DAVID W. GERBING and JAMES C. ANDERSON*

The authors outline an updated paradigm for scale development that incorporates confirmatory factor analysis for the assessment of unidimensionality. Under this paradigm, item-total correlations and exploratory factor analysis are used to provide preliminary scales. The unidimensionality of each scale then is assessed simultaneously with confirmatory factor analysis. After unidimensional measurement has been acceptably achieved, the reliability of each scale is assessed. Additional evidence for construct validity beyond the establishment of unidimensionality then can be provided by embedding the unidimensional sets of indicators within a nomological network defined by the complete structural model.

An Updated Paradigm for Scale Development Incorporating Unidimensionality and Its Assessment

The purpose of measurement in theory testing and development research is to provide an empirical estimate of each theoretical construct of interest. Because of the limitations inherent in single-item measures (cf. Churchill 1979), respondents usually are administered two or more measures, often referred to as a scale, that are intended to be alternative indicators of the same underlying construct. A composite score defined by the respondent's scores on these measures, generally calculated as an unweighted sum, provides an estimate of the corresponding construct. Our central thesis is that the computation of this composite score is meaningful only if each of the measures is acceptably unidimensional. Unidimensionality refers to the existence of a single trait or construct underlying a set of measures (Hattie 1985; McDonald 1981).[1] The importance of unidimensionality has been stated succinctly by Hattie (1985 p. 49): "That a set of items forming an instrument all measure just one thing in common is a most critical and basic assumption of measurement theory."

Because the meaning of a measure intended by the researcher may not be the same as the meaning imputed to it by the respondents, the scale development process must include an assessment of whether the multiple measures that define a scale can be acceptably regarded as alternative indicators of the same construct. Building on the earlier work of Churchill (1979) and Peter (1979, 1981), we outline an updated paradigm for scale development that incorporates confirmatory factor analysis (cf. Bentler 1985; Jöreskog and Sörbom 1984) for the evaluation of unidimensionality. The key aspect of this updated paradigm is that confirmatory factor analysis affords a stricter interpretation of unidimensionality than can be provided by more traditional methods such as coefficient alpha, item-total correlations, and exploratory factor analysis and thus generally will provide *different* conclusions about the acceptability of a scale.

*David W. Gerbing is Associate Professor of Business Administration, Department of Management, Portland State University. James C. Anderson is the William L. Ford Professor of Marketing and Wholesale Distribution, and Associate Professor of Behavioral Science in Management, J. L. Kellogg Graduate School of Management, Northwestern University.

Support for this research was provided in part by the McManus Research Chair awarded to the second author and is gratefully acknowledged.

[1]Measures that satisfy this definition of unidimensionality have been referred to as "congeneric" measurements (Jöreskog 1971) or "point" variables (Burt 1973).

Journal of Marketing Research
Vol. XXV (May 1988), 186–92

As background for understanding, we first briefly review the concept of unidimensionality, then present the confirmatory factor analysis model.

UNIDIMENSIONAL MEASUREMENT AND CONFIRMATORY FACTOR ANALYSIS

The mathematical definition of unidimensionality adopted here is based on the traditional common factor model in which a set of indicators share only a single underlying factor, ξ (McDonald 1981). Assuming linearity, the model for a single indicator can be given as

$$x_i = \lambda_i \xi + \delta_i \qquad (1)$$

where x_i is the i^{th} indicator from the set of unidimensional indicators, λ_i is the corresponding factor loading, and δ_i is the corresponding regression residual, uncorrelated with any factors or residuals (Jöreskog and Sörbom 1984). Because most studies investigate posited relationships among several constructs, they are characterized by *several* sets of postulated unidimensional measures. The relationship of these measures to their respective constructs is represented formally by a multiple-indicator measurement model in which each estimated construct is defined by at least two indicators and each indicator is intended as an estimate of only one construct (Anderson and Gerbing 1982; Hunter and Gerbing 1982).

