Robust Structural Equation Models: Implications for Developmental Psychology

George J. Huba

Western Psychological Services

Lisa L. Harlow

University of Rhode Island

HUBA, GEORGE J., and HARLOW, LISA L. *Robust Structural Equation Models: Implications for Developmental Psychology*. CHILD DEVELOPMENT, 1987, **58**, 147–166. Advances in structural equation modeling techniques have made it possible to test models in data that are not normally distributed. This can lead to more realistic model testing in developmental psychology. Several alternate techniques are illustrated in structural equation models here in order to compare the results that are obtained. Maximum-likelihood and generalized least-squares estimators for normally distributed data are compared with Browne's asymptotically distribution-free technique for continuous nonnormally distributed data and Muthen's estimator for dichotomous indicators. While different critical ratios are found for some parameters estimated in models, the results are generally comparable so long as one does not consider absolute fit to be a critical factor in "accepting" a causal model as a good one.

A major revolution in the study of important psychological phenomena as they develop during the lifespan has been the recent advances in analytical and statistical techniques for modeling the covariation of major constructs in the areas of behavior, personality, attitudes, and environmental characteristics. This article examines some real structural equation models to illustrate some major points in the estimation of the parameters in these models. The central methodological question asks what types of statistical techniques should be used when it is suspected that the data available for testing models may not have textbook normal distributions. Can causal or structural equation modeling be done with nonnormal data, and are alternate methods for nonnormal data superior with the typical datasets and theoretical models tested?

From the standpoint of the applied developmental psychologist, structural equation models have several important advantages to recommend them. First, structural equation models permit one to unambiguously develop models to represent an important theoretical framework. Second, the structural equation models can study the influence of one "error-free" construct on another "error-free" construct so long as the constructs are measured with multiple indicators or variables. A simple algebraic demonstration is given by Huba and Bentler (1982a) that using true scores to assess theories yields more accurate estimates of causal influence than using total (observed) scores. Thus, structural equation models with latent variables can permit us to eliminate the potentially confounding influences of measurement error in the observed variables. Third, if the data are longitudinal or have other major aspects of "quasi-experimentation" designed to control confounding influences, it may be possible to draw some conclusions about causality in phenomena that cannot be ethically studied through experimentation. Fourth, even if the data are not sufficiently strong to permit causal inference, it is possible to compare and contrast the relative fit of the model to the data using several alternate theoretical frameworks.

Having clearly specified a model, computer programs are available that can simultaneously analyze the various patterns of interrelations implied by the major equations in

The authors are indebted to Drs. Michael W. Browne and Bengt Muthen for the use of their computer programs for structural equation modeling. Three anonymous reviewers made detailed comments that were extremely useful in developing the final draft of the paper. Any remaining errors are, of course, those of the authors. The ideas expressed here are those of the authors and may not reflect official policy of Western Psychological Services (WPS). Address correspondence to G. J. Huba, Western Psychological Services, 12031 Wilshire Boulevard, Los Angeles, CA 90025.

[*Child Development*, 1987, **58**, 147–166. 0009-3920/87/5801-0005$01.00]

the system. Under assumptions of multivariate normality among the observed variables, parameter estimates for a model are generally obtained via a maximum-likelihood estimation procedure that yields both a large-sample chi-square test and asymptotic standard errors of estimates. The parameters estimated include variances and covariances for both the latent variables as well as disturbances in the latent constructs. Widely available computer programs for this procedure include LISREL VI (Jöreskog & Sörbom, 1984), EQS (Bentler, 1987, in this issue), and COSAN (McDonald, 1980), and it would be a fair characterization to say that maximum-likelihood (ML) estimation currently dominates in the scientific marketplace for structural equation modeling specifically because of the historical importance of the widely circulated LISREL program and well-known asymptotic (large-sample) efficiency properties of ML estimation in general. Nonetheless, newer estimation methods exist that permit the data to be distributed in ways other than with multivariate normality, and these alternatives may rate as the most important developments in structural equation modeling since the original systematization by Jöreskog and Sörbom.

In the discussion that follows, we have deliberately not introduced the fundamental equations for some alternate estimators. A far more complete and technical introduction to these issues is provided by Huba and Harlow (1986), who also give the major equations and technical citations.

Potential Problems with Structural Equation Models as Presently Used

Despite the valuable contribution of the latent-variable causal modeling techniques with maximum-likelihood estimators to research methodology, several problems can arise during implementation. The first problem concerns the distribution of the variables. Virtually all applications wishing to employ a statistical estimator so that goodness of fit can be assessed utilize maximum-likelihood estimation, which is based on the assumption that the observed variables follow a multivariate normal distribution. While the efficiency of this statistical estimation procedure has been studied (e.g., Browne, 1968; Jöreskog, 1967) and several investigators have suggested robustness for the maximum-likelihood estimations against violations of normality (e.g., Fuller & Hemmerle, 1966; Huba & Bentler, 1983a; Jöreskog & Sörbom, 1984), the validity of the chi-square test statistics and the standard errors may still be suspect with nonnormal data since a fundamental mathematical assumption is violated (Browne, 1982, 1984). As an alternative, one could utilize estimation procedures that do not require such a restrictive and perhaps unrealistic assumption.

Browne (1982, 1984; Browne & Cudeck, 1983) discuss a class of distributions that permit use of best generalized least-squares estimators even when the variables exhibit excessive kurtosis ("peakedness") or insufficient kurtosis ("flatness") when compared to the multivariate normal distribution. While social scientists often seem to worry about the skewness in their data, Browne points out that in fact it is the kurtosis that is critical since it is a term in the mathematical expression for the covariances of the covariances. That is, when the data are not normally distributed, we must know about the variable kurtoses as well as the variable means and covariances in order to infer facts about individual patterns of scores. Browne introduces an asymptotically distribution-free (ADF) method for obtaining parameter estimates, standard errors, and a fit statistic. This ADF estimate has been successfully applied to many different causal models with nonnormal data typical of those encountered in developmental psychology (Huba & Bentler, 1983b; Huba & Harlow, 1983, 1986; Huba & Tanaka, 1983). Browne implements his estimator within a framework that can be considered, from the standpoint of the user, as identical to that of Jöreskog and Sörbom. Thus, the Browne estimator is one that is applicable to continuous variables that do not appear to be normally distributed.

It might be noted that for data that are assumed to have no extra kurtosis over that of a normal distribution, the Browne ADF estimator is asymptotically equivalent to one called the generalized least-squares estimator (Browne, 1974; Jöreskog & Goldberger, 1972; Jöreskog & Sörbom, 1984). The generalized least-squares estimator (or GLS) for normally distributed variables was initially developed to provide a potentially less expensive (in computer time) way of estimating the parameters in causal models when the data were normally distributed or have multivariate kurtoses equivalent to that expected from a normal distribution.

A common application of Browne's ADF estimator occurs in longitudinal studies where it is observed that personality or intellectual functioning variables with "bell-shaped" distributions at some ages are not "bell-shaped" during adolescence or late in the lifespan. Alternately, interesting deviant (i.e., infrequent) behaviors like criminal activity, alcohol consumption, or days of illness may not have normal distributions at any time in the lifespan. While monotonic transformations can usually be applied to nonnormal data to ensure marginal, but not multivariate, normality, transformations of this type are not appropriate in longitudinal studies if the shape of a distribution changes with age. Applying different normalizing transformations to the same variable at different times would destroy the repeated measurement aspects of the study.

A second problem area in causal or structural equation modeling concerns the measurement level of the variables. Procedures utilizing variances, covariances, and product-moment correlations all make the implicit assumption that indicators are at least measured on an interval scale. This is often not feasible or is at least difficult to ensure with social data. Frequently the data are dichotomous or at best ordinal in scale, and the use of such measures as the product-moment correlation becomes problematical (Carroll, 1961). For instance, the variables in a model may be a series of dichotomous indicators of whether or not an illness has occurred. Or they may be a set of stressful events that have or have not happened.

A structural equation procedure capable of handling these noninterval response scale data sets has been developed by Muthen (1978, 1981, 1982a, 1982b, 1983; Muthen, Huba, & Short, 1985; Muthen & Kaplan, 1985). In his computer program, LISCOMP, Muthen uses a limited-information generalized least-squares estimator for dichotomous and polytomous categorical causal models.

