148

MEASURING TRAIT COMMUNICATION APPREHENSION: A TEST OF RIVAL MEASUREMENT MODELS OF THE PRCA-24

TIMOTHY R. LEVINE AND JAMES C. MCCROSKEY

The purpose of the present investigation was to test three rival measurement models of a common measure of trait communication apprehension, the PRCA-24. The rival models include a linear, unidimensional model, a Guttman simplex, and a second-order factor structure. These models are discussed in terms of their conceptual implications, their psychometric properties, the empirical evidence needed for their support, and their implications for both prior and subsequent research. To test these models, a nationwide sample of responses to the PRCA-24 (N = 8879) was analyzed. The results of a variety of statistical procedures indicate that the second-order factor model best fits the data. This model is consistent with prior conceptual work by McCroskey (1984a) and was replicated with two additional data sets.

Communication apprehension (CA) is one of the most frequently investigated variables in the field of human communication. Payne and Richmond (1983) found 876 published articles and convention papers on topics related to the CA construct; a current review of communication writings would demonstrate that this interest has not diminished.

One of the most popular methods for measuring CA has been the use of self-report measures (McCroskey, 1984b; Richmond & McCroskey, 1985) and in particular, the PRCA-24 (McCroskey, 1982). There is substantial evidence to support both the reliability and the construct validity of the PRCA-24 (e.g., McCroskey, Beatty, Kearney & Plax, 1985). The PRCA-24 exhibits high inter-item correlations and the total score correlates with other trait and outcome variables in a manner consistent with its validity. The PRCA-24 was designed to measure trait CA in four communication contexts: dyadic, group, meeting, and public speaking. The scale consists of four subscales, each measuring respondents' CA in a particular context, that can be summed to measure the more global CA construct. Through this design, McCroskey (1984a) implicitly hypothesized a second-order factor structure. This model, however, has never been explicitly delineated in such terms nor has it been subjected to a direct empirical test.

The typical CA researcher tests the PRCA-24's factor structure with principal axes or principal components analyses. The results of these exploratory factor analyses, however, often lead to mixed results. Three patterns are common. These include a four-factor solution with each factor corresponding to each subscale, a unidimensional solution, and a two-dimensional solution with the dyadic and the group items loading on one factor and the meeting and public speaking items loading on the other factor. Such mixed results have led some to believe that the scale has an unstable factor structure, which would challenge the validity of the PRCA-24 and would lead us to question the results of prior research in which the scale was used. If

Timothy R. Levine is a doctoral candidate in the department of communication at Michigan State University. James C. McCroskey is professor and chair of the department of communication studies at West Virginia University. *The data collection for this research was supported by a Special Projects grant from the American Association of Colleges of Pharmacy Council of Sections through the American Foundation for Pharmaceutical Education.*

we are to assume the PRCA-24 is valid, the question which must be addressed is, What causes the apparent instability in the factor structure? Since a second-order factor structure would predict substantial correlations among the first-order factors (i.e., between context scores), this measurement model may explain the conflicting factor solutions obtained from exploratory factor techniques.

A rival explanation might be that the scale is in fact unidimensional and that two- and four-factor solutions are spurious. Multiple factor solutions may be a function of item means. This explanation is consistent with assumptions made by a Guttman simplex.

These models are subject to empirical test. Since the results of social scientific research are only as valid as the measurement used to obtain the results, understanding the psychometric properties of our measures is crucial. This is particularly important for frequently used measures. For example, Christie and Geis's (1970) Mach IV scale was used in numerous studies and thought to be valid and second-order unidimensional until Hunter, Gerbing, and Boster (1982) demonstrated that the scale was, in fact, radically multidimensional. Thus, understanding the psychometric properties of our measures is of great practical import since such knowledge is essential for distinguishing between valid and invalid uses of a given measure.

These practical considerations are by no means the only reason for investigating the psychometric properties of our measures. Such issues have theoretical implications as well. Measurement is the vehicle for translating our conceptualization of a construct into observable units. Good measures should not only be capable of yielding valid results but also provide an accurate rendition of our conceptualization of a construct. Only when there exists a high degree of correspondence between our conceptualization of a construct and the psychometric properties of our measures can valid results be interpreted as theoretically meaningful.

The purpose of the present investigation is to test three rival measurement models of the PRCA-24. These models include (1) a linear, unidimensional model, (2) a Guttman simplex, and (3) a second-order factor structure. In the next sections, each of the models, along with their conceptual and mathematical assumptions and their empirical indicators, will be discussed.