Two criteria follow directly from equation 1, expressed as product rules; each represents a necessary condition for unidimensionality. The correlation of two indicators, i and j, of the same construct, ξ, is described by the product rule for internal consistency:[2]

$$\rho_{ij} = \rho_{i\xi}\rho_{j\xi}. \qquad (2)$$

The correlation of two indicators, i and p, where p is an indicator of another construct ξ^*, is described by the product rule for external consistency:

$$\rho_{ip} = \rho_{i\xi}\rho_{\xi\xi^*}\rho_{p\xi^*}. \qquad (3)$$

When ξ is the same as ξ^*, equation 3 reduces to equation 2; that is, internal consistency represents a special case of external consistency. A useful methodological consideration follows from an analysis of external consistency: indicators from *other* scales provide a means by which to assess unidimensionality of the items that define the given scale. Simply stated, given equation 3, measures that are truly alternate indicators of the same underlying construct will evince a parallel pattern of relationships, to within sampling error, with all other measures in the set.

[2]Note that equation 2 represents a generalization of the parallel test model from classical test theory (Lord and Novick 1968) in which

$$\rho_{ii'} = \rho_{i\xi}^2$$

where i and i' are parallel measures. Unidimensional measures are more general in that they may have unequal true score variances and unequal error variances (cf. Jöreskog 1971).

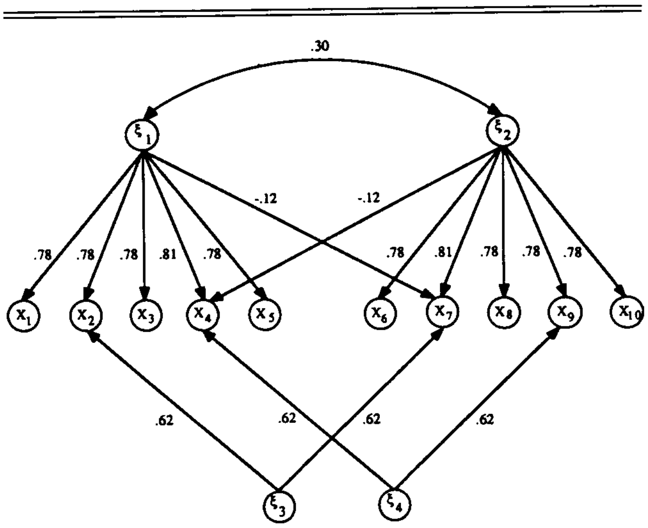
A PARADIGM FOR DEVELOPING UNIDIMENSIONAL SCALES

The development and evaluation of measurement scales traditionally have relied on one or more of the following analyses: coefficient alpha, item-total correlations, and exploratory factor analysis. The use of confirmatory factor analysis for building and evaluating measurement scales is a relatively recent development. Because each of these four methods is based on a different criterion, *different* conclusions may be drawn depending on the choice of method. Our key point is that, of these methods, only a confirmatory factor analysis of a multiple-indicator measurement model directly tests unidimensionality according to the definition provided in equation 1. Using the obtained parameter estimates, one can compute the indicator correlations predicted by the model with the product rules for internal consistency and external consistency. The overall goodness of fit is evaluated according to the similarity of the predicted with the actual correlations.

An Illustrative Model

For pedagogical clarity, we illustrate this updated paradigm with synthetic data. The model in Figure 1 represents the underlying known structure that generates these data. This model specifies two moderately correlated factors (ξ_1 and ξ_2). Each factor is indicated primarily by five items, though one item in each set also indicates to a smaller degree the other factor. Each of the two additional factors (ξ_3 and ξ_4) provides a source of common covariance for two pairs of items across the two sets. Substantively, ξ_3 and ξ_4 might represent either unwanted constructs or method factors (Gerbing and Anderson 1984).