Muthen's procedures again look to the user like LISREL. However, while Muthen's *observed* variables are assumed to be dichotomies or ordered categories, he assumes that for each observed variable there is a corresponding *unobserved and normally distributed latent indicator.* If the level on this unobserved latent variable is over some threshold value, then a response of "yes" or "category *a*" is observed, while a response below this threshold is observed as a "no" or "category *b*." Muthen's causal models connect these corresponding (to observed variables) underlying normally distributed latent variables in ways analogous to LISREL models.

It is critical to recognize that in Muthen's method, while we model the responses in the presumed latent or unobserved normally distributed variables, we can only observe dichotomous or ordered-categorical indicators of them. That is, underlying each dichotomous or categorical response is a presumed normally distributed variable that is dichotomized by our powers to detect it beyond some threshold. If we make this assumption, then we can obtain statistical estimates for the parameters and use statistical testing methods paralleling those used currently in LISREL. Such a statistical model seems especially appropriate when we assume that an underlying normally distributed latent proclivity, for example toward criminality, "causes" the performance of specific behaviors such as crimes in this example.

Muthen's approach should be contrasted to the approach for categorical variables employed by Jöreskog and Sörbom, which is only approximate and does not necessarily lead to a numerically proper test statistic. Jöreskog and Sörbom calculate tetrachoric correlations for the case of dichotomous variables or polychoric correlations for the case of polytomous variables. They then use this matrix as input to a "regular" statistical estimation of the parameters. The user may specify either maximum-likelihood estimators or ordinary least-squares estimates. Muthen, on the other hand, uses a "best" weight matrix for the population tetrachoric or polychoric correlations. Muthen's technique is a proper statistical one supported by appropriate statistical theory (e.g., Muthen, 1984), while Jöreskog and Sörbom's technique is an approximate one using robust correlation estimates. Their procedure does not ensure that there are proper standard errors and a correct global significance test. As an approximate technique, Jöreskog and Sörbom's method may yield better numbers than incorrectly applying quantitative, normal-data techniques, but its only major advantage over Muthen's related technique might be significantly cheaper cost. In practice, however, such an advantage has not been observed in several problems.

An Empirical Comparison of the Approaches with Quantitative Variables

It is illuminating to compare the results that might be obtained with the different ways of calculating structural equation model parameters in a real data set. Specifically, the following techniques are compared: maximum-likelihood (ML) estimators, generalized least-squares (GLS) estimators, and asymptotically distribution-free (ADF) estimators for continuous data, and Muthen's dichotomous variable (DV) technique for qualitative variables. After transforming the data into dichotomous indicators, tetrachoric correlations were calculated and used as input to both ML and ULS (unweighted or ordinary least-squares) estimation called TETRA-ML and TETRA-ULS here. Finally, after transforming the data into indicators with four or five categories, polychoric correlations were calculated and input into both ML and ULS estimation (POLY-ML and POLY-ULS). The ML and GLS methods are appropriate for normally distributed continuous variables, while the ADF estimator is appropriate for nonnormally distributed continuous variables. The ULS method does not have any distributional assumptions and also does not provide standard errors or a fit statistic. The DV method is appropriate for dichotomous (or categorical) variables, while the TETRA-(ML or ULS) and POLY-(ML or ULS) methods are approximate ones appropriate when the data have, respectively, two or more than two categories. The ML, GLS, and ADF estimates were obtained in LISREL-V (Jöreskog & Sörbom, 1981). The program EQS (Bentler, 1985, 1987) can also provide ADF, ML, and GLS estimates and several other alternatives. The DV estimates were obtained in Muthen's (1982b) program LACCI. Polychoric and tetrachoric correlations and POLY-(ML and ULS) and TETRA-(ML and ULS) estimates were obtained in LISREL-V. It might be noted that the examples given here were developed before Muthen's (1986) LISCOMP computer program became generally available. LISCOMP will allow the user to do causal modeling on polychoric correlations, thus using the full number of categories in the original data.

It should be noted that in order to adequately compare the estimates from the various methods, the data for each of the examples were standardized, yielding correlation coefficients instead of covariances. This was necessary as two programs only allow correlations (LACCI and POLY and TETRA in LISREL). Huba and Harlow (1986) discuss why this is not a problem for the particular examples presented here.

The first example is a structural equation model that relates six personality variables to four drug and alcohol variables through four latent constructs. The data were obtained from a sample of 257 students at Rutgers University by Drs. R. Pandina, E. Labouvie, and D. Lester. The six personality variables included: Law Abidance, Liberalism, Religious Commitment, Self-Acceptance, Invulnerability, and Depression. The four drug and alcohol variables included: Frequency of Beer consumption, Quantity of Marijuana use on a typical day, Frequency of Marijuana use, and Quantity of Marijuana use on a typical day. These variables were conceptualized as indicators of the latent constructs Law Abidance, Self-Acceptance, Beer Consumption, and Marijuana Use. This four-factor model is depicted in Figure 1, where the parameters to be estimated are indicated by Greek letters. In their raw forms (as originally measured),

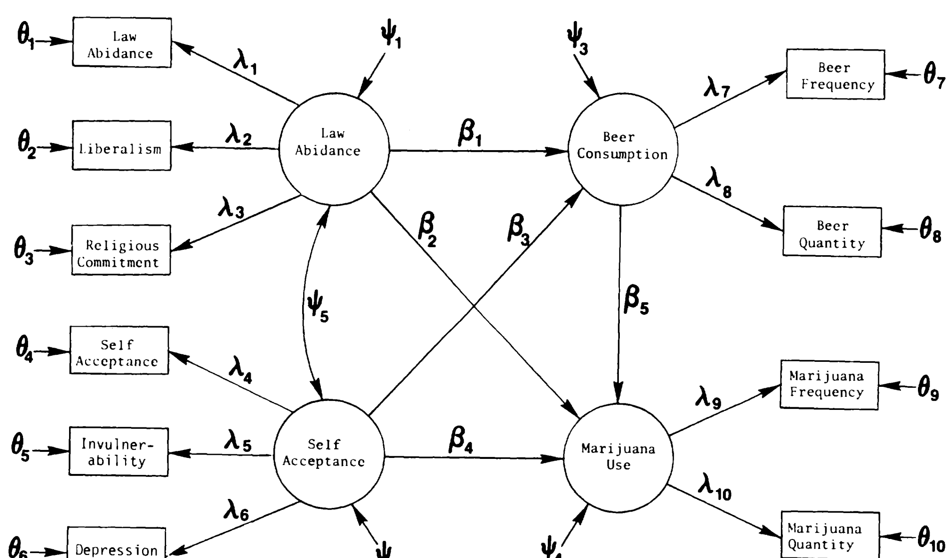

FIG. 1.—Structural equation model for 10 personality and drug variables

the 10 variables had skewnesses and kurtoses of −.20, −.27; −.68, −.59; −.28, −.56; −.44, −.23; −.16, −.12; −.47, −.15; −.01, −1.14; .14, −1.16; .86, −.57; and .71, −.13, respectively. Liberalism was significantly too peaked, while the distributions for Beer Frequency and Beer Quantity were significantly too flat. The multivariate kurtosis coefficient is 4.81, which is statistically different from the value obtained for multivariate normal data. It should also be noted that this model was parameterized with the error variance (04,4) for Self-Acceptance set to zero because of previous model results with this variable (Huba & Bentler, in press).

In this example, the number of response categories for each variable was relatively large (i.e., between 8 and 17). Thus, for the ML, GLS, and ADF solutions, the variables could be thought of as at least approximately continuous. In order to use the other three estimators (i.e., DV, POLY- and TETRA-ML, POLY- and TETRA-ULS), it was necessary to form discrete variables each having two or four response categories. To facilitate more interesting comparisons, two different procedures were utilized to transform the data. In the first the variables were arbitrarily split into fewer response variables by systematically dividing the range into equal segments, often resulting in rather uneven distributions. For instance, the variable Law Abidance was arbitrarily reduced from 16 to 4 response categories for the polychoric case. This results in a distribution with the following percentages in the four response categories: 9%, 31%, 44%, and 16%.