The Linear Unidimensional Model

Virtually all previous research has treated the PRCA-24 as a unidimensional scale. Researchers simply sum the items, note the reliability (which is usually high), and assume that since the scale is reliable, it must be unidimensional.

This model assumes that each respondent has a certain amount of trait CA which is not measured directly but treated as a latent variable. It is implicitly assumed that the trait causes individuals to respond to various scale items in a particular manner and these responses are summed as a measure of the trait. Thus, the trait of interest drives the responses to each item, with a certain amount of error. The internal consistency of this model may be tested applying a simple product rule. The predicted correlation between any two items measuring the same trait is the product of their factor loadings (Hunter & Gerbing, 1982). The predicted inter-item correlations are then compared to the obtained correlations as a test of the model's fit. If the obtained correlations are within sampling error of the predicted correlation, then the data may be said to fit the model. This is the rationale behind confirmatory factor analysis (Hunter & Gerbing, 1982).

In addition to testing internal consistency, Hunter and Gerbing (1982) also argue for testing the parallelism of the measure in question with other valid measures. The parallelism test determines if each item tapping a particular trait correlates with a different construct in a manner similar to other items tapping the trait. Again applying the product rule, the correlation between any two items tapping separate traits should be the product of the factor loadings of each item and its corresponding trait and the correlation between the two traits. As with internal consistency, testing parallelism involves comparing predicted correlations to obtained correlations. If the deviations are small enough to be explained by chance variation, then the model is said to fit the data.

If the data meet both criteria of internal consistency and parallelism, one may have some confidence that a measure fits a linear, unidimensional model. Confirmatory factor analysis is superior to exploratory factor analysis techniques insofar as it provides a more precise test of dimensionality. Since a model is specified a priori, and the obtained data are compared to those predicted by the model, the possibility of detecting specification error is increased.

Guttman Simplex

As mentioned earlier, results of exploratory factor analytic techniques testing the dimensionality of the PRCA-24 often yield divergent solutions. One potential explanation for these results is that responses to the PRCA-24 form a Guttman simplex. That is, responses to individual items may be related nonlinearly to CA. If the PRCA forms a Guttman simplex, differing subscale means may result in spurious multi-factor solutions (Guttman, 1955; Hunter & Boster, 1987). The presence of nonlinearity produces spurious factors since items with similar means show higher inter-item correlations than items with different means.

Not only does the Guttman model explain the results of exploratory factor analytic techniques, it also may make conceptual sense. The four subscales comprising the PRCA-24 may represent increasing thresholds for becoming apprehensive. For example, those who are apprehensive in public speaking contexts may not be apprehensive in the other contexts. Similarly, individuals who are anxious in meetings may also be anxious in public speaking situations but not necessarily in group or dyadic settings. Following this reasoning, individuals who experience CA in dyadic situations will also experience it in the other settings as well. Thus, the four subscales may tap different levels of CA, not different types of CA, producing a step-like function. Theoretically, this model differs radically from the current conceptualization of CA.

Guttman (1955) and Hunter and Boster (1987) discuss empirical tests for a Guttman simplex. First, if the multidimensional solutions yielded by exploratory factor analyses are spurious, we should expect the means of the subscales to be ordered. Moreover, the discrepancies between means should be negatively related to the between-subscale correlations. A strong negative correlation would be consistent with the hypothesis that multiple-factor solutions are a function of subscale means. Second, the regression of the subscale scores onto the total CA score should be non-linear and ogival. Third, when the factor loadings from the first factor are plotted on the x-axis and loadings from the second factor are plotted on the y-axis, the pattern on the points should form the arc of a semicircle with the rank on the points corresponding to the rank of the item means. A final test of the Guttman simplex involves dichotomizing scores on any two subscales, creating a two by two frequency table. If the Guttman simplex model fits the data, we would expect people to fall into only three of the four cells in the table. For example, when dichotomizing responses to the public speaking and dyadic subscales into high and low, we would expect few if any individuals to fall into the high dyad, low public speaking cell. This rationale can be applied to all two-context combinations of the subscales. In all combinations, few individuals should fall into the critical high/low cell.

Second-Order Factor Structure

McCroskey (1984a) intended each of the four subscales as somewhat independent measures of communication apprehension in a particular context. These subscales, however, were also designed to be combined into a global measure of trait CA. Put differently, each scale measures a distinct latent variable, CA in a given context, but each subscale is a measure of a broader latent variable, trait CA. This conceptualization is consistent with what Hunter and Gerbing (1982) term a second-order factor structure.