Figure 1
MODEL UNDERLYING CORRELATION MATRIX IN TABLE 2



This measurement model was chosen to illustrate two types of violations of multiple-indicator measurement. First, unidimensional measurement for *each* set of items as formalized by a multiple-indicator measurement model could not be achieved successfully until the items that indicate both ξ_1 and ξ_2, x_4 and x_7, are deleted from the model. Second, deleting these items also means that the other factors, ξ_3 and ξ_4, are no longer common factors. Once x_4 and x_7 are removed, neither ξ_3 nor ξ_4 provides sources of common variance across two or more items, illustrating that the composition of indicator-specific variance is defined in relation to the particular set of indicators in the analysis. The initial common factors ξ_3 and ξ_4 would be redefined as specific factors that are indistinguishable in the resulting factor analysis from the random error variance of classical test theory (Gerbing and Anderson 1984; Smith 1974). That is, deleting x_4 and x_7 from the model means that ξ_3 and ξ_4 will contribute only to the uniquenesses (Harman 1976) of the affected indicators, x_2 and x_9. Related to this, deleting x_4 and x_7 means that ξ_3 and ξ_4 contribute only to the variances of x_2 and x_9.

Using the product rules for internal consistency and external consistency, in conjunction with the specified parameter values, we generated the population item correlation matrix in Table 1 from the model in Figure 1. A preliminary examination of the patterning of the correlations, without knowledge of the underlying structure, is consistent with the hypothesis that the 10 items potentially could represent two distinct scales. The first five indicators all correlate .61 with one another, as do the last five indicators. The mean correlation between the indicators in these two sets is only .19.

Item-Total Correlations

The use of item-total correlations in the construction of unidimensional scales has been long advocated. For example, Nunnally (1978, p. 274) first establishes the need for unidimensionality by stating, "Items within a measure are useful only to the extent that they share a common core—the attribute which is to be measured," but then writes, ". . . the items that correlate most highly with total scores are the best items for a general-purpose test" (p. 279). Kline (1983, p. 48, 49) maintains, "The standard item-analytic method of producing a homogeneous test is to correlate each item with the total score on the item pool or, if several scales are being considered simultaneously, to correlate each item with their total score." Though both authors recognize problems that may arise from an analysis of item-total correlations, they do not consider them so great as to diminish the usefulness of the analysis. Nunnally prefers the item-total analysis over alternative analyses such as factor analysis for scale construction and though Kline (p. 49) recognizes the advantage of factor analysis in some situations, he concludes that "there is little essential difference between the methods."

The problem is that the item-total method does not account for external consistency. By not accounting for the relations of the posited alternate indicators with indicators of different factors, an item-total analysis may fail to discriminate between sets of indicators that represent different, though correlated, factors. This problem can be explicated within the context of multiple-indicator measurement models. Though an indicator, x, that is consistent with a multiple-indicator measurement model will correlate more highly with its own factor, ξ, than with another factor, ξ^*, if the factors ξ and ξ^* are highly correlated, then the correlation of x with the other factor, $\rho_x\xi^*$, may be substantial. Let C represent a composite arbitrarily formed by combining items that represent indicators of distinct, though correlated, factors ξ and ξ^*. The correlation of each factor with this composite, $\rho_{\xi C}$ and $\xi_{\xi^* C}$, would also be high. As the following equation demonstrates, indicators from different factors as specified by the true underlying multiple-indicator measurement model, especially indicators that are highly related to their own factor, may correlate highly with the arbitrarily formed composite:

$$\rho_{xC} = \rho_{x\xi}\rho_{\xi C}. \tag{4}$$

As a specific example, consider falsely combining the 10 items whose correlations appear in Table 1 into a single scale. For these data, the correlations of each item