In the second procedure, the variables were split to form approximately normal distributions. Thus, in the four-way split, variables had about 25% of the responses in each of the response categories, while the two-way split resulted in approximately 50% of the cases in each of the response categories. Product-moment, tetrachoric, and polychoric correlations for this example are given in Table 1.

The 10-variable drug and personality example was analyzed using the six different methods (i.e., ML, GLS, ADF, DV, POLY- and TETRA-ML and -ULS). However, when conducting the analyses, it was not possible to obtain a converged solution for the TETRA-ML technique with the uniformly split data. Thus, this column is omitted from the Table. The remaining estimates are presented in Table 2 along with a brief description of each parameter. As can be seen, the estimates are all rather comparable, with the exception of those from the TETRA-ML case with the arbitrary split.

One major lack of comparability that might be explored is the fact that the causal regression coefficients β_2 and β_5 for the ADF solution are discrepant from the ML and GLS results. The large standard errors and position of the parameters within the model suggest that under the ADF estimator the two parameters may be so highly correlated as to be effectively collinear. That is, are the estimates of the causal effects called β_2 and β_5 redundant with one another, which is what would be suggested if the parameter estimates were correlated at a level close to 1.0? In fact, this is the case, since it was found that β_2 and β_5 had a correlation of .99 in the ADF solution but only .71 in the ML one. Of course, the smaller correlation of .71 under the ML estimation is found with an incorrect assumption about the distribution of the variables. Apparently the elimination of distribution effects served to make some of the parameters in the model redundant.

A typical fix in models with collinear parameters is to eliminate one of the parameters. Theoretically, here it made sense to set the path (β_5) from Beer Use to Marijuana Use at zero. This alternate model was then reestimated with all methods, and the resultant parameters and their standard errors are shown in Table 3. Again, with this model it was not possible to obtain a converged solution form the TETRA-ML technique either with the arbitrary or uniform split data. Hence, estimates are presented only for 11 methods. Notice that the parameter estimates are relatively stable in all instances across the ML, GLS, ADF, DV, and ULS methods. It should especially be noted that the parameters representing the causal influences of one latent variable upon another (β_1, β_2, β_3, and β_4) now seem to be more comparable across methods. In addition, while the deletion of the parameter had a negligible effect on the global chi-square fit statistic for the ADF solution, it did have an appreciable (significant) effect on chi-square for the ML and GLS solutions. Thus, using the ADF estimator we may accept a "simpler" solution for the observed data. In this first example, then, the different estimators yield fairly comparable results, although we might want to place somewhat greater reliance on the parameter estimates derived from the ADF solution.

A Second Example Comparing the Approaches

A second example is a structural equation model in which data from 601 individuals

TABLE 1

Correlations for the 10 Variables in Example 1

	1	2	3	4	5	6	7	8	9	10
Product-moment coefficients:										
1. Law Abidance	1.00									
2. Liberalism	−.28	1.00								
3. Religious Commitment	.23	−.31	1.00							
4. Self-Acceptance	.12	−.24	.11	1.00						
5. Invulnerability	−.17	.11	−.04	.31	1.00					
6. Depression	.12	−.30	.19	.69	.15	1.00				
7. Beer Frequency	−.45	−.02	−.13	.00	.01	.04	1.00			
8. Beer Quantity	−.42	.01	−.10	−.05	.06	−.03	.78	1.00		
9. Marijuana Frequency	−.43	.21	−.13	−.05	.07	−.08	.53	.45	1.00	
10. Marijuana Quantity	−.42	.18	−.16	−.06	.00	−.05	.53	.52	.81	1.00
Tetrachoric coefficients:										
1. Law Abidance	1.00	−.26	.21	.15	−.03	.21	−.43	−.37	−.40	−.52
2. Liberalism	−.20	1.00	−.26	−.25	.06	−.41	−.01	−.15	.23	.23
3. Religious Commitment	.18	−.29	1.00	.29	.13	.20	−.14	−.13	−.16	−.21
4. Self-Acceptance	.03	−.23	.12	1.00	.46	.73	.03	.00	−.13	−.35
5. Invulnerability	−.08	−.14	.07	.26	1.00	.11	−.01	.06	.10	.05
6. Depression	.11	−.17	.15	.68	.11	1.00	.13	.02	−.07	−.24
7. Beer Frequency	−.45	−.11	−.06	−.05	−.08	.01	1.00	.79	.65	.50
8. Beer Quantity	−.39	−.04	−.12	−.12	.05	−.01	.84	1.00	.54	.68
9. Marijuana Frequency	−.42	.21	−.12	−.12	−.04	−.04	.65	.58	1.00	.78
10. Marijuana Quantity	−.33	.13	−.28	−.09	−.18	.00	.48	.65	.85	1.00
Polychoric correlations:										
1. Law Abidance	1.00	−.28	.24	.15	−.12	.10	−.46	−.45	−.44	−.47
2. Liberalism	−.22	1.00	−.29	−.22	.18	−.30	−.04	.00	.26	.27
3. Religious Commitment	.25	−.30	1.00	.13	−.01	.19	−.12	−.10	−.11	−.09
4. Self-Acceptance	.07	−.23	.09	1.00	.31	.71	.03	−.05	−.07	−.13
5. Invulnerability	−.16	.10	−.07	.31	1.00	.15	.01	.08	.06	−.01
6. Depression	.09	−.25	.15	.71	.15	1.00	.05	−.07	−.09	−.09
7. Beer Frequency	−.41	−.06	−.13	−.05	−.04	.03	1.00	.80	.59	.58
8. Beer Quantity	−.36	−.07	−.10	−.08	.05	−.03	.82	1.00	.51	.60
9. Marijuana Frequency	−.46	.20	−.14	−.07	−.03	−.07	.61	.53	1.00	.89
10. Marijuana Quantity	−.45	.18	−.17	−.09	−.05	−.06	.57	.57	.91	1.00

Note.—Entries below the diagonal are based on a uniform split of the response categories for each of the variables. Entries above the diagonal are based on an arbitrary split of the response categories for each of the variables.

TABLE 2

PARAMETER ESTIMATES AND STANDARD ERRORS FOR EXAMPLE 1

Parameter and Meaning	ML	GLS	ADF	DV_A[a]	$POLY_A$-ML	$TETRA_A$-ML	$POLY_A$-ULS	$TETRA_A$-ULS	DV_U[a]	$POLY_U$-ML	$POLY_U$-ULS	$TETRA_U$-ULS
λ_1: Factor loading for Law Abidance on Law Abidance factor	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*
λ_2: Factor loading for Liberalism on Law Abidance factor	−.44 (.11)	−.34 (.11)	−.21 (.10)	−.32 (.14)	−.43 (.10)	−.88 (.13)	−.43	−.51	−.48 (.14)	−.42 (.11)	−.34	−.36
λ_3: Factor loading for Religious Commitment on Law Abidance factor	.40 (.10)	.36 (.11)	.44 (.11)	.40 (.14)	.36 (.10)	.44 (.12)	.35	.55	.39 (.12)	.45 (.12)	.40	.44
λ_4: Factor loading for Self-Acceptance on Self-Acceptance factor	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*
λ_5: Factor loading for Invulnerability on Self-Acceptance factor	.31 (.06)	.37 (.06)	.45 (.06)	.43 (.10)	.31 (.06)	.46 (.06)	.23	.28	.18 (.10)	.31 (.06)	.25	.24
λ_6: Factor loading for Depression on Self-Acceptance factor	.69 (.05)	.64 (.05)	.59 (.04)	.60 (.13)	.71 (.04)	.73 (.04)	.73	.74	.51 (.23)	.71 (.04)	.71	.67
λ_7: Factor loading for Beer Frequency on Beer Consumption factor	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*
λ_8: Factor loading for Beer Quantity on Beer Consumption factor	.92 (.07)	.88 (.06)	.81 (.05)	.99 (.09)	.97 (.06)	2.91 (.62)	.96	1.01	1.04 (.07)	.91 (.06)	.92	1.07
λ_9: Factor loading for Marijuana Frequency on Marijuana Use factor	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*
λ_{10}: Factor loading for Marijuana Quantity on Marijuana Use factor	1.02 (.07)	1.03 (.07)	1.01 (.06)	.97 (.10)	1.06 (.05)	2.03 (.23)	1.07	1.14	.94 (.05)	.98 (.04)	1.00	.92
β_1: Path from Law Abidance factor to Beer Consumption factor	−.66 (.14)	−.78 (.21)	−1.45 (.49)	−.84 (.37)	−.65 (.14)	.00 (.05)	−.54	−.72	−.68 (.22)	−.61 (.15)	−.45	−.56