As with the linear unidimensional model, empirical support for the second-order factor model can be obtained from the use of confirmatory factor analyses. First, each of the subscales must be consistent with the linear, unidimensional model. That is, they must be tested separately for internal consistency and parallelism. If each of the four subscales (first-order factors) fits this model, the subscales are then treated as items tapping CA (the second-order factor) and tested again for internal consistency and parallelism (Hunter & Gerbing, 1982). For the second-order factor model to fit the data, both the first- and second-order factors must fit. If the first-order factors pass internal consistency and parallelism tests but the factors do not combine to form the second-order factor, then the scale is multidimensional. If the PRCA-24 is found to be consistent with a linear, multidimensional model, then summing all items would be inappropriate.

Implications of the Rival Models

Each of the three models reviewed has important conceptual and empirical implications for the measurement of CA. While all three are consistent with the treatment of the PRCA-24 as a measure of global trait CA, each is based on different assumptions about the nature of CA and, to a lesser extent, each differs with respect to the legitimacy of using context scores. If the PRCA-24 is found to fit the linear, unidimensional model or the Guttman simplex model, subsequent researchers would want to avoid using subscale scores since both models suggest the entire scale is unidimensional. These models differ, however, on a conceptual level. The linear unidimensional model makes no distinctions between subscale items. All 24 PRCA items are assumed to measure the same construct. Thus, using a subscale would merely lower the reliability of the scale (due to fewer items) and result in a less precise measure of trait CA. The Guttman simplex model, on the other hand, views the subscale scores as tapping different levels of trait CA. From this perspective, using only some of the subscales would result in a restriction of range and/or lack of precision. Thus, according to the Guttman model, all four subscales would be necessary to gain an accurate reflection of a person's trait CA.

The second-order model would also allow all items to be summed as a measure of trait CA. Unlike the others, it conceptually allows for the use of subscores since each subscale is viewed as a distinct sub-construct. Thus, if this model receives support, subsequent research using the PRCA-24 could legitimately use subscores, total scores, or both. However, if the scale is found to fit a multidimensional model, prior research in which total scores were computed would be drawn into question.

The three models reviewed above are mutually exclusive. While any one empirical indicator, when considered alone, does not constitute unambiguous support or refutation of any model, when taken together the criteria reviewed above can distinguish each of these models. Therefore, solid support for one model constitutes disconfirmation of the others.

These three models are by no means a comprehensive list of possible measurement models. If none of these models receives support, the validity of the PRCA-24 as a measure of trait apprehension would be in question. Disconfirmation of all three models would be inconsistent with a large body of prior findings.

METHOD

Research Participants

To provide an appropriate test of the models under study, it was important that we obtain a sample of subjects with a high likelihood of generalizability to the general population. Several large-sample data sets were available. The one judged best for our purposes was collected from a national college student population (Berger, Baldwin, McCroskey, & Richmond, 1983). The sample was drawn from 51 of the 71 pharmacy schools in the United States and constituted over 40 percent of all the students in pharmacy at the time of the study. About half of the subjects in the sample were female and half male. The mean score on the PRCA-24 for the sample was 65.2, less than half a point from the combined normative mean drawn from over 40,000 college student subjects (65.6) and from the mean (65.4) of a national non-college student (chiropractic assistants) sample (Allen, Richmond, & McCroskey, 1984). Data from a total of 8879 subjects were included in our analyses.

Measurement

Each participant completed a short questionnaire containing the PRCA-24, a 5-item semantic differential-type immediacy scale, and a number of demographic items. The PRCA-24 is a 24-item Likert-type scale with a 5-point response format anchored by "strongly agree" and "strongly disagree."

The immediacy scale also used a 5-point response format and served the test of parallelism. The scale was found to be highly reliable (alpha = .97) and to fit a unidimensional structure based on confirmatory factor analysis using PACKAGE (Hunter, Cohen, & Nicol, 1982).

Statistical Analyses

The three rival models, as well as their empirical indicators (i.e., the criteria necessary for their support), were specified a priori. The Guttman simplex model was tested first since the nonlinear regressions required for its confirmation would make linear OLS estimation procedures (e.g., confirmatory factor analyses) nonapplicable. If this model was not supported, the remaining linear models would be tested with confirmatory factor analyses.