Table 1
AN ILLUSTRATIVE CORRELATION MATRIX

Item	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
1	1.00	.61	.61	.61	.61	.18	.10	.18	.18	.18
2	.61	1.00	.61	.61	.61	.18	.48	.18	.18	.18
3	.61	.61	1.00	.61	.61	.18	.10	.18	.18	.18
4	.61	.61	.61	1.00	.61	.10	.01	.10	.48	.10
5	.61	.61	.61	.61	1.00	.18	.10	.18	.18	.18
6	.18	.18	.18	.10	.18	1.00	.61	.61	.61	.61
7	.10	.48	.10	.01	.10	.61	1.00	.61	.61	.61
8	.18	.18	.18	.10	.18	.61	.61	1.00	.61	.61
9	.18	.18	.18	.48	.18	.61	.61	.61	1.00	.61
10	.18	.18	.18	.10	.18	.61	.61	.61	.61	1.00

with the 10-item total score range from .64 to as high as .71. Even with the correction for the overlap of item and total score (Nunnally 1978), the correlations range from .58 to .67. Thus item-total correlations in the .60's can be obtained from the analysis of a scale defined by two sets of five items with an average intercorrelation of only .19!

Exploratory Factor Analysis

Exploratory factor analysis is a useful scale development technique for reducing a large number of indicators to a more manageable set. It is particularly useful as a preliminary analysis in the absence of sufficiently detailed theory about the relations of the indicators to the underlying constructs. As Churchill (1979, p. 69) wrote, ". . . factor analysis can indeed be used to suggest dimensions, and the marketing literature is replete with articles reporting such use." A typical usage of exploratory factor analysis within this context is to factor an overall set of items and then construct scales on the basis of the resulting factor loadings. Scales are formed by assigning to the same scale the items that load at least moderately on the same factor (e.g., .4) and do not load as highly on other factors.

A primary conceptual difference between exploratory and confirmatory analyses, though, is that the exploratory analysis typically does not provide an explicit test of unidimensionality as defined by the application of equation 1 to each indicator. Factors in an exploratory analysis do not correspond directly to the constructs represented by each set of indicators because each factor from an exploratory analysis is defined as a weighted sum of *all* observed variables in the analysis. For example, even if the factors are orthogonal, the resulting scales generally are correlated. The construction of scales from an analysis of the size of the factor loadings does not provide an evaluation of the unidimensionality of these scales, as would be accomplished by a confirmatory factor analysis in which each factor is antecedent to a mutually exclusive *subset* of the indicators. The more rigorous specification that is required for a confirmatory factor analysis of a multiple-indicator measurement model, in turn, affords a more rigorous evaluation of unidimensionality according to the constraints imposed by internal and external consistency.

Thus, exploratory factor analysis can be a useful *preliminary* technique for scale construction but, if the definition of unidimensionality in equation 1 is used, a subsequent confirmatory analysis would be needed to evaluate, and likely refine, the resulting scales. As an illustration, consider the first three eigenvalues extracted from the correlation matrix in Table 1: 4.33, 2.54, and .77, respectively. Unlike the results obtained from using item-total correlations, the resulting exploratory solution does show a clear two-factor solution. Only two eigenvalues are larger than 1.0 and according to the scree test (Cattell 1978) the differences between successive eigenvalues are 1.78, 1.77, and .15, indicating a steep gradient after the second factor. Accordingly, two principal-axis factors were extracted by iterating for communalities. The oblique (promax) rotation of the resulting two factors also shows a clean two-factor solution:[3] all loadings of the first five items on the first factor range in magnitude from .74 to .80, as do the loadings of the second five items on the second factor, whereas the magnitude of all other loadings is less than .10. Thus the substantive researcher would be likely to form two 5-item scales on the basis of this exploratory analysis.

However, because of the lack of external consistency for items 2 and 7, these 5-item scales do not conform to a multiple-indicator measurement model and, accordingly, a confirmatory factor analysis using the LISREL program (Jöreskog and Sörbom 1984) demonstrates a decided lack of fit for the two-factor multiple-indicator model of all 10 items.[4] The goodness-of-fit index (GFI) value is .79 and the root mean square residual (RMR) value is .067. Moreover, two of the normalized residuals are 9.34 and seven additional normalized residuals are greater than 2.5.

The pattern of normalized residuals from the LISREL analysis clearly indicates the need for a respecification: all the normalized residuals for items 4 and 9 with every other item are greater than 2.5. The next largest normalized residual for any other item is only .10. Unlike the information provided by the exploratory analysis, this pattern strongly suggests that each of these two items is not tapping a single underlying construct as are the other four items in their respective scales.[5] In practice the measurement model would be respecified by eliminating these two items. Performing this respecification results in a model that provides perfect goodness of fit, demonstrating that unidimensional measurement has been achieved.

