall, 23.1% of the variance in the ASCs could be explained in terms of the set of 10 correlates. This figure is both the average of the Mult R^2 values in Table 1 and the cumulative variance explained in the ASCs by the canonical variates based on the set of correlates. However, the Mult R^2 values in Table 1 demonstrate that the variance explained varies substantially for the different ASCs. Given the nature of the correlates, it is not surprising that the VerbSC (41%) and particularly MathSC (60%) are substantially better predicted than any of the other ASCs and, perhaps, that ProbSC is least well predicted (4%).

Overall, 26.3% of the variance in the 10 correlates can be explained in terms of the set of eight ASCs—the average of the Mult R^2 in the correlates (Table 1; see also Table 2). However, again the variance explained in the different correlates varies substantially. Math correlates tend to be the best predicted, particularly math grades. English correlates tend to be the least well predicted, apparently due to the fact that there was not a specific ASC component for English (VerbSC refers to German). Also, the test scores and school grades tend to be better predicted than the course selection variables. Although six of the canonical correlations are statistically significant (Table 2), most of the variance can be explained by the first three pairs of canonical variates, with some additional variance attributable to the fourth and fifth.

A somewhat different picture emerges when evaluating how much of the variance in the ASCs is explained by differing numbers of canonical variates based on the ASCs themselves. Each pair of canonical variates based on the ASCs and the correlates is constructed to maximally predict variables in one set with variables in the other. Hence, it is also important to evaluate how much variance in the ASCs can be explained by the canonical variates that are actually based on the ASCs themselves. Whereas the first two of these canonical variates account for most of the explained variance in the correlates, they explain only 30.7% of the variance in the ASCs themselves. Appreciable amounts of variance in the ASC scores can be explained by the third (20.1%), fourth (8.9%), fifth (14.4%), sixth (6.4%), seventh (10.5%), and even the eighth (8.9%) canonical variates. This will typically be the case in a set of variables that are as highly differentiated (relatively uncorrelated) as the ASCs.

TABLE 2
Multivariate Relations Between the 8 Academic Self-Concept Scales and the 10 Correlates: Canonical Correlation Analysis and Redundancy Statistics

	%Var Explained in ASCs by Canon Variates Based on				*%Var Explained in CORR by Canon Variates Based on*				
	ASC Scales		*Criteria*		*ASC Scales*		*Criteria*		
Canonical Variate	*%*	*Cum%*	*%*	*Cum%*	*%*	*Cum%*	*%*	*Cum%*	*Canonical Correlation*
1	18.6	18.6	11.8	11.8	12.4	12.4	19.6	19.6	.798*
2	12.2	30.7	4.4	16.2	8.7	21.2	24.1	43.7	.602*
3	20.1	50.9	6.0	22.2	4.6	25.8	15.5	59.2	.543*
4	8.9	59.8	0.6	22.7	0.4	26.1	5.8	64.9	.252*
5	14.4	74.3	0.3	23.1	0.1	26.2	4.0	68.9	.155*
6	6.4	80.7	0.0	23.1	0.0	26.3	6.9	75.8	.077*
7	10.5	91.2	0.0	23.1	0.0	26.3	5.2	81.0	.041
8	8.8	100.0	0.0	23.1	0.0	26.3	7.8	88.7	.013

Note. Canonical correlation was used to relate the set of eight academic self-concept scales (ASCs) to the set of 10 correlates (see Table 1). Although it is possible to construct eight pairs of canonical variates—one based on ASCs and one based on correlates—correlations between the two are only significant for the first six. Of particular relevance is the cumulative percentage of variance in the correlates explained by the canonical variates based on the ASCs and the cumulative percentage of variance in the correlates explained by the canonical variate based on the ASCs.
*$p < .01$.

In summary, much variance in the correlates (26.3%) can be explained by the set of ASCs and much variance in the ASCs (23.1%) can be explained by the correlates. However, even though as few as two canonical variates based on the ASCs can explain most of the explicable variance in the correlates, these two canonical variates are able to explain only a small portion of the variance in the ASCs. Furthermore, these summary statistics do not do a very good job of capturing the highly differentiated pattern of correlations between specific components of ASC and specific correlates presented in Table 1.

RESULTS: LATENT PROFILE ANALYSIS OF ASCS—A PERSON-CENTERED APPROACH

Now we consider an alternative perspective—an LPA that focuses on groups of individuals with similar profiles rather than relations among variables. In demonstrating the application of LPA, we focus on particular issues faced by applied LPA researchers: number of groups and goodness of fit, alternative ways of representing the data, and validating the grouping variables in relation to correlates.

Number of Groups and Goodness of Fit

Sometimes researchers have a well-established, a priori hypothesis as to the number of groups used to represent the ASCs. However, typically researchers have only a limited basis for predicting the nature and number of groups. Here, for example, we posit that groups should be defined by a combination of quantitative differences in overall ASC (e.g., high, medium, and low mean ASC averaged across the eight ASCs) and qualitative differences among the various ASC scales (e.g., high MathSC and low VerbSC, or vice versa). Consistent with typical practice, we explored solutions with varying numbers of groups and selected the one that makes the most sense in relation to theory, previous research, the nature of the groups, and interpretability. On this basis we selected a five-group solution as best representing our data.

We also evaluated models positing between one and eight groups in relation to five indexes of fit commonly used in this area (Table 3). For the three information indexes (AIC, BIC, and the SSA–BIC), the values continued to decrease across the range of models considered, suggesting that we should consider at least eight groups. This is not surprising given the large sample size and sample size dependency of these measures. None of the eight models resulted in groups with less than 1% of the cases, whereas models positing more than five groups each resulted in at least one group with less than 5% of the cases. Finally, although results based on the LMR and BLRT tests were nearly identical to each other, the results were not entirely consistent in terms of choosing the appropriate number of groups. Both LMR and BLRT were highly significant for the two- and three-group solutions ($p \leq .01$), whereas values for the four-group solution were marginally nonsignificant. Although this might suggest a three-group solution, the values for the five-group solution were also statistically significant, and the values for the six-group solution were not, suggesting that the five-group solution was best. However, both LMR and BLRT values were marginally significant ($p \leq .05$) for the seven-group solution, even though the seven-group solution was not as interpretable as the five-group solution.

In summary, goodness-of-fit indexes, tests of statistical significance, and so-called "golden rules" for evaluating the number of groups were not consistent or particularly useful in this investigation. Based on our evaluation of goodness of fit, we concluded that the results were not inconsistent with our decision—based on a subjective evaluation of the results of alternative

TABLE 3
Goodness of Fit for Models Based on Different Numbers of Groups

						Group Sizes	
No. Group	*No. Parm*	*AIC*	*BIC*	*SSA–BIC*	*p LMR (BLRT)*	*LT1%*	*LT5%*
1	16	101,473	101,576	101,525	—	0	0
2	33	99,236	99,448	99,343	.0000 (.0000)	0	0
3	50	97,943	98,263	98,104	.0003 (.0003)	0	0
4	67	97,132	97,561	97,348	.0865 (.0854)	0	0
5	84	96,473	97,011	96,744	.0096 (.0094)	0	0
6	101	96,070	96,717	96,396	.2051 (.2033)	0	1
7	118	95,703	96,459	96,084	.0274 (.0269)	0	1
8	135	95,527	96,392	95,963	.2917 (.2894)	0	2

Note. AIC = Akaike's Information Criterion; BIC = Bayesian Information Criterion; SSA–BIC = sample-size-adjusted Bayesian Information Criteria; p LMR (BLRT) = p values for the Lo–Mendell–Rubin likelihood and the bootstrap likelihood ratio test for K versus $K - 1$ classes. Group sizes refer to the number of groups with less than 1% and less than 5% of the cases. $N = 4{,}475$.

models in relation to a priori hypotheses—to consider five groups. (We subsequently return to this issue in a discussion of ongoing issues in the application of LPA.)

Latent Profile Analysis of the Eight ASCs

In Model 1, only the eight ASCs were used to form the five ASC–LPA groups. From this analysis, we constructed sets of the five group probabilities (i.e., the probability that a given individual is from each of the five groups). Because the group probabilities are ipsative, the average correlation among the group probabilities is negative (Table 4, Model 1). Although not a focus of this investigation, correlations between the matching groups dichotomies (not presented) and corresponding group probabilities were very high, ranging from .92 to .96.