TABLE 2 (*Continued*)

Parameter and Meaning	ML	GLS	ADF	DV_A[a]	$POLY_A$-ML	$TETRA_A$-ML	$POLY_A$-ULS	$TETRA_A$-ULS	DV_U[a]	$POLY_U$-ML	$POLY_U$-ULS	$TETRA_U$-ULS
β_2: Path from Law Abidance factor to Marijuana Use factor	−.38 (.13)	−.35 (.17)	−1.09 (1.18)	−.37 (.31)	−.32 (.12)	−.40 (.10)	−.33	−.48	−.29 (.18)	−.47 (.14)	−.37	−.40
β_3: Path from Self-Acceptance factor to Beer Consumption factor	.09 (.06)	.01 (.06)	.09 (.09)	.23 (.12)	.11 (.06)	−.01 (.02)	.12	.28	−.06 (.07)	.01 (.06)	.02	.03
β_4: Path from Self-Acceptance factor to Marijuana Use factor	.01 (.05)	.04 (.05)	.04 (.10)	−.02 (.10)	−.05 (.05)	−.09 (.03)	−.02	−.06	.02 (.06)	.01 (.05)	−.01	.00
β_5: Path from Beer Consumption factor to Marijuana Use factor	.43 (.09)	.45 (.11)	.03 (.56)	.59 (.18)	.51 (.08)	.31 (.06)	.52	.54	.67 (.13)	.49 (.08)	.53	.60
θ_1: Error variance for Law Abidance	.36 (.12)	.34 (.14)	.54 (.11)	.39	.32 (.13)	.64 (.08)	.25	.50	.33	.41 (.12)	.26	.43
θ_2: Error variance for Liberalism	.88 (.08)	.61 (.07)	.57 (.06)	.94	.87 (.08)	.72 (.08)	.86	.87	.84	.90 (.08)	.92	.93
θ_3: Error variance for Religious Commitment	.90 (.08)	.79 (.08)	.83 (.06)	.90	.91 (.08)	.93 (.08)	.91	.85	.90	.88 (.08)	.88	.89
θ_4: Error variance for Self-Acceptance	.00*	.00*	.00*	−.29	.00*	.00*	.00*	.00*	−.36	.00*	.00*	.00*
θ_5: Error variance for Invulnerability	.90 (.08)	.74 (.07)	.72 (.06)	.76	.90 (.08)	.79 (.07)	.95	.92	.95	.90 (.08)	.94	.94

θ_6: Error variance for Depression	.52	.47	.44	.53	.50	.46	.47	.44	.65	.49	.49	.55
	(.05)	(.04)	(.04)		(.04)	(.04)				(.04)		
θ_7: Error variance for Beer Frequency	.14	.09	.04	.10	.17	.73	.16	.22	.12	.10	.11	.22
	(.05)	(.05)	(.06)		(.04)	(.08)				(.05)		
θ_8: Error variance for Beer Quantity	.28	.27	.30	.11	.22	−1.30	.23	.20	.04	.26	.25	.10
	(.05)	(.04)	(.05)		(.04)	(.42)				(.04)		
θ_9: Error variance for Marijuana Frequency	.21	.18	.12	.20	.16	.62	.16	.32	−.04	.07	.09	.08
	(.04)	(.04)	(.04)		(.03)	(.06)				(.03)		
θ_{10}: Error variance for Marijuana Quantity	.17	.15	.12	.25	.05	−.58	.05	.11	.07	.11	.09	.22
	(.05)	(.04)	(.04)		(.03)	(.12)				(.03)		
ψ_1: Variance of Law Abidance factor	.64	.56	.33	.61	.68	.36	.75	.50	.67	.59	.74	.57
	(.14)	(.16)	(.12)	(.25)	(.15)	(.09)			(.20)	(.14)		
ψ_2: Variance of Self-Acceptance factor	1.00	.92	.87	1.29	1.00	1.00	1.00	1.02	1.36	1.00	1.01	1.01
	(.09)	(.08)	(.07)	(.24)	(.09)	(.09)			(.59)	(.09)		
ψ_3: Residual variance for Beer Consumption factor	.59	.51	.30	.52	.55	.27	.64	.57	.55	.68	.74	.61
	(.09)	(.12)	(.23)	(.20)	(.09)	(.08)			(.13)	(.10)		
ψ_4: Residual variance for Marijuana Use factor	.41	.40	.34	.21	.40	.27	.39	.18	.41	.43	.42	.40
	(.06)	(.06)	(.14)	(.07)	(.05)	(.05)			(.09)	(.06)		
ψ_5: Covariance between Law Abidance and Self-Acceptance factors	.16	.11	.08	.29	.18	.24	.22	.33	.18	.11	.15	.16
	(.06)	(.06)	(.05)	(.09)	(.06)	(.06)			(.09)	(.06)		
χ^2	102.01	81.99	92.14	60.95	177.96	475.44	...	...	66.81	140.13	...	...
df	30	30	30	29	30	30	...	...	29	30	...	...
P	.000	.000	.000	.000	.000	.000	...	...	.000	.000	...	...

NOTE.—Standard errors are given in parentheses. Estimates with asterisks denote fixed values.

[a] The "A" subscript in DV_A, $POLY_A$, and $TETRA_A$ indicates that the variables were transformed by an arbitrary split of the response categories, while the "U" subscript refers to variables transformed by a uniform split of the response categories.

TABLE 3

PARAMETER ESTIMATES AND STANDARD ERRORS FOR ALTERNATE SOLUTION OF EXAMPLE 1

Parameter and Meaning	ML	GLS	ADF	DV_A[a]	$POLY_A$-ML	$POLY_A$-ULS	$TETRA_A$-ULS	DV_U[a]	$POLY_U$-ML	$POLY_U$-ULS	$TETRA_U$-ULS
λ_1: Factor loading for Law Abidance on Law Abidance factor	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*
λ_2: Factor loading for Liberalism on Law Abidance factor	−.45 (.12)	−.25 (.11)	−.21 (.10)	−.29 (.14)	−.47 (.12)	−.47	−.52	−.32 (.13)	−.40 (.13)	−.25	−.35
λ_3: Factor loading for Religious Commitment on Law Abidance factor	.42 (.12)	.37 (.12)	.44 (.11)	.45 (.14)	.34 (.11)	.37	.58	.41 (.13)	.42 (.13)	.37	.47
λ_4: Factor loading for Self-Acceptance on Self-Acceptance factor	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*
λ_5: Factor loading for Invulnerability on Self-Acceptance factor	.31 (.06)	.38 (.06)	.45 (.06)	.43 (.10)	.31 (.06)	.22	.28	.13 (.10)	.31 (.06)	.25	.23
λ_6: Factor loading for Depression on Self-Acceptance factor	.69 (.05)	.63 (.05)	.59 (.04)	.58 (.13)	.71 (.04)	.73	.75	.35 (.24)	.71 (.04)	.71	.67
λ_7: Factor loading for Beer Frequency on Beer Consumption factor ..	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*
λ_8: Factor loading for Beer Quantity on Beer Consumption factor ..	.91 (.07)	.88 (.06)	.82 (.05)	1.01 (.09)	.97 (.06)	.96	1.01	1.02 (.07)	.90 (.06)	.92	1.07
λ_9: Factor loading for Marijuana Frequency on Marijuana Use factor	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*
λ_{10}: Factor loading for Marijuana Quantity on Marijuana Use factor	1.02 (.07)	1.03 (.07)	1.01 (.06)	.96 (.10)	1.06 (.05)	1.06	1.14	.97 (.05)	.98 (.04)	1.00	.92
β_1: Path from Law Abidance factor to Beer Consumption factor	−1.17 (.15)	−1.37 (.18)	−1.47 (.19)	−1.51 (.24)	−1.23 (.15)	−1.22	−1.55	−1.35 (.20)	−1.29 (.17)	−1.25	−1.38
β_2: Path from Law Abidance factor to Marijuana Use factor	−1.20 (.16)	−1.15 (.15)	−1.16 (.15)	−1.50 (.24)	−1.31 (.17)	−1.36	−1.77	−1.29 (.16)	−1.57 (.22)	−1.58	−1.75
β_3: Path from Self-Acceptance factor to Beer Consumption factor ...	.19 (.07)	.03 (.08)	.09 (.09)	.36 (.14)	.24 (.07)	.31	.60	−.05 (.08)	.11 (.08)	.13	.19
β_4: Path from Self-Acceptance factor to Marijuana Use factor	.16 (.07)	.06 (.07)	.04 (.07)	.24 (.13)	.15 (.07)	.25	.43	−.01 (.06)	.13 (.08)	.14	.21