Without specifically referring to confirmatory factor analysis, Churchill (1979, p. 69) called for more theory testing with factor analysis: "Though this application [exploratory factor analysis] may be satisfactory during the early stages of research on a construct, the use of factor analysis in a confirmatory fashion would seem better at later stages." The subsequent refinement of confirmatory factor analysis programs (Bentler 1985; Jöreskog

[3]An oblique rotation was used because it more accurately reflects the underlying structure of the data than that provided by the more restrictive orthogonal solution. However, for completeness, an orthogonal rotation (varimax) also was performed. The conclusion that the exploratory factor analysis supports the hypothesis of two 5-item scales remained unchanged. The only practical difference in the pattern of loadings between the oblique and orthogonal rotations was that, in the orthogonal solution, the highest loading of an item on the factor corresponding to the other scale was .21 instead of less than .10.

[4]For purposes of estimation, a sample size of 1000 was assumed.

[5]This comparative example underscores the need to include all posited scales within a single analysis so that an assessment of external consistency can be made. Separate confirmatory factor analyses of each scale cannot provide this assessment.

and Sörbom 1984) with the growing acceptance of this methodology provides an almost tailormade means of accomplishing this testing.

Coefficient Alpha and Reliability

Unidimensionality alone is not sufficient to ensure the usefulness of a scale. According to the scale development paradigm advocated here, the reliability of the composite score should be assessed *after* unidimensionality has been acceptably established. Even a perfectly unidimensional (and otherwise construct valid) scale would be of little or no practical use if the resultant composite score were determined primarily by measurement error, with the values of the scores widely fluctuating over repeated measurements.

The three types of reliability coefficients in Cronbach's (1947) classification are the coefficient of stability, the coefficient of equivalence, and the coefficient of stability and equivalence. Both the coefficient of stability and the coefficient of stability and equivalence involve administering forms of the scale on two or more different occasions. Only the coefficient of equivalence can be estimated from the administration of the scale at a single time; consequently it is the reliability coefficient most often used by marketing researchers (Peter 1979).

Coefficient alpha (Cronbach 1951), the most widely used coefficient of equivalence, sometimes has been misinterpreted as an index of unidimensionality rather than reliability (cf. Danes and Mann 1984; Green, Lissitz, and Mulaik 1977; McDonald 1981). Consistent with this misinterpretation, one widely employed practice for developing scales has been to select items that maximize reliability (e.g., Churchill 1979). This practice has been encouraged by computer packages such as SPSS^{x}, which has a specialized routine for accomplishing it. However, unidimensionality and reliability are distinct concepts (Green, Lissitz, and Mulaik 1977). This distinction between dimensionality and reliability can be made meaningfully in the following way. The dimensionality of a scale can be evaluated by examining the patterning of its component indicator correlations, whereas the reliability of a scale is determined by the number of items that define the scale and the reliabilities of those items. In particular, as the following equation demonstrates, coefficient alpha is an application of the Spearman-Brown formula of classical test theory (Lord and Novick 1968), the formula also used in the computation of split-half reliability. The difference is that instead of two splits, the number of splits for coefficient alpha is equal to the number of items on the scale (p):

$$\alpha = \frac{p(\bar{r})}{1 + (p-1)\bar{r}} \tag{5}$$

where $\bar{r}$ represents the average off-diagonal correlation. Hence, *regardless* of the dimensionality of the scale, its reliability tends to increase as the average off-diagonal item correlation increases and/or the number of items increases.

The 10-item correlation matrix in Table 1 provides an example of the distinction between reliability and dimensionality. Each item in the matrix indicates at least one of two distinct factors (ξ_1 and ξ_2), resulting in high within-scale item correlations of .61 relative to the mean between-scale item correlations of only .19, yet coefficient alpha for the 10 items considered as a *single* scale is a respectable .85. After the refinement of the initial 10-item scale with a confirmatory factor analysis, coefficient alpha for each 4-item scale is .86, a slightly higher value despite each scale having less than half as many items.