The nature of the five groups based on Model 1 is illustrated by their ASC profiles (Figure 2a), represented as box plot graphs (in which the 25th, 50th, and 75th percentiles are represented by the top of the box, the horizontal line in the middle of the box, and the bottom of the box, respectively). As expected, the groups reflected combinations of quantitative differences in the level of ASC (relatively higher or lower values of mean ASC) and qualitative differences in the ASC scales (although not used in the LPA, we also included the mean ASC and the difference between MathSC and VerbSC as marker variables to clarify the nature of each profile in relation to a priori predictions). Thus Group 1 is characterized primarily by low ASC (i.e., the box for mean ASC falls completely below the horizontal reference line representing the mean self-concept). However, individuals in Group 1 also have relatively higher VerbSCs than MathSCs (although there was substantial overlap between the two boxes and the MathSC–VerbSC difference box extended above the horizontal reference line). Group 2 reflects primarily an average level of ASC and similar MathSC and VerbSC scores (although MathSC is somewhat higher than VerbSC). Group 3 reflects primarily large MSC–VSC differences (MathSC scores are much lower than VerbSC, and the difference box was well below the horizontal reference line) although the mean ASC is somewhat above average. Also, in Group 3 ProbSC, IntlSC, and ArtsSC are relatively higher than TechSC and CompSC. Group 4 represents a combination of high ASCs overall and large MathSC–VerbSC differences (MathSC > VerbSC). However, in Group 4, IntlSC is highest overall, whereas ArtsSC is lowest. Group 5 represented primarily a MathSC–VerbSC difference (MathSCs > VerbSCs). However, in Group 5 ProbSC, IntlSC, ArtsSC, and PoltSC are relatively lower, whereas TechSC and CompSC are relatively higher. In this respect, profiles in Groups 3 and 5 represent opposite patterns.

We also used a discriminant function analysis to depict the profile differences (see Figure 2b). The discriminant analysis based on the eight ASCs was able to correctly classify 83.4% of the individual students into their appropriate group (based on the five dichotomous LPA grouping variables from Model 1). All four of the discriminant functions were statistically significant (because the five groups are ipsative, only four discriminant functions are possible). However, most of the explicable variance could be explained by the first (58.5%) and second (40.1%) discriminant functions that are presented in Figure 2b. Inspection of group centroids (the black boxes in Figure 2b) for the two functions clearly demonstrates that the first (x-axis) represents ASC level, whereas the second axis represents MathSC–VerbSC (in which higher VerbSC values are above the horizontal reference line and higher MathSCs values are below the horizontal reference line). Thus, for example, the most extreme groups along the x-axis are Groups 1 (low ASC) and 4 (high ASC), and Group 2 is almost exactly at the intersection of the two axes

TABLE 4
Relations Between the Original Variables (Eight Academic Self-Concept Scales and 10 Correlates) and the Sets of Five Grouping Variables (Group Probabilities and Group Dichotomies Based on Model 1 and Model 2)

	8 ASC Scores								*10 Correlates*										*Model 1 Groups*					*Model 2 Groups*				
	Math	*Verb*	*Prob*	*Intl*	*Arts*	*Polt*	*Tech*	*Comp*	*Sex*	*MAch*	*EAch*	*MCrs*	*ECrs*	*GCrs*	*MGrd*	*GGrd*	*EGrd*	*TGPA*	*1*	*2*	*3*	*4*	*5*	*1*	*2*	*3*	*4*	*5*
Model 1: Groups based on ASC only																												
1	−.46	−.11	−.49	−.68	.04	−.37	−.38	−.33	.27	−.32	−.16	−.20	.10	.07	−.25	−.11	−.15	−.24	1.0									
2	.03	−.05	−.06	−.03	−.06	.01	.02	.00	.02	−.01	−.03	.02	.01	−.04	.01	−.02	−.02	−.02	−.25	1.0								
3	−.32	.62	.37	.40	.34	.38	−.06	−.07	.04	−.18	.22	−.32	.17	.25	−.19	.27	.22	.06	−.34	−.27	1.0							
4	.48	.13	.37	.47	−.04	.22	.22	.19	−.16	.35	.15	.23	−.13	−.11	.32	.17	.18	.30	−.28	−.19	−.15	1.0						
5	.45	−.64	−.12	−.06	−.34	−.21	.29	.29	−.23	.28	−.16	.37	−.20	−.24	.22	−.30	−.21	−.03	−.27	−.15	−.39	−.15	1.0					
Model 2: Groups based on ASC + correlates																												
6	−.71	.17	−.17	−.35	.16	−.12	−.25	−.24	.25	−.49	−.01	−.44	.23	.21	−.55	−.04	−.07	−.33	.56	−.09	.04	−.23	−.39	1.0				
7	.20	−.66	−.32	−.36	−.22	−.31	.04	.06	−.08	.11	−.37	.36	−.20	−.23	.09	−.42	−.38	−.20	.14	−.16	−.40	−.21	.61	−.33	1.0			
8	.10	.00	−.02	.06	−.03	.02	.04	.00	−.00	.00	.03	−.02	.04	−.05	.11	.03	.03	.03	−.18	.57	−.08	−.20	−.05	−.31	−.23	1.0		
9	−.14	.60	.30	.36	.24	.34	−.07	−.06	.05	−.06	.25	−.25	.12	.27	−.04	.37	.31	.21	−.27	−.24	.73	−.02	−.30	−.16	−.27	−.27	1.0	
10	.62	−.05	.28	.39	−.14	.13	.27	.27	−.24	.50	.14	.38	−.20	−.20	.45	.11	.17	.37	−.33	−.14	−.22	.75	.13	−.29	−.19	−.26	−.17	1.0

Note. ASC = academic self-concepts (see Appendix). Groups are group probabilities for five groups in each analysis. Values shaded in gray are relations between the most closely matching groups based on Model 1 and Model 2.

