Assumptions of the Two-Step Approach

Reply to Anderson and Gerbing

Claes Fornell
Youjae Yi
University of Michigan

The purpose of our article (Fornell and Yi 1992 [this issue]) was to identify the assumptions implicitly made under two-step approaches and to evaluate the extent to which they can be met. In total, four sets of assumptions were discussed. In a response to our article, Anderson and Gerbing (1992 [this issue]) dismiss all of them as a result of misunderstandings and incomprehension on our part. Although we do admit to a very different understanding of basic statistical principles, we do not believe that Anderson and Gerbing have made a convincing case for the merits of the two-step approach.

A1. THEORY-MEASUREMENT INDEPENDENCE

Anderson and Gerbing proclaim that the two-step approach does not make this assumption. However, the two-step approach consists of two separate steps: an examination of the "measurement portion" isolated from the structural part, and a test of the "substantive theory." Such separation of measurement and theory must be based on an implicit assumption that "theory and measurement are independent of one another or can be treated as such" (Fornell and Yi 1992 [this issue], p. 295), whether it is acknowledged or not. Citing themselves or appealing to "a philosophy of science of *epistemological fallibilism*" (Anderson and Gerbing 1992 [this issue], p. 323) does not change the state of affairs. Neither does the fact that confirmatory factor analysis

SOCIOLOGICAL METHODS & RESEARCH, Vol. 20, No. 3, February 1992 334-339

is theory-based (including an implicit but loose specification of "the substantive" theory).

Anderson and Gerbing also comment on our illustration of the interdependence between measurement and theory (Fornell and Yi 1992 [this issue], Figure 1). First, they claim that finding an equally fitting model is not surprising and that the alternative specification is atheoretical. However, this claim misses the point. Our example was not intended to show an alternative model with comparable fit or to find a true model. Its purpose was to illustrate that changing the structural model (theory) can produce changes in the parameters and reliabilities of the measurement model.

Second, Anderson and Gerbing discard the example as interpretational confounding (Burt 1976). However, it is not useful to dismiss theory-measurement interdependence as "interpretational confounding." Instead, theory-measurement interdependence needs to be explicitly acknowledged. Only then can we begin to understand the building blocks of construct validity and how theory and data are linked together.

Third, they claim that our example illustrates "the known instability and problematic nature of estimated constructs defined by two indicators" (Anderson and Gerbing 1992 [this issue], p. 324). It is true that the number of indicators affects the stability of loadings, but we are now considering modeling in practice. The researcher often has only a limited number of measures available. For example, a construct with only two indicators was used in the original model by Anderson and Narus (1984). No construct had more than two indicators in their extended model (Anderson and Narus 1990, Figure 3).

A2. MEASUREMENT VALIDITY FROM THE FIRST STEP CAN BE GENERALIZED

Anderson and Gerbing contend that measurement quality or measurement error is not dependent on the specification of a structural submodel. Two points can be made. First, by definition, the estimation of parameters (including error variances) is dependent on the whole

specification (including the structural model) under full information estimation procedures. Second, our illustration demonstrates that different structural models can produce different conclusions about the quality of one and the same measurement. Bagozzi and Fornell (1989) have also shown that acceptable unidimensionality is not sufficient for assuming that factor analysis loadings will hold in other specifications.

Anderson and Gerbing make several comments regarding our example on the theory of reasoned action. They assert that Model 2 is equivalent to Model 3. Furthermore, they claim that this "is true of any structural model that uses these same four latent variables, irrespective of whether or not the model is recursive or the latent variables are ordered in the same direction" (Anderson and Gerbing 1992 [this issue], p. 325). The implication is that Models 1 and 3 are equivalent, which is difficult to accept. From a theoretical viewpoint, Models 1 and 3 are different. Model 1 posits that attitudes affect behavior indirectly through intentions, whereas Model 3 predicts that attitudes can influence behavior directly. From an empirical viewpoint, the two models are not identical either. They do not provide the same fit, as was illustrated in our original article and is documented in the literature (e.g., Bagozzi, Baumgartner, and Yi 1989; Fredricks and Dossett 1983).

A3. ESTIMATORS HAVE DESIRABLE PROPERTIES

Anderson and Gerbing accept our argument from the viewpoint of statistical theory but raise several issues from the viewpoint of *practice*. First, they object to our argument on the ground that the estimators will be biased if the (parsimonious) model is not the true model. This is hardly surprising because most statistical tests (including chi-square tests) are based on the assumption that a model under consideration is correct. Also, the bias problem was noted and illustrated mainly for the two-step approach using *separate* factor analysis. For the two-step approach using confirmatory factor analysis, we noted and illustrated the inefficiency problem. Unless the saturated theoretical model is the true model, which is rare in practice, the confirmatory measurement model will have superfluous parameters and suffer from inefficiency.