θ_1: Error variance for Law Abidance	.61	.54	.54	.63	.62	.61	.68	.59	.69	.68	.73
	(.07)	(.06)	(.06)		(.06)				(.07)		
θ_2: Error variance for Liberalism	.92	.61	.57	.97	.92	.91	.91	.96	.95	.98	.97
	(.08)	(.07)	(.06)		(.08)				(.09)		
θ_3: Error variance for Religious Commitment	.93	.79	.83	.93	.95	.95	.89	.93	.95	.96	.94
	(.08)	(.08)	(.06)		(.09)				(.08)		
θ_4: Error variance for Self-Acceptance	.00*	.00*	.00*	−.34	.00*	.00*	.00*	−.90	.00*	.00*	.00*
θ_5: Error variance for Invulnerability	.91	.75	.72	.76	.90	.95	.92	.97	.90	.94	.94
	(.08)	(.07)	(.06)		(.08)				(.08)		
θ_6: Error variance for Depression	.52	.47	.44	.55	.50	.47	.44	.77	.49	.49	.55
	(.05)	(.04)	(.04)		(.04)				(.04)		
θ_7: Error variance for Beer Frequency	.14	.09	.04	.12	.17	.16	.22	.10	.10	.11	.21
	(.05)	(.05)	(.06)		(.04)				(.05)		
θ_8: Error variance for Beer Quantity	.28	.27	.30	.11	.22	.23	.20	.07	.26	.25	.10
	(.05)	(.04)	(.05)		(.04)				(.04)		
θ_9: Error variance for Marijuana Frequency	.20	.18	.12	.19	.16	.16	.32	−.03	.07	.09	.08
	(.05)	(.04)	(.04)		(.03)				(.03)		
θ_{10}: Error variance for Marijuana Quantity	.18	.15	.12	.26	.06	.05	.11	.04	.11	.09	.22
	(.05)	(.04)	(.04)		(.03)				(.03)		
ψ_1: Variance of Law Abidance factor	.39	.34	.32	.37	.38	.39	.32	.41	.31	.32	.27
	(.08)	(.07)	(.07)	(.08)	(.08)			(.09)	(.07)		
ψ_2: Variance of Self-Acceptance factor	1.00	.91	.87	1.34	1.00	1.00	1.01	1.90	1.00	1.01	1.01
	(.09)	(.08)	(.07)	(.25)	(.09)			(1.28)	(.09)		
ψ_3: Residual variance for Beer Consumption factor	.37	.23	.28	.19	.31	.35	.30	.13	.42	.42	.33
	(.08)	(.08)	(.08)	(.14)	(.07)			(.12)	(.08)		
ψ_4: Residual variance for Marijuana Use factor	.28	.32	.34	.11	.25	.22	.02	.33	.21	.16	.17
	(.07)	(.06)	(.06)	(.13)	(.06)			(.12)	(.08)		
ψ_5: Covariance between Law Abidance and Self-Acceptance factors	.18	.06	.08	.28	.20	.25	.35	.09	.13	.14	.17
	(.06)	(.06)	(.05)	(.09)	(.06)			(.09)	(.06)		
χ^2	114.81	85.72	92.15	62.30	192.56	...	...	70.42	155.75	...	...
df	31	31	31	30	31	...	...	30	31	...	...
P	.000	.000	.000	.000	.000	...	...	.000	.000	...	...

NOTE.—Standard errors are given in parentheses. Estimates with asterisks denote fixed values.

[a] The "A" subscript in DV_A, $POLY_A$, and $TETRA_A$ indicates that the variables were transformed by an arbitrary split of the response categories, while the "U" subscript refers to variables transformed by a uniform split of the response categories.

in the UCLA Study of Adolescent Growth (Huba & Bentler, 1982b, in press) are used. For this example there were three indicators of a latent variable of Law Abidance measured in Year I of the longitudinal study with two cannabis use (marijuana, hashish) indicators in Year I, one cannabis use indicator in Year II, two cannabis use indicators in Year IV, and three cannabis use indicators (marijuana frequency, marijuana quantity, hashish frequency) in Year V.

A diagrammatic representation of the model is given in Figure 2, which shows how the latent variable of Law Abidance in Year I relates to the sequence of Cannabis Use over the 5-year period of the study. Of the 11 variables used in the analysis, five have seemingly nonnormal distributions with high levels of kurtosis and skew, while the remaining variables have kurtosis levels indicative either of normal distributions or distributions that are too flat. The multivariate kurtosis is 117.24, which is highly significant.

Parameter estimates were obtained for the model under the ML, GLS, and ADF techniques. Furthermore, to obtain tetrachoric correlations, the drug use variables were dichotomized as "ever" versus "never," while the three indicators of Law Abidance were dichotomized at the mean. For the polychoric correlations, the drug use variables were retained in five categories, while the Law Abidance indicators were split so that there would be approximately even distributions among the five categories. Parameter estimates, standard errors, and goodness of fit are given in Table 4.

While the model does not fit the data under ML and GLS estimation (assuming normally distributed variables), it does fit when the data are analyzed in ADF or Muthen's DV method. Nonetheless, notice that the parameter estimates for the ML and GLS technique are about the same as those for the ADF estimator.

While the interpretation of the model is not a major point of the article, a few major conclusions might be noted. Examining the factor loadings, there is general agreement among methods that the indicators assess the factors as hypothesized. Examining the structural regression coefficients (parameters β_1 through β_7), it can be seen that ADF, GLS, and ML all agree on about the same value, and that in all cases the critical ratio test (dividing the parameter estimate by its standard