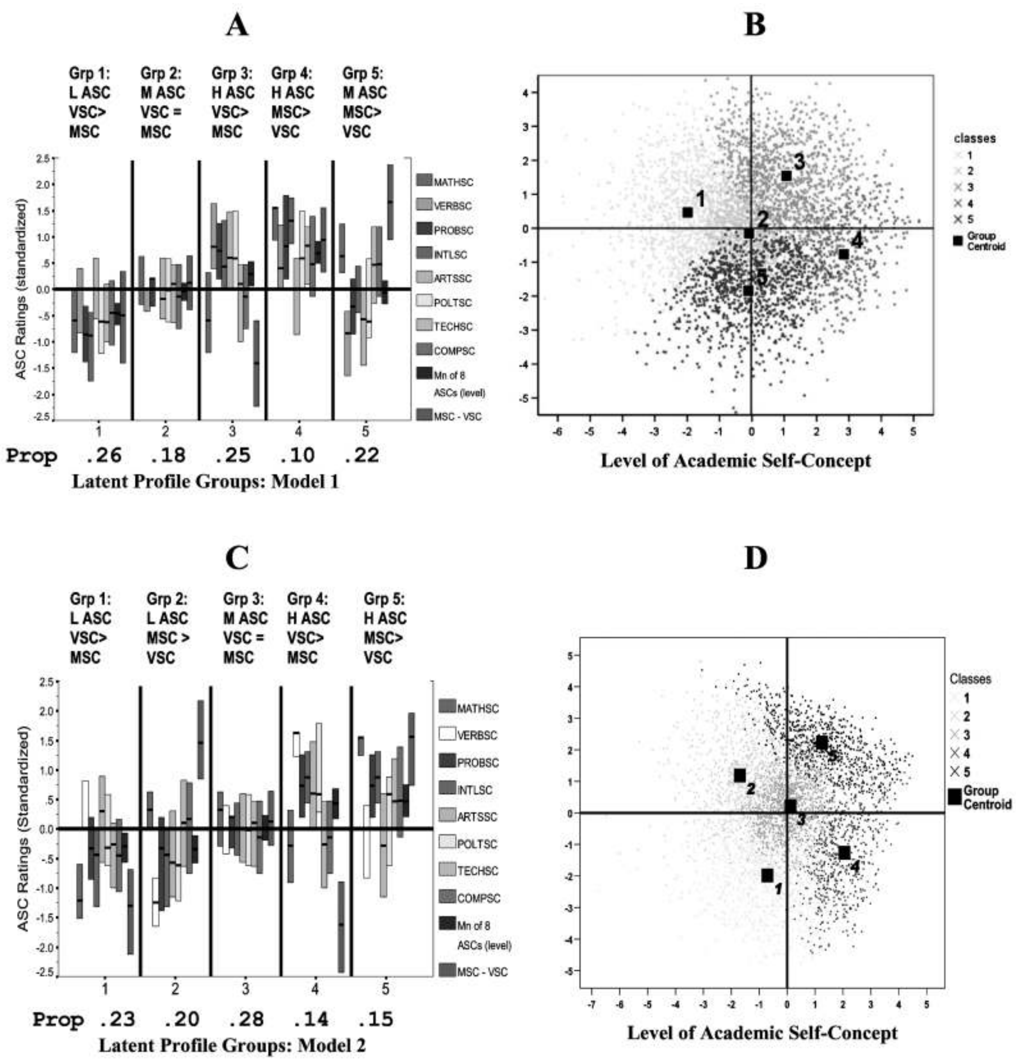

FIGURE 2 A and C are box-plots illustrating the profiles of academic self-concept scores (ASCs) based on latent profiles based on only ASCs (A, Model 1) and ASCs plus covariates (C, Model 2). The order of the boxes and the self-concepts represented by each represents is as follows: Verbal (VerbSC); Math (MathSC); Problem Solving (ProbSC); Intellectual (IntlSC); Artistic (ArtsSC); Political (PoltSC); Technical (TechSC); Computers (CompSC); Mean of all 8 academic self-concept scales (Mn of 8 ASCs, an indicator over overall self-concept level); the difference between Math and Verbal self-concepts (MSC—VSC, an aspect of shape). B and D are the corresponding results of a discriminant function analysis used to depict profile differences in relation to two discriminant functions representing ASC level (x-axis) and the difference between math self-concept and verbal self-concept (y-axis) for Model 1 (B) and Model 2 (B). Each set of profiles (Grp1 − Grp5) is characterized in terms of ASC level (L = low, M = medium, H = high) and Shape relative to differences between math self-concept (MSC) and verbal self-concept (VSC). Prop = proportion of cases in each group.

(average ASC, with little difference between MathSC and VerbSC). On the y-axis, the most extreme groups are Groups 3 (VerbSC > MathSC) and 5 (MathSC > VerbSC). These results provide strong support for a priori predictions about the nature of the LPA grouping variables.

In summary, consistent with a priori predictions, the five LPA groups are well represented by a combination of level (high vs. low ASC) and shape (MathSC vs. VerbSC). This finding provides preliminary support for the construct validity of the five groups derived from the LPA. Particularly useful were the inclusion of marker variables based on our a priori predictions and the discriminant function analysis showing that most of the explicable variance could be explained in terms of two discriminant functions that are clearly consistent with a priori predictions. We now turn to a more demanding test of the construct validity, an evaluation of relations between the grouping variables and the set of 10 correlates.

Validating the Latent Profile Analysis in Relation to External Validity Criteria

In evaluating relations between the ASC–LPA variables (group probabilities based on the LPA that included only ASCs) and the set of correlates, we begin by correlating the two sets of variables (Table 4). In support of the construct validity of the LPA groups, there is a clear pattern of relations between the five ASC–LPA group profiles and the set of 10 correlates. For example, Group 1 was previously characterized as having the lowest ASCs overall, but also having VerbSCs that were somewhat higher than MathSCs (see Figure 2a and 2b). Consistent with this characterization of the ASC–LPA, Group 1 has mostly low scores across the set of correlates (particularly for total GPA, math grades, math achievement, and math coursework selection) but has somewhat above-average scores for coursework selection in English and German. Group 1 is also made up of predominantly females, consistent with the gender differences in ASC overall and particularly gender differences in MathSc and VerbSc.

In contrast to Group 1, Group 5 (medium ASC, MathSC > VerbSC) has a mixture of positive and negative scores on the correlates, but almost no correlation with total GPA. Consistent with the characterization of Group 5, these students have high math-related correlates (coursework selection, test scores, and grades) but low verbal-related correlates (German coursework selection and grades and, to a lesser extent, English grades, achievement, and coursework selection).

Group 3 and particularly Group 4 have high ASCs overall, but Group 3 has much higher VerbSCs than MathSCs, whereas Group 4 has higher MathSCs than VerbSCs. Again these characterizations of the LPA group profiles are reflected in group differences in the correlates. Although Group 3, and especially Group 4, tend to have high scores for all the academic correlates, Group 4 has substantially higher scores for math achievement, grades, and coursework selection, whereas Group 3 has higher scores for the German- and English-related correlates.

Group 2 (medium ASC, MathSC = VerbSC) is particularly interesting. It might be reasonable, based on the ASC scores alone, to posit that these students were undifferentiated, perhaps, because they did not carefully complete the ASC instrument. However, this same group of students was also remarkably undifferentiated in terms of the set of correlates (all correlations varied between $\pm.04$). Hence, their undifferentiated ASC profile (in terms of both level and shape) was closely paralleled by the undifferentiated pattern of criterion measure scores.

Validating LPAs: Discriminant Function, Canonical Correlations, and Auxiliary Variable Analyses

Discriminant functions. In this discriminant function analysis (Figure 2d), we related ASC-LPA grouping variables (dichotomies) to the set of 10 correlates. All four discriminant functions were statistically significant, but almost all (97.9%) of the explicable variance could be explained in terms of only two functions. In this analysis, the percentage of correct classification based on four functions (46.8%) was only marginally better than that based on two functions (45.3%). As the N is very large, the cross-validation percentages were nearly as large (45.3% and 44.4%). In summary, almost half of the individuals can be correctly classified into five ASC–LPA groups based on their scores on the correlates.

Canonical correlations. In the canonical correlation analysis, we compare results based on the five ASC–LPA groups (see Table 5) and the original eight ASCs on which the groups were based (see earlier discussion of Table 2). The canonical variates based on the five ASC–LPA group probabilities were able to explain 18.6% of the variance in the set of 10 correlates, whereas canonical variates based on the 10 correlates were able to explain 14.6% of the variance in five ASC–LPA groups. However, these values were still substantially below what could be explained by the original eight ASC variables (23.1% and 26.3% respectively, see Table 2). In summary, five ASC–LPA grouping variables were able to explain substantially less variance than the eight ASCs on which they were based.

Auxiliary variable analysis. In version 5.0 of M*plus*, it is also possible to relate correlates to grouping variables by treating them as auxiliary variables that are not included in the actual LPA model. This method—the auxiliary (e) function from M*plus*—provides a test of statistical significance for the hypothesis of the equality of the means of each criterion variable across five latent groups (see Table 6). This approach is based on pseudo-class draws that take into account the probabilities that a particular case will fall into a particular group, thus avoiding the issue of only considering the most likely discrete class assignment (L. Muthén & Muthén, 2008). In this respect, it differs from relations between correlates based on observed group dichotomies and might represent a compromise between group dichotomies and group probabilities. The pattern of group differences (Table 6) based on the auxiliary variable approach is nearly identical to those based on group probabilities. Although the auxiliary variable approach also provides an omnibus test of the group differences on each criterion variable considered separately, all tests are highly significant for the large sample size considered here (as, indeed, are all tests of statistical significance reported in this investigation).

Latent Profile Analysis of the Eight ASC Scales and 10 Correlates (ASC + CORR)

An important, unresolved issue in LPA is whether correlates should be included in the LPA used to define the LPA groups (i.e., treated as covariates in M*plus*). In this investigation, we addressed this issue by comparing the results presented thus far (ASC–LPA groups in Model 1) with further analyses in which correlates were included in the analysis as covariates (ASC + CORR–LPA groups in Model 2). Although clearly recognizing that this is a complex issue, our

TABLE 5
Results of Five Canonical Correlation Analyses Relating Each Set of the Five Latent Profile Grouping Variables to Either the Eight Academic Self-Concept Scales or the 10 Correlates

Canonical Corr Analysis	*Cum % Var Explain by 4 Canonical Vars Based on*		*Cum % Var Explain by 4 Canonical Vars Based on*	
Model 1	8 ASCs in		5 Groups in	
Group (ASC Only)	8 ASCs	5 Groups	8 ASCs	5 Groups
With 8 ASCs	62.9	37.6	44.5	100.0
Model 1	10 Corr in		5 Groups in	
Group (ASC Only)	10 Corr	5 Groups	10 Corr	5 Groups
With 10 Corr	63.3	14.6	18.6	100.0
Model 2	8 ASCs in		5 Groups in	
Group (ASC & Corr)	8 ASCs	5 Groups	8 ASCs	5 Groups
With 8 ASCs	60.2	32.1	39.1	100.0
Model 2	10 Corr in		5 Groups in	
Group (ASC & Corr)	10 Corr	5 Groups	10 Corr	5 Groups
With 10 Corr	62.4	24.6	29.8	100.0
Models 1 & 2	5 ASC Groups in		5 ASC + CORR Groups in	
ASC groups with	5 ASC	5 ASC + CORR	5 ASC	5 ASC + CORR
ASC + Corr groups	100.0	49.0	43.7	100.0

Note. In each of the five canonical correlation analyses, a set of five grouping variables (group probabilities) was related to either a set of eight academic self-concept scales (ASCs), a set of 10 correlates (CORR), or another set of grouping variables. Separate analyses were conducted on grouping scores based on only the eight ASCs or based on the eight ASCs and 10 correlates. Each canonical correlation analysis resulted in one set of four canonical variates based on the grouping variables and a second set of four canonical variates based on the eight ASCs, the 10 correlates, or on another set of grouping variables (the maximum number of canonical variates is four rather than five because all sets of grouping variables are ipsative). Presented are the cumulative variance percentages explained by these two sets of canonical variates. Of interest is the variance explained by each set of canonical variates for their own variables (e.g., variance in eight ASCs explained by the set of four canonical variates based on the ASCs; this will be less than 100% when the number of variables is greater than the number of variates used to represent them). Of particular relevance is the variance in one set of scores explained by the canonical correlates based on the other set of scores (e.g., variance in 10 correlates explained by the five canonical variates based on the ASC LPA groups) that are shaded in gray.

intent initially is to simply compare the results based on the two LPA models to determine the extent to which the qualitative nature of the groups is changed by the inclusion of correlates.