Moreover, in the computation of coefficient alpha, one *assumes* that (1) the items already form a unidimensional set and (2) the items have equal reliabilities (Nunnally 1978). Computing alpha for items with unequal reliabilities will lead to the underestimation of the reliability of the composite score (cf. Smith 1974). Beginning with Mosier (1943), the unbiased estimate of this reliability for the general case for both weighted and unweighted composites has been given in a variety of forms.[6] Jöreskog (1971) has provided an expression for reliability that does not assume equal item reliabilities within the context of confirmatory factor analysis. For a single unweighted composite for standardized measures, $\bar{x}$, this expression can be given as

$$\rho(\bar{x}\bar{x}') = \frac{\left(\sum_i \lambda_i\right)^2}{\left(\sum_i \lambda_i\right)^2 + \sum_i (1 - \lambda_i^2)} \tag{6}$$

where λ_i represents the i^{th} factor loading on its corresponding factor. Because the items in each 4-item scale all have the same intercorrelation, for this example the values for alpha and Jöreskog's reliability coefficient are identical. It is interesting to note that even for the 10-item scale the two values are very similar, identical until the fourth decimal point. In practice, unless the number of items on the scale is very small and/or the item reliabilities are very discrepant, the underestimation of the composite reliability by alpha is likely to be of no practical consequence.

[6]Mosier's formula accounts for the items' reliabilities as well as their intercorrelations, as can be seen in the following formula.

$$\rho(\bar{x}\bar{x}') = \frac{\sum\sum_{j \neq k} \rho(x_j x_k) + \sum_j \rho(x_j x_j')}{\sum\sum_{j \neq k} \rho(x_j x_k) + p}$$

Heise and Bohrnstedt (1970) developed the same formula from a path analytic context in which they proposed using factor analysis to estimate the item communalities and then, in turn, using these communality estimates as approximations of the item reliabilities. The resulting expression was called *coefficient omega*, which also will underestimate reliability to the extent that item-specific variance is nonzero, though the underestimation will not be as large.

SUBSTANTIVE CONSIDERATIONS

One consequence of the paradigm for scale development outlined here is that the isolation of unidimensional scales with a confirmatory factor analysis of a multiple-indicator measurement model tends to delineate rather specific domains of content. As an example, consider an analysis of the Likert measure of Machiavellianism, the Mach IV scale, which originally had been constructed on the basis of the item-total correlational method. A confirmatory factor analysis revealed that the Mach IV measure was not, itself, unidimensional though it comprises unidimensional subscales (Hunter, Gerbing, and Boster 1982). One of the resulting scales, for example, contained items pertaining only to "duplicity" and another scale contained items pertaining only to "flattery of important others." Though these scales represented narrow content domains, because of their nature they were themselves of substantive interest, as shown by their incorporation into a subsequent structural equation model (cf. Hunter, Gerbing, and Boster 1982).

In some cases, however, the items are unidimensional with respect to a single common construct, but that construct is itself not of general interest because the resulting content domain from which the items were sampled is too restrictive. An alternative is to embed the unidimensional scales, as indicators themselves, within a higher order factor structure (cf. Gerbing and Anderson 1984). The concept of multiple-indicator measurement can be preserved, but the level of analysis shifts so that a more broadly defined construct enters into a structural model or other application. An advantage of this approach is that the first-order factors, estimated with multiple measures, have their estimated construct correlations implicitly corrected for attenuation due to measurement error (cf. Howell 1987). The primary analysis then shifts from an analysis of individual items to an analysis of more reliable versions of those items, the first-order constructs. An example of developing such second-order constructs is found in the analysis of impulsivity by Gerbing, Ahadi, and Patton (1987).[7]