Second, they argue that the practical extent of the differences in standard errors is unknown. However, there is no reason to believe that latent variable models would be more immune to the addition of parameters than would multiple regression. Estimating all possible links between latent variables is somewhat akin to step-wise regression and runs the risk of adding sampling error more than anything else. Because all latent variable correlations are nonzero for the particular model (i.e., Model 1) used in our article, Anderson and Gerbing contend that these correlations are not superfluous. Such an argument lacks generality because one could easily construct an example in which some of the latent variable correlations were zero.

A4. TESTS ARE INDEPENDENT

Anderson and Gerbing defend their recommendation that respecification and testing can be done on the same data by (a) letting on that we focused on the incorrect test (sequential tests instead of sequential difference tests), (b) reiterating results from Steiger, Shapiro, and Browne (1985), and (c) pronouncing us incorrect in our assertion that either Model 1 or Model 3 is true, but not both.

There is no need to respond to the first point. Simple reading of our article will do. However, the practice of model respecification and conditioning of statistical tests by knowledge of the data goes to the essence of statistical theory. Our understanding of the basic principles of statistical hypothesis testing is clearly at odds with that of Anderson and Gerbing. One such principle is that probabilistic interpretations cannot be given to a result obtained from a series of respecifications within the same data. It does not matter if sequential difference tests have been used. For example, after a series of respecifications of the initial measurement model, Anderson and Narus (1984) concluded that "The resulting measurement model retained 14 of the 24 measures in the initial model, and provided an *acceptable fit* ($\chi^2(71) = 81.36, p = .19$)" (p. 68, emphasis added). In the second step, Anderson and Narus (1984) again conclude that the modified theoretical model "provided an *acceptable fit* to the data ($\chi^2(74) = 89.21, p = .11$)" (p. 68, emphasis added). Note that chi-square tests are used and interpreted

as a conventional inferential statistic in deciding acceptable fit, ignoring the fact that a search process within the data tends to yield an inflated fit for a modified model.

Anderson and Gerbing assert that Model 1 and Model 3 can be correct in the sense that both approximate a true model. Although such reasoning may, under certain circumstances, have merit, it too is incompatible with extant statistical theory. Inferential statistics do not deal with approximate hypotheses. On a substantive basis, it is also difficult to accept the notion that Model 1 and Model 3 are both correct. It would mean that all research on the direct effect of attitudes on behavior was futile (e.g., Bentler and Speckart 1981; Fredricks and Dossett 1983).

CONCLUDING REMARKS

Our article has focused on the implicit assumptions associated with the two-step approach that have not been discussed by its proponents or examined in the literature. We have not elaborated on the difficulties with a one-step approach because they are well documented elsewhere. To evaluate the appropriateness of any research approach, its assumptions must obviously be identified. Anderson and Gerbing's prompt (and, in our view, mistaken) dismissal of every single assumption does not further their advocacy of the two-step approach and will not contribute to a balanced perspective of the relative strengths and weaknesses of a one-step versus a two-step approach. Only by a full accounting of the assumptions and properties associated with each will researchers be assisted in making decisions between them.

REFERENCES

Anderson, James C. and David W. Gerbing. 1992. "Assumptions and Comparative Strengths of the Two-Step Approach: Comment on Fornell and Yi." *Sociological Methods & Research* 20:321-333.

Anderson, James C. and James A. Narus. 1984. "A Model of the Distributor's Perspective of the Distributor-Manufacturer Working Relationships." *Journal of Marketing* 48:62-74.

———. 1990. "A Model of Distributor Firm and Manufacturer Firm Working Partnerships." *Journal of Marketing* 54:42-58.

Bagozzi, Richard P., Johann Baumgartner, and Youjae Yi. 1989. "An Investigation into the Role of Intentions as Mediators of the Attitude-Behavior Relationship." *Journal of Economic Psychology* 10:35-63.

Bagozzi, Richard P. and Claes Fornell. 1989. Consistency Criteria and Unidimensionality: An Attempt at Clarification. Pp. 321-25 in *Advances in Consumer Research*, edited by Thomas Srull. Ann Arbor, MI: Association for Consumer Research.

Bentler, Peter M. and George Speckart. 1981. "Attitudes Cause Behaviors: A Structural Equation Analysis." *Journal of Personality and Social Psychology* 40:226-38.

Burt, Ronald S. 1976. "Interpretational Confounding of Unobserved Variables in Structural Equation Models." *Sociological Methods & Research* 5:3-52.

Fornell, Claes and Youjae Yi. 1992. "Assumptions of the Two-Step Approach to Latent Variable Modeling." *Sociological Methods & Research* 20:291-320.

Fredricks, Arlene J. and Dennis L. Dossett. 1983. "Attitude-Behavior Relations: A Comparison of the Fishbein-Ajzen and the Bentler-Speckart models." *Journal of Personality and Social Psychology* 45:501-12.

Steiger, James H., Alexander Shapiro, and Michael W. Browne. 1985. "On the Multivariate Asymptotic Distribution of Sequential Chi-Square Statistics." *Psychometrika* 50:253-64.