ASC + CORR latent profiles: Relations with ASCs and correlates. Simple correlations for LPA grouping variables based on Model 1 (ASC–LPA, ASCs only) and Model 2 (ASC + CORR–LPA, ASCs and correlates) clearly demonstrate that the nature of the latent classes has changed (see Table 4). In a trivial sense, the order of the LPA groups has changed. Thus, for example, ASC–LPA Group 2 based on Model 1 is most highly correlated with ASC +

TABLE 6
Relations of the Five Latent Classes (From Model 1) to the Correlates (Auxiliary) Variables

	Class 1	*Class 2*	*Class 3*	*Class 4*	*Class 5*	*Chi-Square*
GENDER	0.360	0.040	0.066	−0.436	−0.350	422.010*
MATHACH	−0.417	−0.029	−0.258	−0.936	0.412	969.630*
ENGLACH	−0.210	−0.069	0.324	0.391	−0.229	304.536*
MATHCRS	−0.257	0.041	−0.458	−0.613	0.532	937.251*
ENGLCRS	0.127	0.018	0.240	−0.346	−0.290	261.052*
GERMCRS	0.098	−0.069	0.356	−0.305	−0.341	394.782*
MATHGRD	−0.327	0.019	−0.275	−0.860	0.305	900.350*
GERMGRD	−0.150	−0.032	0.379	0.447	−0.434	513.526*
ENGSGRD	−0.196	−0.036	0.315	0.489	−0.300	368.591*
TOTGRD	−0.342	−0.068	0.063	0.799	−0.073	478.454*

Note. Model 1 was based on eight grouping variable indicators (academic self-concept scores) and 10 correlates that were treated as auxiliary variables of the latent groups. For each auxiliary variable, the mean is presented for each latent class and a test of the significance of the of group differences across the five latent groups.

$^{*}p < .001$ ($df = 4$).

CORR–LPA Group 5 for Model 2 ($r = .61$). More important, sizes of the correlations for the most closely matching groups based on the two models clearly do not approach 1.0 (.56–.75). Hence, LPA groups based on the two models are clearly measuring something different.

The discriminant function analyses based on the two models (Table 3) provide further insight into these differences. Correct classification into LPA groups based on the original eight ASCs is only marginally poorer for Model 2 (ASC + CORR, 82.3%) than for Model 1 (ASCs only, 83.4%). However, correct classification into LPA groups based on the original 10 correlates is substantially worse for Model 1 (46.6%) than Model 2 (62.3%), which incorporates the correlates into the construction of the LPAs. Hence, although clearly reflecting differences in LPA ASC profiles identified in Model 1, ASC + CORR LPA profiles in Model 2 are much more closely aligned to the set of correlates than the ASC–LPA profiles.

The group differences are also clearly demonstrated by a comparison of the box plots and discriminant functions plots for the first two disciminant functions (Figure 2). For example, even though ASC–LPA Group 1 in Model 1 is most closely matched to ASC + CORR–LPA Group 1 in Model 2 based on correlations in Table 4, the profiles for the two groups on the ASCs are quite different. In Model 1, Group 1 is characterized by small differences in MathSC and VerbSC but by substantially low ASCs overall. In Model 2, Group 1 is characterized by slightly below-average overall ASC but substantially higher VerbSC than MathSC. Similarly, ASC–LPA Group 5 (Model 1) is most closely matched to ASC + CORR–LPA Group 2 (Model 2) based on correlations, but the profiles are qualitatively different. Group 5 in Model 1 is close to the mean of overall ASC and substantially in the direction of MathSC on the MathSC–VerbSC dimension. Although Group 2 in Model 1 also has higher MathSCs than VerbSCs, the overall mean ASC is below average. Based on a visual comparison of the two sets of figures, the most closely matching profiles are for ASC–LPA Group 2 (Model 1) and ASC + CORR–LPA Group 3 (Model 2). Although the correlations suggest that these are matching groups, the size of the correlation is not as high, for example, as those relating Groups 3 and 4 from Model 1 to Groups 4 and 5 from Model 2 and similar in size to those already considered (correlations relating Groups 1 and 5 from Model 1 to Groups 1 and 2 from Model 2).

ASC + CORR and ASC latent profiles: Canonical correlations. We now have two sets of LPA group probabilities, one based on ASCs only (ASC–LPA) and one based on a combination of ASCs and correlates (ASC + CORR–LPA). How much of the variance in one set of LPAs can be explained by the other? In pursuing this question, we conducted an additional canonical correlation analysis (see Table 5). Each set of LPAs is able to explain less than half of the variance in the other. The set of five ASC–LPAs explains 49% of the variance in the set of five ASC + CORR–LPAs, whereas the set of five ASC + CORR–LPAs explains 43.7% of the variance in the set of five ASC–LPAs. These analyses demonstrate that the LPA–ASC groups are quite different from those based on the ASC + CORR–LPA groups.

In summary, although there are apparent similarities in the set of LPA groups based on Models 1 and 2, there are also important differences. Quantitatively, the ASC + CORR–LPA groups based on Model 2 are much more closely aligned to the correlates. Furthermore, even though successful classification into groups based on ASCs is similar for Models 1 and 2, there are qualitative differences in the nature of ASC profiles based on the two models. In conclusion, the inclusion of the correlates in the LPA quantitatively and qualitatively altered the nature of the observed groups, violating the assumption that it should not do so. However, even though this problem is relevant to all LPA studies, in this investigation it is exacerbated by the observation that the covariates (correlates) are not antecedents to the latent classes. Indeed, the observed differences between the two models reflect this influence as the resulting classes are more closely related to the correlates in Model 2 than Model 1. Fortunately, the various new procedures presented here provide a stronger basis for evaluating the implications of this violation and alternate methods for investigating the external validity of latent profiles.

DISCUSSION AND IMPLICATIONS

Distinction Between Quantitative (Level) and Qualitative (Shape) Differences in LPA Profiles

This investigation is a uniquely appropriate application of LPA because—consistent with a priori predictions—the different profiles reflect a combination of quantitative differences in overall level of ASCs and qualitative differences in the shape of the profiles (i.e., some with MathSC > VerbSc and some with VerbSC > MathSC). Historically, this distinction has been a critical issue in the evaluation of profiles based on a traditional analysis of variance (ANOVA) approach (e.g., Tabachnick & Fidell, 2001) in which the "level" is just the effect of the grand mean averaged over all components (e.g., overall ASC), whereas the "shape" is the extent to which there are distinct profile differences. Although clearly useful, the focus of the ANOVA approach and related techniques is to identify sources of variance and not to identify the actual profiles or the individuals who fit within each profile group.

Although the overall level effect is important, in many cases it should not be a primary focus of LPAs and might even be of limited interest in some applications. This is particularly the case when LPA groups reflect primarily level differences that can be more effectively explained in relation to the mean levels of the original set of variables (or a higher order factor that represents them) than a small number of grouping variables in which a reasonably continuous score (the mean of the original variables or latent factor appropriately constructed factor score) is divided into a small number of discrete categories. However, even when the LPA indicators measure primarily one underlying trait, LPAs might be useful in providing cutoff values and prevalence rates for different categories—particularly when the focus of the research is on diagnostic categories, as in some mental health studies (see Meehl, 1995; Muthén, 2006). When the LPA groups reflect a combination of level and shape differences, as in this investigation, LPAs provide a potentially important, alternative perspective to the traditional variable-centered approaches. Indeed, as emphasized by Muthén (2006) and in this investigation, the use of person- and variable-centered approaches should be seen as complementary rather than competing approaches.