SUMMARY AND CONCLUSION

Contributing to the tradition of articles by Churchill (1979) and Peter (1979, 1981), we outline an updated paradigm for scale development that incorporates a more recent methodological development: confirmatory factor analysis. In doing so, we attempt to provide a better understanding of the concept of unidimensional measurement and the ways in which it can be assessed and, in particular, to demonstrate that an explicit evaluation of unidimensionality is accomplished with a confirmatory factor analysis of the individual measures as specified by a multiple-indicator measurement model. Coefficient alpha is important in the assessment of reliability, but it does not assess dimensionality. Though item-total correlations and exploratory factor analysis can provide useful preliminary analyses, particularly in the absence of sufficiently detailed theory, they do not directly assess unidimensionality. The reason is that a confirmatory factor analysis makes possible an assessment of the internal consistency *and* external consistency criteria of unidimensionality implied by the multiple-indicator measurement model.

Following the paradigm of scale development outlined here, after the unidimensionality of a set of scales has been acceptably established, one would assess its reliability. Even a perfectly unidimensional scale will not be useful in practice if the resultant scale score has unacceptably low reliability. Because most measures in marketing are administered at a single point in time, coefficient alpha or some other coefficient of equivalence reliability would probably be used for this assessment.

The goal of most research projects is not just to develop unidimensional and reliable measurement scales, but to build and test theory. Essential to this undertaking is the assessment of construct validity. A construct achieves its meaning in two ways (Anderson 1987; Cronbach and Meehl 1955): (1) through observed indicators for which it is posited to be causally antecedent (and through observed measures for which it is *not*) and (2) through the set of relationships of the construct with other constructs as specified by some theory (the nomological network). Unidimensionality, then, is necessary but not sufficient for construct validity. Not only should all the indicators that define a scale provide estimates of exactly one factor, but the meaning of the underlying factor should correspond to the construct of interest.

The nomological network can be explored within the context of the full structural equation model. One means for accomplishing this is the approach developed by Anderson and Gerbing (1988) that allows an assessment of nomological validity that is asymptotically independent of the assessment of the measurement model. It is called a "two-step" approach because the measurement model first is developed and evaluated separately from the full structural equation model that simultaneously models measurement and structural relations. The measurement model in conjunction with the structural model makes possible a comprehensive confirmatory assessment of construct validity (Bentler 1978). Hence, the assessment of unidimensionality provided by a confirmatory factor analysis represents but a first step in the establishment of meaning for the estimated factors.

[7]This analysis of impulsivity measures is a substantive example of the scale development paradigm outlined here. Item-total correlations and exploratory factor analysis were used to reduce a large number of items to a more manageable set and to provide preliminary scales, which then were refined with a confirmatory factor analysis.

REFERENCES

Anderson, James C. (1987), "An Approach for Confirmatory Measurement and Structural Equation Modeling of Organizational Properties," *Management Science*, 33 (April), 525–41.

——— and David W. Gerbing (1982), "Some Methods for Respecifying Measurement Models to Obtain Unidimensional Construct Measurement," *Journal of Marketing Research*, 19 (November), 453–60.

——— and ——— (1988), "Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach," *Psychological Bulletin*, forthcoming.

Bentler, Peter M. (1978), "The Interdependence of Theory, Methodology, and Empirical Data: Causal Modeling as an Approach to Construct Validation," in *Longitudinal Drug Research*, D. B. Kandel, ed. New York: John Wiley & Sons, Inc., 267–302.

——— (1985), "Theory and Implementation of EQS, A Structural Equation Program." Los Angeles: BMDP Statistical Software.

Burt, Ronald S. (1973), "Confirmatory Factor-Analytic Structures and the Theory Construction Process," *Sociological Methods & Research*, 2 (November), 131–90.

Cattell, Raymond B. (1978), *The Scientific Use of Factor Analysis in Behavioral and Life Sciences*. New York: Plenum Press.

Churchill, Gilbert A., Jr. (1979), "A Paradigm for Developing Better Measures of Marketing Constructs," *Journal of Marketing Research*, 16 (February), 64–73.