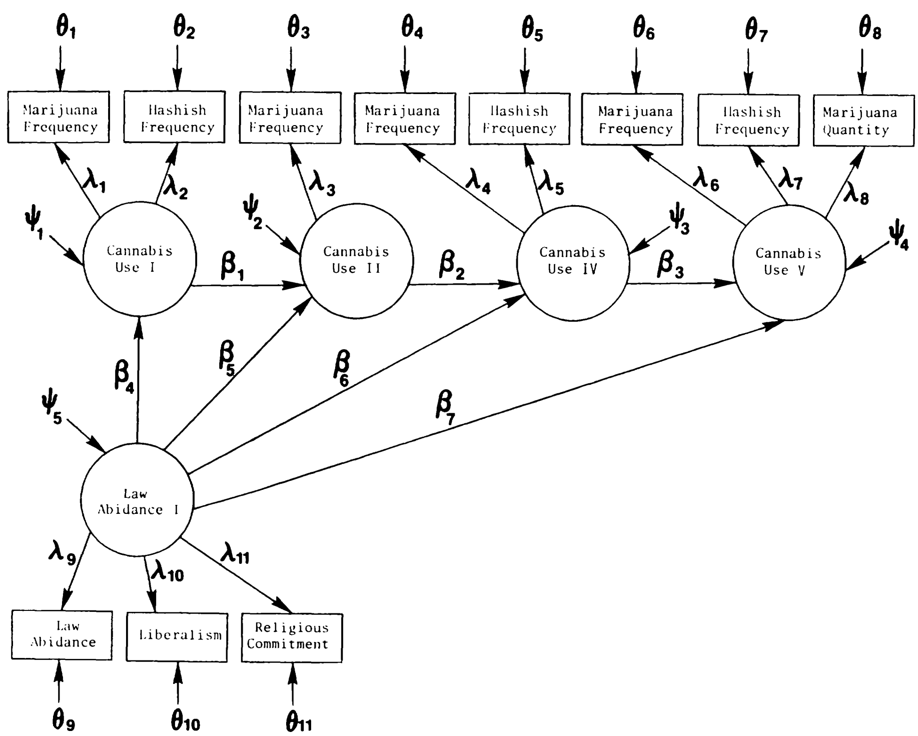

FIG. 2.—Cannabis use over 5 years and law abidance

TABLE 4

PARAMETER ESTIMATES AND STANDARD ERRORS FOR EXAMPLE 2

Parameter and Meaning	ML	GLS	ADF	DV	POLY-ML	TETRA-ML	POLY-ULS	TETRA-ULS
λ_1: Loading for Marijuana Frequency1 on Cannabis Use I factor	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*
λ_2: Loading for Hashish Frequency1 on Cannabis Use I factor	.56 (.07)	.56 (.07)	.45 (.08)	.96 (.08)	.83 (.04)	.86 (.04)	.90	.95
λ_3: Loading for Marijuana Frequency2 on Cannabis Use II factor	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*
λ_4: Loading for Marijuana Frequency4 on Cannabis Use IV factor	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*
λ_5: Loading for Hashish Frequency4 on Cannabis Use IV factor	.76 (.04)	.75 (.04)	.63 (.06)	.88 (.04)	.86 (.03)	.79 (.02)	.88	.83
λ_6: Loading for Marijuana Frequency5 on Cannabis Use V factor	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*
λ_7: Loading for Hashish Frequency5 on Cannabis Use V factor	.81 (.03)	.80 (.03)	.77 (.04)	.96 (.02)	.90 (.02)	.91 (.02)	.94	.94
λ_8: Loading for Marijuana Quantity5 on Cannabis Use V factor	.91 (.03)	.90 (.03)	.88 (.04)	1.02 (.01)	.93 (.02)	.98 (.01)	.95	.99
λ_9: Loading for Law Abidance1 on Law Abidance I factor	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*
λ_{10}: Loading for Liberalism1 on Law Abidance I factor	−.80 (.10)	−.82 (.10)	−.82 (.10)	−.58 (.12)	−.73 (.08)	−.56 (.08)	−.68	−.59
λ_{11}: Loading for Religious Commitment1 on Law Abidance I factor	.46 (.09)	.46 (.08)	.43 (.09)	.38 (.10)	.42 (.08)	.50 (.08)	.44	.46
β_1: Path from Cannabis Use I factor to Cannabis Use II factor	.23 (.07)	.25 (.08)	.27 (.09)	1.10 (.23)	.34 (.10)	.33 (.19)	.34	1.03
β_2: Path from Cannabis Use II factor to Cannabis Use III factor	.13 (.05)	.12 (.06)	.09 (.06)	.87 (.19)	.17 (.06)	.04 (.08)	.14	.42

TABLE 4 (*Continued*)

Parameter and Meaning	ML	GLS	ADF	DV	POLY-ML	TETRA-ML	POLY-ULS	TETRA-ULS
β_3: Path from Cannabis Use III factor to Cannabis Use IV factor	.75 (.05)	.73 (.05)	.70 (.08)	.87 (.07)	.80 (.04)	.67 (.04)	.77	.73
β_4: Path from Law Abidance I factor to Cannabis Use I factor	−.88 (.11)	−.89 (.11)	−.79 (.12)	−.89 (.23)	−1.23 (.11)	−1.40 (.11)	−1.33	−1.28
β_5: Path from Law Abidance I factor to Cannabis Use II factor	−.64 (.14)	−.64 (.14)	−.63 (.14)	.18 (.30)	−.58 (.19)	−.82 (.32)	−.66	.24
β_6: Path from Law Abidance I factor to Cannabis Use IV factor	−.75 (.12)	−.78 (.13)	−.78 (.14)	−.02 (.23)	−.81 (.11)	−1.23 (.16)	−.90	−.63
β_7: Path from Law Abidance I factor to Cannabis Use V factor	−.20 (.08)	−.24 (.08)	−.24 (.11)	.01 (.11)	−.14 (.06)	−.22 (.08)	−.23	−.19
θ_1: Error variance for Marijuana Frequency1	.09 (.10)	.10 (.09)	.02 (.12)	.22	.04 (.03)	.12 (.02)	.12	.20
θ_2: Error variance for Hashish Frequency1	.71 (.05)	.68 (.05)	.49 (.13)	.29	.34 (.03)	.34 (.03)	.28	.28
θ_3: Error variance for Marijuana Frequency2	.00*	.00*	.00*	.26	.00*	.00*	.00*	.00*
θ_4: Error variance for Marijuana Frequency4	.11 (.03)	.09 (.03)	.04 (.05)	−.02	.03 (.02)	−.09 (.02)	.05	−.04
θ_5: Error variance for Hashish Frequency4	.49 (.03)	.39 (.03)	.38 (.05)	.22	.27 (.02)	.31 (.02)	.26	.28
θ_6: Error variance for Marijuana Frequency5	.09 (.02)	.08 (.02)	.08 (.02)	.04	.01 (.01)	.01 (.00)	.05	.04

θ_7: Error variance for Hashish Frequency5	.40 (.03)	.34 (.02)	.33 (.03)	.12	.20 (.01)	.19 (.01)	.17	.15
θ_8: Error variance for Marijuana Quantity5	.24 (.02)	.24 (.02)	.24 (.03)	.00	.15 (.01)	.05 (.00)	.15	.05
θ_9: Error variance for Law Abidance1	.62 (.05)	.61 (.05)	.60 (.06)	.38	.60 (.05)	.64 (.04)	.65	.63
θ_{10}: Error variance for Liberalism1	.76 (.05)	.73 (.05)	.72 (.05)	.79	.79 (.05)	.89 (.05)	.84	.87
θ_{11}: Error variance for Religious Commitment1	.92 (.06)	.90 (.05)	.91 (.05)	.91	.93 (.05)	.91 (.05)	.93	.92
ψ_1: Residual variance for Cannabis Use I factor	.62 (.11)	.58 (.10)	.58 (.15)	.29 (.14)	.35 (.06)	.18 (.04)	.26	.19
ψ_2: Residual variance for Cannabis Use II factor	.70 (.05)	.68 (.05)	.64 (.07)	.00*	.56 (.04)	.40 (.04)	.53	.33
ψ_3: Residual variance for Cannabis Use IV factor	.60 (.06)	.61 (.06)	.62 (.07)	.43 (.07)	.56 (.04)	.51 (.04)	.55	.52
ψ_4: Residual variance for Cannabis Use V factor	.29 (.03)	.30 (.03)	.31 (.04)	.21 (.04)	.28 (.02)	.34 (.02)	.23	.28
ψ_5: Variance of Law Abidance I factor	.38 (.06)	.38 (.06)	.34 (.06)	.62 (.17)	.40 (.06)	.36 (.05)	.35	.37
χ^2	102.61	90.84	50.22	44.70	197.15	1796.52	...	...
df	38	38	38	38	38	38		
P	.000	.000	.089	.211	.000	.000		

NOTE.—Standard errors are given in parentheses. Estimates with asterisks denote fixed values. (Note that θ_3 was fixed to zero in order to identify the model as variable 3 was a single indicator of the Cannabis Use II factor; likewise, ψ_2 was set to zero in the DV case.)

error) indicates the same decision about whether the parameter is significant or not. However, the DV method of Muthen tends to yield substantially different conclusions about the paths from Law Abidance I to Canabis Use II, IV, and V, with the DV method indicating that the paths are not significant while all other methods do suggest that the paths are significant. In this case, dichotomization of the variables may throw away so much information that these significant effects cannot be detected. Similar values are found for all of the error and residual variances. In the approximate TETRA-ML and TETRA-ULS methods, a negative error variance is again found.