Number of Groups: Goodness of Fit and Appropriate Interpretations

Based on initial analyses, we selected a five-group solution as being most reasonable in relation to theory, previous research, the nature of the groups, and interpretability. This approach, reminiscent of strategies used in exploratory factor analysis to determine the number of factors, is highly subjective and more appropriate to an exploratory analysis than a confirmatory one. Nevertheless, here we found the supposedly less subjective approaches based on fit indexes and significance tests to be of limited use. This apparent problem is understandable in relation to the broader research literature on goodness of fit. Although there is a well-developed literature on the use of goodness-of-fit indexes in CFA and SEM research (e.g., Bentler, 1990; Marsh, Balla, & McDonald, 1988; Marsh et al., 2005), researchers pursuing latent mixture models have taken a very different direction. In applied CFA and SEM research, there is a predominant focus on indexes that are sample-size independent (e.g., root mean squared error of approximation, Tucker–Lewis Index, comparative fit index; Marsh, 2007a; Marsh, Balla, & McDonald, 1988; Marsh et al., 2005). In particular, the information criterion indexes like those emphasized in LPA research are not given much attention because they are so sample-size dependent. However, as noted in the Marsh et al. (2005) review, sample-size dependency is not an appropriate criticism of the information criterion indexes in that these indexes are explicitly intended to be highly sample-size dependent in that there is more information available when sample sizes are larger. From a statistical perspective, this is justified because it is appropriate to consider more complex models (i.e., LPA models with more groups) when the sample size is larger—there is less danger in capitalizing on chance. Nevertheless, for real data based on increasingly large sample sizes, the information criteria will always choose the most complex, and ultimately, the saturated model, as evident in this investigation.

There is a growing literature on mixture models based on data simulated from a "true model" in which the "right" number of groups is known (e.g., Collins et al., 1993; Henson et al., 2007; Magidson & Vermunt, 2004; Nylund et al., 2007; Tofighi & Enders, 2007; Yang, 2006). This literature seems to have converged on the use of SSA–BIC as the best of the information criterion indexes and of the use of statistical tests of significance, particularly the BLRT, as the best guides to the right number of groups. Nevertheless, this research might have limited generalizability for applied researchers who seek to test models with real data that are only intended to approximate reality (see Marsh et al., 2005) rather than simulated data based on a known true population model used to generate the data. This is particularly the case for complex data where LPA groups represent a combination of differences in level and shape. For real data it will generally be the case that measures of statistical significance (based on LRTs) and information criterion tests will be heavily dependent on sample size such that small sample sizes will result in fewer groups and larger sample sizes will result in more groups. Although this situation might be appropriate from a statistical perspective, it calls into question the assumption that there is one and only one "right" number of groups and that this right number of groups can be identified by indexes typically used in LPA studies. Although LPA and mixture model research have given some attention to the issue of sample size dependency (e.g., Henson et al., 2007), this issue has been largely ignored. To the extent that an applied researcher believes that there is one and only right number of groups, it might not make a lot of sense to use sample-size-dependent indexes to determine the number of groups. This is relevant in this investigation in which our sample size ($N = 4{,}475$) is reasonably large. Here we found information criterion indexes to be of limited usefulness. Although the two tests of statistical significance gave nearly identical results (i.e., p values were nearly the same) in this investigation, there was also a degree of inconsistency in their interpretation.

As demonstrated here, the selection of the "right" number of groups cannot be based on a mechanical application of recommendations about fit indexes. As argued by Marsh et al. (2004) in another context, there are no "golden rules" that will allow researchers to select the best model—there is always a degree of subjectivity and need for informed, professional judgment. Indeed, it is unclear whether there is even an internally consistent rationale for the application of these indexes in LPA and, more generally, mixture models. To the extent that more complex models are more appropriate when sample size is large, then it might be appropriate to use information indexes and tests of significance to select the most appropriate number of groups. However, this strategy means that the "right" number of groups will vary with sample size. To the extent that there is a posited, absolute right number of groups, then information indexes and statistical tests of significance might not be appropriate. Evidently, this is an area in need of further research and more guidance for the applied researcher.

The potentially counterproductive practice of arguing for a single right number of groups is also evident when the underlying groups represent primarily levels along a single underlying dimension—as is typical in LPA studies. As argued earlier, summarizing a reasonably continuous score in terms of a finite number of discrete grouping variables is likely to involve the loss of potentially useful information. The loss of information will vary inversely with the number of groups, whereas the statistical significance of the information lost will vary directly with sample sizes. For small sample sizes, a modest number of groups might be able to explain a sufficient amount of variance so that the information lost is not statistically significant. However, for a large sample size, the information lost will be statistically significant so that more groups are needed to achieve a sufficiently good fit. For an infinitely large sample size, the information lost will always be statistically significant when based on a finite number of groups.

Although not particularly the focus of this investigation, this apparent ambiguity in the selection of the right number of groups also makes problematic the assumption that there even is a right number of groups. For example, based on the original Bauer and Curran (2003) article in *Psychological Methods* and the set of responses and rejoinders that it prompted, Bauer (2007) offered some observations about growth mixture models that are also relevant to this investigation. He argued that "most of these applications are unlikely to reproduce the underlying taxonomic structure of the population. At a more fundamental level, in many cases there is no taxonomic structure to be found" (p. 757). Acknowledging that classes of individuals can approximate the true underlying continuum of individual differences in ways that can be useful, the categorical classifications can also be problematic so that their usefulness needs to be scrutinized carefully. In particular, Bauer (2007) argued that assumptions underlying mixture models are often problematic, whereas the theoretical rationale is often dubious. Citing Maughan (2005), Nagin (2005), Bauer and Curran (2003), and others, Bauer (2007) argued that it is reasonable to assume that most behaviors, traits, and abilities will differ continuously by degrees rather than by types. Finally, acknowledging the potential value of mixture modeling even when distinct groups are not thought to exist (what he referred to as *indirect applications*), Bauer argued that there was a danger in overinterpreting the results to reify groups that are heuristic at best and often merely convenient fictions. Making a similar point, Muthén and Asparouhov (2008; also see Muthén, 2004) stressed that:

> It is important to note that the need for latent classes may be due to non-normality of the outcomes rather than substantively meaningful subgroups (see, e.g., McLachlan & Peel, 2000, pp. 14–17; Bauer & Curran, 2003). To support a substantive interpretation of the latent classes, a researcher should consider not only the outcome variable in question, but also antecedents (covariates predicting latent class membership), concurrent outcomes, and distal outcomes (predictive validity). (pp. 28–29)

Particularly relevant to this discussion, the results of this investigation offer the application of well-known multivariate approaches—discriminant function analysis and canonical correlation—to further interrogate the data in terms of what is lost—or gained—in terms of variance explained by using classes of persons rather than continuous variables. Approaches demonstrated here also offer a stronger basis for interrogating the construct validity of the latent groups in terms of diverse correlates (antecedent, concurrent, and distal variables—or simply correlates that do not rely on a hypothesized causal ordering).

Alternative Styles of Presentation and Analysis

Particularly when LPAs are based on many indicators, it might not always be clear on what dimensions the classes primarily differ and how these relate to a priori hypotheses about the nature of the groups. In relation to this issue, the box plots (Figure 2) seem more useful than the typical line plots. To these we added marker measures of profile level (the mean of all ASCs) and shape (difference between math and verbal self-concepts) that were based on a priori predictions. Based on these pure measures of shape and level, we demonstrated that all the observed groups were readily interpretable in relation to a priori hypotheses. These interpretations were based on a combination of central tendency and variability—a particular strength of box plots. Whereas the level indicator (mean value of all indicators) might be broadly applicable in different studies, a priori hypotheses in relation to shape should necessarily be idiosyncratic to each particular study. Nevertheless, support for a priori predictions about shape is particularly important in the justification of the LPA. Although there might be some special circumstances in which groups that differ solely in terms of level are important, we suggest that the reduction of results based on reasonably continuous variables into a small number of discrete groups that differ only in terms of level typically is not a useful application of LPA. Our claim is based on the recognition that the "right" number of groups in this situation is likely to be arbitrary, idiosyncratic to characteristics of a particular study (e.g., sample and sample size) and not replicable, and that much information is lost in terms of variance explained (see related discussion by MacCallum, Zhang, Preacher, & Rucker, 2002, on the appropriateness of dichotomizing reasonably continuous variables).