Cronbach, Lee J. (1947), "Test 'Reliability': Its Meaning and Determination," *Psychometrika*, 12 (March), 1–16.

——— (1951), "Coefficient Alpha and the Internal Structure of Tests," *Psychometrika*, 16 (September), 297–334.

——— and Paul E. Meehl (1955), "Construct Validity in Psychological Tests," *Psychological Bulletin*, 52 (July), 281–302.

Danes, Jeffery E. and O. Karl Mann (1984), "Unidimensional Measurement and Structural Equation Models with Latent Variables," *Journal of Business Research*, 12 (September), 337–52.

Gerbing, David W. and James C. Anderson (1984), "On the Meaning of Within-Factor Correlated Measurement Errors," *Journal of Consumer Research*, 11 (June), 572–80.

———, Stephan A. Ahadi, and Jim H. Patton (1987), "Toward a Conceptualization of Impulsivity: Components Across the Behavioral and Self-Report Domains," *Multivariate Behavioral Research*, 22 (July), 357–79.

Green, Samuel B., Robert W. Lissitz, and Stanley A. Mulaik (1977), "Limitations of Coefficient Alpha as an Index of Test Unidimensionality," *Educational and Psychological Measurement*, 37 (October), 827–38.

Harman, Harry H. (1976), *Modern Factor Analysis*, 3rd ed. Chicago: The University of Chicago Press.

Hattie, John (1985), "Methodology Review: Assessing Unidimensionality of Tests and Items," *Applied Psychological Measurement*, 9 (June), 139–64.

Heise, David R. and George W. Bohrnstedt (1970), "Validity, Invalidity, and Reliability," in *Sociological Methodology*, Edgar F. Borgatta and George W. Bohrnstedt, eds. San Francisco: Jossey-Bass, Inc., 104–29.

Howell, Roy D. (1987), "Covariance Structure Modeling and Measurement Issues: A Note on 'Interrelations Among a Channel Entity's Power Sources'," *Journal of Marketing Research*, 24 (February), 119–26.

Hunter, John E. and David W. Gerbing (1982), "Unidimensional Measurement, Second Order Factor Analysis, and Causal Models," in *Research in Organizational Behavior*, Vol. 4, B. M. Staw and L. L. Cummings, eds. Greenwich, CT: JAI Press, Inc., 267–99.

———, ———, and Frank J. Boster (1982), "Machiavellian Beliefs and Personality: The Construct Invalidity of the Machiavellian Dimension," *Journal of Personality and Social Psychology*, 43 (December), 1293–1305.

Jöreskog, Karl G. (1971), "Statistical Analysis of Sets of Congeneric Tests," *Psychometrika*, 36 (June), 109–33.

——— and Dag Sörbom (1984), *LISREL VI: Analysis of Linear Structural Relationships by the Method of Maximum Likelihood*. Chicago: National Educational Resources, Inc.

Kline, Paul (1983), *Personality: Measurement and Theory*. New York: St. Martin's Press.

Lord, Fred M. and Melvin R. Novick (1968), *Statistical Theories of Mental Test Scores*. Reading, MA: Addison-Wesley Publishing Company, Inc.

McDonald, Roderick P. (1981), "The Dimensionality of Tests and Items," *British Journal of Mathematical and Statistical Psychology*, 34 (May), 100–17.

Mosier, Charles I. (1943), "On the Reliability of a Weighted Composite," *Psychometrika*, 8 (September), 161–8.

Nunnally, Jum C. (1978), *Psychometric Theory*, 2nd ed. New York: McGraw-Hill Book Company.

Peter, J. Paul (1979), "Reliability: A Review of Psychometric Basics and Recent Marketing Practices," *Journal of Marketing Research*, 16 (February), 6–17.

——— (1981), "Construct Validity: A Review of Basic Issues and Marketing Practices," *Journal of Marketing Research*, 18 (May), 133–45.

Smith, Kent W. (1974), "On Estimating the Reliability of Composite Indexes Through Factor Analysis," *Sociological Methods and Research*, 2 (May), 485–510.