Robust versus Resistant Modeling Techniques

The procedures that have been discussed thus far are new, revolutionary statistical estimators that are robust over violations of the assumption of multivariate normality. That is, the Browne ADF estimator can, with sufficiently large samples, be expected to perform reasonably well when the data are not normally distributed but are continuous. Assuming certain facts about the underlying latent distributions, Muthen's procedures are robust over problems that are caused by the observed data being dichotomous or measured in ordered categories.

Frequently the word "robust" is used by applied data analysis workers to refer also to the fact that ideally we would like a statistical procedure to be relatively unaffected by a gross pathology in the data like a miscoded subject or an outlier case, or we would like most of the parameters estimated in a model to have values that are minimally affected by some local areas of misspecification, such as neglecting to include a necessary parameter. Such techniques are called *resistant* ones here. Basically a resistant technique is one that should not be unduly affected by one or a few "bad" observations, or poor specification in some small part of the model.

Thus far there has been relatively little work on resistant estimation for structural equation modeling parameters, although the techniques are relatively well known in other areas of statistics and, most notably, in the related area of multiple linear regression (see Mosteller & Tukey, 1977). The exception to this rule has been some pioneering work by Browne (1982) that applies resistant techniques to structural equation modeling methods.

In resistant modeling, two different approaches might be contrasted. In the first approach, we would estimate a set of causal modeling parameters on the data as they occur in the data file, but we would somehow weight the observations differentially when we were actually minimizing some function so that the parameters would be based on fit to data where each observation did not count equally. That is, data that were relatively different from the other data would not count as heavily. Thus, outlier subjects would be counted less, or not at all, and accordingly it would be expected that the parameter estimates obtained in such a procedure would be reasonably stable against data anomalies such as a mispunched record, or a client who deliberately sabotaged responses to a questionnaire. In such an approach, the major issue is whether to develop the weights from discrepancies between observed and reproduced covariance matrices or between observed and reproduced raw data and what weighting scheme to use. Browne (1982) adopts this general approach to resistant fitting in causal modeling by basing his weighting scheme on the bisquare procedure of Mosteller and Tukey (1977). Browne derives his weights from the discrepancies between observed and reproduced covariance matrix elements. This is the most computationally effective technique. In the Browne resistant fitting procedure. there will be a tendency to fit most of the elements in the covariance matrix very well but to leave a few very large discrepancies that would tend to be attributed either to gross data pathologies or bad specifications of the theoretical model.

A second approach to resistant fitting might be to argue that since structural equation modeling techniques are methods for determining whether a model adequately describes a *covariance matrix*, one might use some weighting scheme to develop a resistant covariance matrix estimate and then use this robust covariance matrix as the input to a "regular" structural equation modeling procedure. Again, any number of weighting schemes might be used here (see, e.g., Huber, 1981). The first author has experimented extensively with the estimation of a resistant covariance matrix using the Mosteller-Tukey bisquare weight scheme (see Huba & Bentler, in press) and found that in general very similar results are found for structural regressions based on either the "regular" or the "resistant" covariance matrix in some data sets that had been carefully cleaned of errant observations. Jöreskog and Sörbom (1984) discuss a similar, although computationally slightly different, approach.

It might be noted that Tanaka (1987, in this issue) discusses some resistant fitting methods that are especially applicable for small data sets. Tanaka's recommended procedures are most like the notion of developing a resistant estimate of the covariance matrix where "extra" correlation due to the small sample size has been reduced through statistical manipulation.

In theory, both of these resistant estimation techniques should yield generally comparable results, although an "outlier" observation would tend to be identified as an "outlier covariance-producing element" in the first case and as a "pairwise outlier" in the second case. It is not known in practice if there are differences between the two procedures when the data are dirty to different degrees.

Resistant structural equation modeling techniques are especially useful when large and potentially "dirty" data sets are to be used in a causal modeling example. In such cases, a few aberrant observations could potentially bias the results *whether or not a robust modeling technique such as ADF is used.* In general, it seems unreasonable to expect that statistically based methods for nonnormal data would be strongly effective against gross abnormalities such as would be caused by bad outlier cases. However, in general, we will want to get the best statistical estimates possible, so a robust method may be preferable over a (nonstatistical) resistant technique.

It is the position of the authors that resistant techniques for causal modeling parameter estimation need to be studied in much greater detail. Nonetheless, it may be that the major use of resistant fitting procedures will be to verify that the same results can be obtained as have been gotten using a "regular" robust method such as ADF. That is, if we can run the data through both ADF (or ML or GLS) estimation and find almost exactly the same parameter estimates as we do when we run the data through a resistant fitting method such as Browne's (1982) bisquare weighting scheme, then we probably would want to give great credence to claims of validity for the statistical estimator in the data set. On the other hand, if the resistant estimates of parameters depart greatly from those obtained in the statistical methods of ML, GLS, or ADF, then we might want to carefully examine the raw data set and eliminate "bad" data points that can be identified as the result of data miscoding or poor keying of the data. That is, the comparison of the results from robust and resistant parameter estimation techniques in causal modeling may serve to indicate whether or not a data set should be cleaned carefully again or not. Unfortunately, various resistant estimators for causal modeling parameters and correlations/covariances are not generally available in widely circulated computer programs for structural equations modeling.

Fit Coefficients for Robust and Resistant Modeling Methods

As noted by Tanaka and Huba (1985), it has frequently been argued that statistical indices of the fit of structural equation models to data tend to emphasize that not all of the covariation has been explained as opposed to how much covariation is accounted for. Several alternate types of fit coefficients, or correlation-like indices of amount of variance accounted for, have been proposed for causal models. Of these, the general coefficient discussed by Tanaka and Huba (1985) as modified from work by Jöreskog and Sörbom (1981) seems to be the most applicable to the robust structural equation modeling techniques that are discussed here. The generalized "goodness of fit" or GFI index seems most appropriate for two reasons. First, Tanaka and Huba (1985) show that a general form of the GFI index can be demonstrated to have an optimal value when the causal modeling fit functions reach their minimum. That is, causal modeling techniques that minimize chi-square or chi-square-like coefficients will maximize GFI coefficients. Second, from this general result, Tanaka and Huba were able to show that specific coefficients for maximum-likelihood, generalized least-squares, asymptotically distribution-free, and unweighted least-squares estimation can be derived. Thus, the general coefficient is appropriate for these different estimators, and about the same metric for the coefficient applies irrespective of the method of parameter estimation. That is, GFI coefficients obtained from different methods of parameter estimation can be compared to one another.

Discussion

A very old criticism in latent-variable causal modeling or structural equation modeling has been to state that maximum-likelihood parameter estimates and goodness-of-fit statistics are derived under the assumption of multivariate normality, go on to point out that any reasonable person knows that this or that variable cannot possibly be normally distributed, and then go on to snicker. Thus, many existing causal models in the literature are frequently thought to be somewhat suspect because the variables did not religiously ring the normal bell. In response to the supposed restrictions of the maximum-likelihood method of parameter estimation, and with a full realization that social scientists often have their hands tied in the measurement arena by considerations opposite to those that might guide the development of measurement instruments with nice bell-shaped distributions, statisticians have suggested some elegant alternate estimation procedures. These extremely promising newer techniques, such as Browne's asymptotically distribution-free (ADF) and Muthen's dichotomous and polytomous estimation techniques, appear to be the methods of choice when their requirements of very large samples can be met. Monte Carlo (random-number) studies are still needed to determine the statistical power and bias properties of these methods when they are used in small samples, but for large samples these alternate techniques represent a very major advance in the statistical theory of causal modeling.

Of course, it is also important to ask if older models in the literature that had been estimated with maximum-likelihood parameter estimates can be "trusted" in their major features. A number of investigations have compared the results of the newer estimators with those of ML estimation for many "real" developmental problems, and it has generally been concluded that the parameter estimates are about the same, although the global goodness-of-fit chi-square values for the model and the standard errors for the parameter estimates may differ somewhat when the data are not normally distributed (Huba & Bentler, 1983a, in press; Huba & Harlow, 1983, 1986; Huba & Tanaka, 1983). What these comparative studies seem to illustrate is what many methodologists have been saying about causal models for a number of years (at least as evinced by their behavior): when the data are not normally distributed, trust the ML parameter estimates but not necessarily the goodness-of-fit statistic or the standard errors for the individual parameter estimates. Using ML estimates with data that are nonnormal can conceivably add some extra junk parameters like correlated errors to a "fitting model," although the major parameters are generally quite stable (Browne, 1982; Huba & Bentler, in press; Huba, Wingard, & Bentler, 1981).