We also felt that results of the discrimant function analyses and particularly the canonical correlations were an important contribution to the construct validation of interpretations based on the LPA. For instance, in the discriminant function analysis, the first two discriminant functions explained the majority of the explained variance and these functions were clearly supportive of the a priori predictions in relation to the nature of the groups (in terms of level and shape). A critical issue in the evaluation of the LPAs should be the amount of variance that is explained by alternative LPA models as well as the original variables on which they are based. This is particularly useful when the variables on which the LPAs are based (the ASCs in this study) are supplemented by an extensive set of correlates. Finally, the juxtaposition of these supplemental presentations—box plots, discriminant function analyses, and canonical correlations—were critical in demonstrating that groups based on ASCs and correlates were qualitatively different from groups based on only the ASCs. More generally, presentational approaches and subsequent interrogation of LPA results are consistent with a broader approach to construct validity that should be the hallmark of applied research and particularly substantive-methodological synergies. This construct validity approach where the researcher takes the role of a data detective is particularly important in the application of new statistical techniques where well-established guidelines of best practice are still evolving (see Bauer, 2007; Hipp & Bauer, 2006; see also Marsh, Byrne, & Yeung, 1999; Marsh & Hau, 2007).

Integration of Correlates

The applied researcher is left with the task of how to incorporate correlates into the analysis. There is no clear resolution to this issue and the decision must necessarily be idiosyncratic to each application and a priori assumptions about the nature of the correlates.

One way to do so, as in the present investigation (ASC + CORR–LPAs; see Figure 2b), is to allow the correlates to directly affect the LPA latent grouping variable (i.e., arrows in the path diagram go from the correlates to the latent grouping variable). However, there are alternative approaches (Lubke & Muthén, 2005, 2007) in which, for example, the correlates can be affected by grouping variables (which essentially eliminates the distinction between the original variables used to define the latent groups and the correlates) or the correlates can influence both the LPA grouping variables and the indicator variables (although some additional constraints are needed for this model to be identified).

This increased flexibility in incorporating correlates into the LPA leaves unanswered the question as to whether correlates should be included at all. In this investigation, we demonstrated that the set of LPA grouping variables based on only the original ASCs was substantially different from those that incorporated the correlates into the LPA. Indeed, because the whole point of considering correlates was to validate the latent profiles and associated grouping variables, we argue that it is inappropriate to include the correlates in the LPA when they influence the definition of the latent groups. Not surprisingly, LPA grouping variables that included the correlates as covariates in the actual LPA model were substantially more related to the correlates. However, this does not mean that the ASC + CORR–LPA groups were substantially more valid than the LPA–ASC groups. Indeed, the inclusion of the correlates in the LPA so confounded the interpretation that it was difficult to determine to what extent the groups even reflected the original ASCs, as opposed to the correlates, and made problematic any attempts to validate the ASC + CORR–LPA groups in relation to the original set of correlates.

It is also useful to think of the ASC + CORR–LPA groups as a combination of reflective indicators (ASCs, arrows going from groups to ASCs) and formative indicators (correlates, arrows going from groups to ASC). Although the appropriate use and interpretation of formative factors is not resolved, the criticisms by Howell et al. (2007) are relevant. In particular, the latent profiles used to explain any well-defined set of variables (e.g., the ASCs) that includes (formative) covariates in the LPA model will be idiosyncratic to that particular study and not comparable to the results of other studies that include different covariates or do not include any covariates (see related concerns discussed earlier based on Anderson & Gerbing, 1988; Howell et al., 2007; Lubke & Muthén, 2005). Although we could describe the ASC + CORR–LPA and ASC–LPA using similar verbal labels (and there seemed to be substantial similarities between the two), they were clearly different. Indeed the apparent similarity is deceptive and potentially worrisome if not interpreted appropriately. There is clearly indeterminacy in the nature of the groups if they change substantially depending on the correlates that are included. Hence, even when there is a strong a priori justification for the inclusion of correlates as covariates in the LPA model, researchers should be careful to evaluate how the inclusion of correlates alters the nature of the groups.

B. Muthén (personal communication, March 2, 2008) recently clarified this issue, emphasizing that the inclusion of covariates into the LPA (or, more generally, mixture models) is based on the assumption that the covariates are strictly antecedent variables such that they meet the strong assumption that the causal ordering is from the (antecedent) covariates to the latent groups (see Figure 1a). Because strong assumptions about the direction of causal ordering are very difficult to test, interpretations of LPAs that included covariates should always be looked at cautiously and the nature of the groups with and without covariates should be examined using approaches such as those demonstrated here, especially when the inclusion of covariates does qualitatively change the nature of the latent groups. When, as here, it is clear that the assumption about the direction of causality of relating covariates and grouping variables is violated, it is generally inappropriate to include the covariates in the analysis (as in Model 2). Indeed, the introduction of the auxiliary variable approach into the latest version of M*plus* is specifically designed to accommodate this situation—allowing researchers to relate the latent groups to correlates without actually incororating the correlates into the LPA as covariates that influence the grouping variables. Although there are potential advantages in the use of the auxiliary variable approach in M*plus* (because they are based on pseudo-class draws that take into account the probabilities that a particular case will fall into a particular class), we did not find this to be an important consideration in this investigation (the pattern of group differences based on auxiliary variable approach in M*plus* was highly similar to those based on grouping variables in our supplemental analyses). Importantly, sole reliance on the use of the auxiliary variable approach does not facilitate the subsequent analyses used here to interrogate the construct validity of interpretations based on the LPA in relation to a priori predictions. However, because appropriate practice and the capability of statistical software is evolving rapidly, it is likely that further developments will facilitate the more flexible analyses like those considered here within a single step that does not necessitate outputting results from one analysis to serve as input data for additional, supplemental analyses. Indeed, this seems particularly important if LPAs and mixture models are seen as preliminary analyses of part of a larger research program rather than an end in themselves.

We are not arguing that it is always or even typically inappropriate to incorporate correlates into the LPA model. Indeed, Muthén (2006) provides examples when it might be inappropriate not to do so. For example, background demographic variables such as gender and age might result in differences in responses on indicators even when individuals belong to the same latent class. In this investigation, for example, it might have been appropriate to include only gender in the LPA analysis to determine the extent to which gender differences in the original ASCs could be explained in terms of the LPA grouping variables and to control for gender differences in validating the ASC–LPA groups in relation to the remaining correlates. Alternatively, when the grouping variables do not provide sufficient information to accurately distinguish between groups, it might be useful to include information from covariate measures to improve accuracy of classification. Nevertheless, the manner in which correlates are included into the LPA, if at all, is a highly complex issue that must be based on the nature of the variables being considered, the purposes of the analyses, and the theoretical issues being pursued. Whenever researchers do incorporate correlates into their LPAs, we recommend a careful consideration of the nature of the LPA grouping variables with and without the inclusion of the correlates and how this alters the interpretation of the results. Although more research is needed to explicate the importance of violating assumptions associated with the inclusion of covariates into LPAs, the applied researcher should defend the decision to include covariates theoretically and empirically. Particularly when a strong a priori case cannot be made for the covariates being antecedent variables in which the causal ordering is from covariates to grouping variables, the inclusion of covariates into LPAs should be done with extreme caution—if at all. Although we have clearly not resolved this issue, we hopefully have provided an alternative perspective and tools for applied researchers to evaluate this issue.

The focus of this investigation has been on the inappropriateness of considering correlates as antecedent covariates when there is not a strong basis for the causal ordering assumption that antecedent covariates “cause” the latent groups (or grouping variables). Although given less attention, it is also possible to include correlates as distal outcomes when there is a strong basis for the causal ordering assumption that the latent groups cause the distal outcomes as might be reasonable in longitudinal studies. This is easily accomplished when latent group probabilities or dichotomies are used in subsequent analyses. However, in an application of growth mixture modeling with longitudinal data, Muthén (2004; also see L. Muthén & Muthén, 2008, Example 8.6) suggested that a distal variable could be included as part of the mixture model in which latent groups were used to predict the distal outcomes. Furthermore, he argued that there might be strategic advantages in doing so in terms of accuracy of prediction. Although we know of no LPAs using this approach to incorporate distal outcomes, this approach clearly warrants further consideration in future research. We have not considered it in this investigation as we argue that ASCs (and groups based on) are reciprocally related to many of our correlates and the strong assumption of causal ordering in this distal outcome approach is typically not appropriate in cross-sectional studies based on a single wave of data. However, even when there is a strong basis for latent groups causing the distal outcomes, we encourage researchers to carefully scrutinize the construct validity of models with and without the inclusion of distal outcomes using procedures outlined here to determine whether the inclusion of distal outcomes has altered the nature of the groups.

SUMMARY

This investigation clearly demonstrates that it is useful to use traditional variable-centered and person-centered approaches in combination. They should be seen as complementary approaches—different ways of looking at the same phenomena. To the extent that latent profiles represent differences in profile shape, the LPAs are likely to be particularly useful. More generally when LPA groups reflect a combination of shape and level, as in this investigation, the complementary nature of the two approaches is more clearly evident. Even here, however, it is important to emphasize that LPA groups are formed to maximize the distinctiveness of the groups so that some—perhaps much—of the variance in the scores that make up those groups is likely to be lost. This is particularly the case when a small number of LPA groups is based on a large number of reasonably distinct indicators. Here, as is typically the case, there is a need to balance the parsimony gained from considering only a small number of groups with this loss of information. Indeed, even though the results of this investigation seem to be a particularly appropriate application of LPA, it is important to emphasize that the original set of ASCs was able to explain about 25% of the variance in the set of correlates, whereas the set of ASC LPA group probabilities only explained about 15% of the variance in correlates (although not reported, LPA group dichotomies explained even less). In fact, the five LPA groups were able to explain less than two thirds of the variance in the ASC indicators on which they were based. Although "variance explained" might be an inherently variable-centered basis for evaluating the results, it does demonstrate that much reliable and valid information in the original group indicators can be lost in the formation of LPA groups. In this sense, we encourage researchers to simultaneously pursue both person- and variable-centered approaches to maximize the benefits of both approaches.

It is important is emphasize that there are new and evolving methodologies that allow researchers considerable flexibility in pursuing alternative approaches to issues raised here. Central to the definition of the classic LPA model is the assumption of conditional independence. However, this assumption can be relaxed in more general mixture models that treat LPA (and the assumption of conditional independence) as a special case. If this assumption is relaxed, it is possible to test a priori assumptions about the nature of the within-group associations between indicator variables and the invariance of these relations over the groups (e.g., Lubke & Muthén, 2007; Uebersax, 1999; Vermunt & Magidson, 2002), or to incorporate a a one-factor mixture model as a compromise between classic LPA assumption of conditional independence and mixture models with an unrestricted covariance matrix. In contrast to this investigation, there are many applications in which it is appropriate to consider the correlates as covariates that have a causal effect on the grouping variables (or their indicators) or as distal outcomes variables that are caused by the grouping variables (e.g., Lubke & Muthén, 2007; Tofighi & Enders, 2007). Clearly, such analyses are facilitated by longitudinal data where assumptions of causal ordering are supported by temporal ordering. It is also possible to consider hybrid models that combine factor models and mixture models (e.g., Lubke & Muthén, 2005, 2007; Lubke & Neale, 2006). In this investigation, the issue of power was not a critical issue because of the large sample size, but particularly in studies with fewer cases it would be relevant to explore the power to detect differences between classes (e.g., Lubke & Spies, 2007). However, even when more complex mixture models are considered, the issues considered here (number of groups and their interpretation, nature of causal assumptions) are relevant, and simpler models like the ones considered here provide appropriate baseline comparisons to evaluate how the nature of groups is changed (e.g., Bauer, 2007).

ACKNOWLEDGMENTS

Work on this investigation was conducted, in part, while Professor Herbert W. Marsh was a Visiting Scholar at the Center for Educational Research at the Max Planck Institute for Human Development in Berlin. This research was funded in part by a grant from the United Kingdom Economic and Social Research Council to the first author and support from the Max Planck Institute. The authors would like to thank Bengt Muthén, Gitta Lubke, Dena Pastor, John Uebersax, and Jacqueline Cheng for comments on earlier drafts of this article. Requests for further information about this investigation should be sent to Professor Herbert W. Marsh, Dept. of Education, University of Oxford, 15 Norham Gardens, Oxford OX2 6PY, UK; E-mail: herb.marsh@education.ox.ac.uk.

REFERENCES

Allport, G. W. (1937). *Personality: A psychological interpretation.* New York: Holt.

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103,* 411–423.

Bartholomew, D. J. (1987). *Latent variable models and factor analysis.* London: Griffin.

Bauer, D. J. (2007). Observations on the use of growth mixture models in psychological research. *Multivariate Behavioral Research, 42,* 557–786.

Bauer, D. J., & Curran, P. J. (2003). Distributional assumptions of growth mixture models: Implications for over-extraction of latent trajectory classes. *Psychological Methods, 8,* 338–363.

Bauer, D. J., & Curran, P. J. (2004). The integration of continuous and discrete latent variable models: Potential problems and promising opportunities. *Psychological Methods, 9,* 3–29.

Baumert, J., Bos, W., & Lehmann, R. (Eds.). (2000). *TIMSS/III: Dritte Internationale Mathematik- und Naturwissenschaftsstudie—Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn* [TIMSS/III: Third international mathematics and science study—Students' knowledge of mathematics and science at the end of secondary education]. Opladen, Germany: Leske & Budrich.

Bentler, P. M. (1990). Comparative fit indices in structural models. *Psychological Bulletin, 107,* 238–246.

Boyle, G. J. (1994). Self-Description Questionnaire II. *Test CORRiques, 10,* 632–643.

Byrne, B. M. (1996). *Measuring self-concept across the life span: Issues and instrumentation.* Washington, DC: American Psychological Association.

Collins, L. M., Fidler, P. L., Wugalter, S. E., & Long, J. D. (1993). Goodness-of-fit testing for latent class models. *Multivariate Behavioral Research, 28,* 375–389.

Educational Testing Service. (1997). *TOEFL 1997 edition: Test & score manual.* Princeton, NJ: Author.
Gibson, W. A. (1959). Three multivariate models: Factor analysis, latent structure analysis, and latent profile analysis. *Psychometrika, 24,* 229–252.
Hattie, J. (1992). *Self-concept.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Henson, J. M., Reise, S. P., & Kim, K. H. (2007). Detecting mixtures from structural model differences using latent variable mixture modeling: A comparison of relative model fit statistics. *Structural Equation Modeling, 14,* 202–226.
Hipp, J. R., & Bauer, D. J. (2006). Local solutions in the estimation of growth mixture models. *Psychological Methods, 11,* 36–53.
Howell, R. D., Breivik, E., & Wilcox, J. B. (2007). Reconsidering formative measurement. *Psychological Methods, 12,* 205–218.
Köller, O., Daniels, Z., Schnabel, K. U., & Baumert, J. (2000). Kurswahlen von Maedchen und Jungen im Fach Mathematik: Zur Rolle von fachspezifischem Selbstkonzept und Interesse [Course selections of girls and boys in mathematics: The role of academic self-concept and interest]. *Zeitschrift für Pädagogische Psychologie, 14,* 26–37.
Köller, O., Watermann, R., Trautwein, U., & Lüdtke, O. (2004). *Wege zur Hochschulreife in Baden-Württemberg. TOSCA—Eine Untersuchung an allgemein bildenden und beruflichen Gymnasien* [Educational pathways to college in Baden-Württemberg. TOSCA—A study of traditional and vocational Gymnasium schools]. Opladen, Germany: Leske & Budrich.
Lazarfeld, P. F., & Henry, N. W. (1968). *Latent structure analysis.* Boston: Houghton Mifflin.
Lo, Y., Mendell, N., & Rubin, D. (2001). Testing the number of components in a normal mixture. *Biometrika, 88,* 767–778.
Loken, E., & Molenaar, P. (2008). Categories of continua? The correspondence between mixture models and factor models. In G. R. Hancock & K. M. Samuelsen (Eds.), *Advances in latent variable mixture models* (pp. 277–297). Charlotte, NC: Information Age.
Lubke, G. H., & Muthén, B. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods, 10,* 21–39.
Lubke, G. H., & Muthén, B. (2007). Performance of factor mixture models as a function of model size, criterion measure effects, and class-specific parameters. *Structural Equation Modeling, 14,* 26–47.
Lubke, G., & Neale, M. (2006). Distinguishing between latent classes and continuous factors: Resolution by maximum likelihood? *Multivariate Behavioral Research, 41*(4), 499–532.
Lubke, G. H., & Spies, J. R. (2008). Choosing a "correct" factor mixture model: Power, limitations, and graphical data exploration. In G. R. Hancock & K. M. Samuelson (Eds.), *Advances in latent variable mixture models* (pp. 343–362). Charlotte, NC: Information Age.
MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7,* 19–40.
Magidson, J., & Vermunt, J. (2004). Latent class models. In D. Kaplan (Ed.), *Handbook of quantitative methodology for the social sciences* (pp. 175–198). Newbury Park, CA: Sage.
Magnusson, D. (1990). Personality development from an interactional perspective. In L. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 193–222). New York: Guilford.
Magnusson, D., & Stattin, H. (1998). Person-context interaction theories. In W. Damon (Series Ed.) & R. M. Lerner (Vol. Ed.), *Handbook of child psychology: Vol. 1. Theoretical models of human development* (pp. 685–759). New York: Wiley.
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. (1989a). Age and sex effects in multiple dimensions of self-concept: Preadolescence to early-adulthood. *Journal of Educational Psychology, 81,* 417–430.
Marsh, H. W. (1989b). Sex differences in the development of verbal and math constructs: The High School and Beyond study. *American Educational Research Journal, 26,* 191–225.
Marsh, H. W. (1990). A multidimensional, hierarchical self-concept: Theoretical and empirical justification. *Educational Psychology Review, 2,* 77–172.
Marsh, H. W. (1992a). The content specificity of relations between academic achievement and academic self-concept. *Journal of Educational Psychology, 84,* 35–42.