As the next generation of computer programs for structural equation modeling are developed in the ensuing decade, it is likely that the developmental psychologist will be offered a number of options about how the parameters in the model are estimated. Browne's (1982, 1984) general framework will allow for a number of elaborations based on modified estimators (see Tanaka, 1984, for one example and the first Monte Carlo evaluation of ADF estimation). As these newer computer programs become widely available on computing equipment that is progressively faster and less costly to use, developmental psychologists will finally be able to choose to use statistical modeling techniques that do not force them to assume that data which are clearly not normally distributed are in fact normally distributed.

In conclusion, it should be noted that structural equation modeling methods can be specialized to most commonly used multivariate analysis techniques, and especially those of a statistical nature, such as the multivariate analysis of variance, discriminant analysis, canonical correlation analysis, and multiple linear regression. It is quite likely that the next few years will also see the development of computer programs for robust canonical correlation analysis, robust linear regression, and robust discriminant analysis using Browne and Muthen estimators as well as further refinements encompassing logistic regression estimators (Tanaka & Huba, 1986). Finally, developmental psychologists will have an arsenal of statistical tools that permit us to assume that the data have the distributional form that the data actually have. These statistical and computational developments of Browne (1982, 1984) and Muthen (1984) will do much to aid the accurate assessment of important developmental psychological theories through statistical model testing.

References

Bentler, P. M. (1985). *Theory and implementation of EQS: A structural equations program.* Los Angeles: BMDP Statistical Software.

Bentler, P. M. (1987). Drug use and personality in adolescence and young adulthood: Structural models with nonnormal variables. *Child Development*, **58**, 65–79.

Browne, M. W. (1968). A comparison of factor analytic technique. *Psychometrika*, **33**, 267–334.

Browne, M. W. (1974). Generalized least squares estimators in the analysis of covariance structures. *South African Statistical Journal*, **8**, 1–24.

Browne, M. W. (1982). Covariance structures. In D. M. Hawkins (Ed.), *Topics in applied multivariate analysis* (pp. 72–141). Cambridge: Cambridge University Press.

Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, **37**, 62–83.

Browne, M. W., & Cudeck, R. (1983). *BENWEE: A computer programme for path analysis with latent variables*. Pretoria: HSRC.

Carroll, J. B. (1961). The nature of data, or how to choose a correlation coefficient. *Psychometrika*, **26**, 347–372.

Fuller, E. L., Jr., & Hemmerle, W. J. (1966). Robustness of the maximum-likelihood estimation procedure in factor analysis. *Psychometrika*, **31**, 255–266.

Huba, G. J., & Bentler, P. M. (1982a). On the usefulness of latent variable causal modeling in testing theories of naturally occurring events (including adolescent drug use): A rejoinder to Martin. *Journal of Personality and Social Psychology*, **43**, 604–611.

Huba, G. J., & Bentler, P. M. (1982b). A developmental theory of drug use: Derivation and assessment of a causal modeling approach. In P. M. Baltes & O. G. Brim, Jr. (Eds.), *Life-span development and behavior* (Vol. **4**, pp. 147–203). New York: Academic Press.

Huba, G. J., & Bentler, P. M. (1983a). Test of a drug use causal model using asymptotically distribution free methods. *Journal of Drug Education*, **13**, 3–14.

Huba, G. J., & Bentler, P. M. (1983b). Causal models of the development of law abidance and its relationship to psychosocial factors and drug use. In W. S. Laufer & J. M. Day (Eds.), *Personality theory, moral development, and criminal behavior* (pp. 165–215). Lexington, MA: Heath.

Huba, G. J., & Bentler, P. M. (in press). *Antecedents and consequences of adolescent drug use: A psychosocial study of development using a causal modeling approach*. New York: Plenum.

Huba, G. J., & Harlow, L. L. (1983). Comparison of maximum likelihood, generalized least squares, ordinary least squares, and asymptotically distribution free estimates in drug abuse latent variable causal models. *Journal of Drug Education*, **13**, 387–404.

Huba, G. J., & Harlow, L. L. (1986). Robust estimation for causal models: A comparison of methods in some development datasets. In P. B. Baltes, D. L. Featherman, & R. M. Lerner (Eds.), *Life-span development and behavior* (Vol. **6**, pp. 69–111). Hillsdale, NJ: Erlbaum.

Huba, G. J., & Tanaka, J. S. (1983). Confirmatory evidence for three daydreaming factors in the Short Imaginal Processes Inventory. *Imagination, Cognition, and Personality*, **3**, 137–145.

Huba, G. J., Wingard, J. A., & Bentler, P. M. (1981). A comparison of two latent variable causal models for adolescent drug use. *Journal of Personality and Social Psychology*, **40**, 180–193.

Huber, P. J. (1981). *Robust statistics*. New York: Wiley.

Jöreskog, K. G. (1967). Some contributions to maximum likelihood factor analysis. *Psychometrika*, **34**, 443–482.

Jöreskog, K. G., & Goldberger, A. S. (1972). Factor analysis by generalized least squares. *Psychometrika*, **37**, 243–260.

Jöreskog, K. G., & Sörbom, D. (1981). *LISREL V: Analysis of linear structural relationships by the method of maximum likelihood*. Chicago: National Educational Resources.

Jöreskog, K. G., & Sörbom, D. (1984). *LISREL VI: Analysis of linear structural relationships by the method of maximum likelihood*. Chicago: National Educational Resources.

McDonald, R. P. (1980). A simple comprehensive model for the analysis of covariance structures: Some remarks on applications. *British Journal of Mathematical and Statistical Psychology*, **33**, 161–183.

Mosteller, F., & Tukey, J. W. (1977). *Data analysis and regression*. Reading, MA: Addison-Wesley.

Muthen, B. (1978). Contributions to factor analysis of dichotomous variables. *Psychometrika*, **43**, 551–560.

Muthen, B. (1981). Factor analysis of dichotomous variables: American attitudes toward abortion. In E. Borgatta & D. J. Jackson (Eds.), *Factor analysis and measurement in sociological research: A multidimensional perspective* (pp. 201–214). San Francisco: Sage.

Muthen, B. (1982a). Some categorical response models with continuous latent variables. In K. G. Jöreskog & Wold (Eds.), *Systems under indirect observation: Causality, structure, prediction* (pp. 65–80). Amsterdam: North-Holland.

Muthen, B. (1982b). *LACCI: Latent variable analysis of dichotomous, ordered categorical, and continuous indicators: An experimental program for researchers*. Los Angeles: University of California at Los Angeles.

Muthen, B. (1983). Latent variable structural equation modeling with categorical data. *Journal of Econometrics*, **22**, 43–65.

Muthen, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, **49**, 115–132.

Muthen, B. (1986). *LISCOMP. International Educational Statistics*. Evanston: Indiana.

Muthen, B., Huba, G. J., & Short, L. (1985). *Applications of LISCOMP structural equation modeling for ordered categorical variables*. Paper presented at the American Psychological Association Meetings, Los Angeles.

Muthen, B., & Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of skewed, ordered categorical variables. *British Journal of Mathematical and Statistical Psychology*, forthcoming.

Muthen, B. (1986). LISCOMP. Mooresville, IN: Scientific Software.

Olsson, U. (1979). On the robustness of factor analysis against crude classification of the observations. *Multivariate Behavioral Research*, **14**, 485–500.

Tanaka, J. S. (1984). *Some results on the estimation of covariance structure models*. Doctoral dissertation, University of California at Los Angeles.

Tanaka, J. S. (1987). How big is big enough? The sample size issue in structural equation modeling. *Child Development*, **58**, 134–146.

Tanaka, J. S., & Huba, G. J. (1985). A fit index for covariance structure models under arbitrary GLS estimation. *British Journal of Mathematical and Statistical Psychology*, **38**, 197–201.

Tanaka, J. S., & Huba, G. J. (1986). *Rasch and quasi-Rasch scalability of health and social behaviors*. Under editorial consideration.