Marsh, H. W. (1992b). *Self-Description Questionnaire III: Manual.* Sydney, Australia: University of Western Sydney, SELF Research Centre.

Marsh, H. W. (1993). Academic self-concept: Theory measurement and research. In J. Suls (Ed.), *Psychological perspectives on the self* (Vol. 4, pp. 59–98). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Marsh, H. W. (2007a). Application of confirmatory factor analysis and structural equation modeling in sport/exercise psychology. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of sport psychology* (3rd ed., pp. 774–798). New York: Wiley.

Marsh, H. W. (2007b). *Self-concept theory, measurement and research into practice: The role of self-concept in educational psychology.* Leicester, UK: British Psychological Society.

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

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., Byrne, B. M., & Yeung, A. S. (1999). Causal ordering of academic self-concept and achievement: Reanalysis of a pioneering study and revised recommendations. *Educational Psychologist, 34,* 155–167.

Marsh, H. W., & Craven, R. G. (1997). Academic self-concept: Beyond the dustbowl. In G. Phye (Ed.), *Handbook of classroom assessment: Learning, achievement and adjustment* (pp. 131–198). Orlando, FL: Academic.

Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science, 1,* 133–163.

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

Marsh, H. W., & Hau, K. T. (2004). Explaining paradoxical relations between academic self-concepts and achievements: Cross-cultural generalizability of the internal-external frame of reference predictions across 26 countries. *Journal of Educational Psychology, 96,* 56–67.

Marsh, H. W., & Hau, K. T. (2007). Applications of latent-variable models in educational psychology: The need for methodological-substantive synergies. *Contemporary Educational Psychology, 32,* 151–171.

Marsh, H. W., Hau, K. T., & Grayson, D. (2005). Goodness of fit evaluation in structural equation modeling. In A. Maydeu-Olivares & J. McArdle (Eds.), *Contemporary psychometrics. A festschrift for Roderick P. McDonald* (pp. 275–340). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis testing approaches to setting cutoff values for fit indexes and dangers in overgeneralising Hu & Bentler's (1999) findings. *Structural Equation Modeling, 11,* 320–341.

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

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

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2006). Integration of multidimensional self-concept and core personality constructs: Construct validation and relations to well-being and achievement. *Journal of Personality, 74,* 403–455.

Marsh, H. W., & Yeung, A. S. (1998). Longitudinal structural equation models of academic self-concept and achievement: Gender differences in the development of math and English constructs. *American Educational Research Journal, 35,* 705–738.

Marsh, H. W., & Yeung, A. S. (2001). An extension of the internal/external frame of reference model: A response to Bong (1998). *Multivariate Behavioral Research, 36,* 389–420.

Maughan, B. (2005). Developmental trajectory modeling: A view from developmental psychopathology. *Annals of the American Academy of Political and Social Science, 602,* 118–130.

McLachlan, G., & Peel, D. (2000). *Finite mixture models.* New York: Wiley.

Meehl, P. E. (1995). Bootstraps taxometrics: Solving the classification problem in psychopathology. *American Psychologist, 50,* 266–275.

Muthén, B. O. (2001). Latent variable mixture modeling. In G. Marcoulides & R. Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp. 1–33). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Muthén, B. O. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *Handbook of quantitative methodology for the social sciences* (pp. 345–368). Newbury Park, CA: Sage.
Muthén, B. O. (2006). Should substance use disorders be considered as categorical or dimensional? *Addiction, 101*(Suppl. 1), 6–16.
Muthén, B. O., & Asparouhov, I. (2008). Growth mixture modeling: Analysis with non-Gaussian random effects. In G. Fitzmaurice, M. Davidian, G. Verbeke, & G. Molenberghs (Eds.), *Longitudinal data analysis* (pp. 143–165). Boca Raton, FL: Chapman & Hall/CRC Press.
Muthén, B. O., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research, 24,* 882–891.
Muthén, L., & Muthén, B. O. (2008). *M*plus *user's guide* (5th ed.). Los Angeles: Muthén & Muthén.
Myers, I. B., & McCaulley, M. H. (1985). *Manual: A guide to the development and use of the Myers-Briggs Type Indicator.* Palo Alto, CA: Consulting Psychologists Press.
Nagin, D. S. (2005). *Group-based modeling of development.* Cambridge, MA: Harvard University Press.
Nylund, K. L., Asparouhov, A., & Muthén, B. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14,* 535–569.
Pastor, D. A., Barron, K. E., Miller, B. J., & Davis, S. L. (2007). A latent profile analysis of college students' achievement goal orientation. *Contemporary Educational Psychology, 32,* 8–47.
Robins, R. W., John, O. P., Caspi, A., Moffitt, T. E., & Stouthamer-Loeber, M. (1996). Resilient, overcontrolled, and undercontrolled boys: Three replicable personality types. *Journal of Personality and Social Psychology, 70,* 157–171.
Schwanzer, A. D., Trautwein, U., Lüdtke, O., & Sydow, H. (2005). Entwicklung eines Instruments zur Erfassung des Selbstkonzepts junger Erwachsener [Development of a questionnaire to measure young adults' self-concept]. *Diagnostica, 51,* 183–194.
Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Validation of construct interpretations. *Review of Educational Research, 46,* 407–441.
SPSS. (2006). *SPSS for Windows*, Version 14. Chicago: Author.
Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). New York: HarperCollins.
Tofighi, D., & Enders, C. K. (2007). Identifying the correct number of classes in a growth mixture models. In G. R. Hancock & K. M. Samuelson (Eds.), *Advances in latent variable mixture models* (pp. 317–341). Greenwich, CT: Information Age.
Uebersax, J. S. (1999). Probit latent class analysis with dichotomous or ordered category measures: Conditional independence/dependence models. *Applied Psychological Measurement, 23,* 283–297.
Vermunt, J. K., & Magidson J. (2002). Latent class cluster analysis. In J. A. Hagenaars & A. McCutcheon (Eds.), *Applied latent class analysis* (pp. 89–106). Cambridge, UK: Cambridge University Press.
Wylie, R. C. (1989). *Measures of self-concept.* Lincoln, NE: University of Nebraska Press.
Yang, C. (2006). Evaluating latent class analyses in qualitative phenotype identification. *Computational Statistics & Data Analysis, 50,* 1090–1104.

APPENDIX
DESCRIPTION OF THE ACADEMIC SELF-CONCEPT SCALES USED IN THIS INVESTIGATION

Verbal (VerbSC; reliability = .795; 4 items). Participants' self-perceptions of their verbal skills, verbal reasoning ability, and interest in verbal activities.

Math (MathSC; reliability = .893; 4 items). Participants' self-perceptions of their mathematical skills, mathematical reasoning ability, and interest in mathematics.

Problem Solving (ProbSC; reliability = .648; 4 items). Participants' self-perceptions of their problem-solving skills and their ability to find new solutions to problems.

Intellectual (IntlSC; reliability = .773; 4 items). Participants' self-perceptions of being bright and intelligent rather than slow and uninspired.

Artistic (ArtsSC; reliability = .902; 4 items). Participants' self-perceptions of their abilities in and understanding of music and arts.

Political (PoltSC; reliability = .945; 4 items). Participants' self-perceptions of their understanding of politics and their ability to contribute in political discussions.

Technical (TechSC; reliability = .743; 4 items). Participants' self-perceptions of their understanding of technical processes related to generic life skills and abilities concerning skillful technical work.

Computers (CompSC; reliability = .876; 4 items). Participants' self-perceptions of how skillfully they use computers and a variety of computer programs.

Reliability estimates are coefficient alpha estimates based on results of this investigation.