All confirmatory factor analyses were conducted with PACKAGE (Hunter et al., 1982). This program provides estimates of factor loadings based on a centroid solution, computes predicted correlations based on the model specified, and provides deviations between the predicted and the actual correlation with which to assess model fit. Unlike most confirmatory factor analyses, however, the fit of the model was determined by the absolute magnitude of the deviations rather than being based on sampling error. Due to the large sample size, deviations of less than .02 differ significantly from the predicted correlations. Therefore, sampling error as a criterion could potentially lead to disconfirmation based on a fraction of a percent of variance. This makes confirmation virtually impossible. If the only sources of error were specification error and sampling error, this criterion might be justified. It would be naive for us, however, to reject the possibility of minor response errors other than specification or sampling errors. For this reason, a somewhat arbitrary deviation magnitude of .10 was set as the critical value. It was reasoned that deviations amounting to less than 1% of the variance would be considered trivial by most communication researchers. This critical value, we reasoned, would avoid possible disconfirmation from trivial deviations while maintaining a somewhat conservative criterion for goodness of fit. Again, this value was set a priori.[1]

RESULTS

Test of the Guttman Simplex

Examination of the four subscale means indicated that these means were ordered in a manner consistent with the Guttman model. The dyadic mean was lowest ($M = 14.42$, $s = 4.39$), followed by the group mean ($M = 15.52$, $s = 4.94$), meeting mean ($M = 16.41$, $s = 4.96$), and the public speaking mean ($M = 18.65$, $s = 5.18$). Moreover, a strong negative correlation ($r = -.82$, df $= 4$, $p > .05$) between the discrepancies in subscale means and the intersubscale correlations was obtained, consistent with the hypothesis that multidimensional factor solutions are spurious and satisfying the first criterion of a Guttman model.

However, the second, third, and fourth criteria were not met. Regressions of each subscale onto the total PRCA-24 score produced four regression lines closely approximating linearity rather than ogival curves. These patterns were inconsistent with the Guttman model. The plotting of factor loadings was also inconsistent with a Guttman simplex. The data were subjected to principal axis factor analysis with a varimax rotation with the factor loadings from the first extracted factor plotted on the x-axis and the loadings from the second factor plotted on the y-axis. With the exception of the four items with the largest means and the three items with the smallest means, the rank of the plots showed little correspondence with the rank of the means. Finally, examination of the 2 by 2 frequency tables for all combinations of subscales revealed that while the critical high/low cell was always less than the low/high cell, the frequency of respondents within the cell was, in all cases, substantially greater than zero (8.6% to 13.6%). These results together with the linear regressions suggest that responses to the PRCA-24 do not meet the assumptions of a Guttman simplex.

Test of the Linear, Unidimensional Model

The linear, unidimensional model involved testing the internal consistency and parallelism of the scale with confirmatory factor analysis. This analysis, consistent with previous results, found that the scale was highly reliable (alpha = .94). The test of internal consistency identified that 34 of the 276 (12.3%) deviations were greater or equal to the critical value of .10. Twelve of these deviations were greater than .20 and three were greater than .30. The test of parallelism identified 5 of the 120 (4.2%) deviations that were greater or equal to the critical value. All of these deviations, however, were attributable to one item, number 17, and all were less than .13.

Thus, the results of these analyses indicated that the PRCA was highly reliable and that all but one item correlated with the items of a separate measure in a manner consistent with that of a linear, unidimensional scale. The internal structure of the scale, however, deviated substantially from that which was predicted by the unidimensional model. Moreover, since the offending deviations were distributed across a majority of items, these discrepancies could not be reconciled with the exclusion of a few items. For these reasons, the linear, unidimensional model was also rejected.

Test of the Second-Order Factor Model

The test of the second-order factor model involved two successive confirmatory factor analyses. The first was a test of the first-order factor structure—that the PRCA-24 formed four reliable, internally consistent, and parallel factors, each containing items tapping CA in a different context. The second tested the second-order factor—that the subscales combined to form one global measure of CA.

Each of the four first-order factors were found to be fairly reliable: .87, .89, .86, and .86 for the group, meeting, dyadic, and public speaking contexts respectively. The test of internal consistency found 9 of 60 (15%) deviations to be greater or equal to .10. Most of the offending deviations, however, were only slightly greater than the critical value. Only one deviation exceeded .13 and all were attributable to relatively few items.

The test of parallelism with the outside measure revealed all 144 deviations were less than the critical value. The test of parallelism between the four PRCA-24 factors identified 5 of 216 (2.3%) deviations greater or equal to the critical value. Again, these deviations were only slightly over the critical value.

The second-order factor was found to be reasonably reliable (alpha = .81), internally consistent, and parallel to the outside factor. None of the deviations met or exceeded the critical value.

Thus, the second-order factor model met all criteria except internal consistency. Since the offending deviations were few in number, relatively small in magnitude, and attributable to only a few items, a secondary analysis was conducted with four items (1, 10, 17, and 24)[2] excluded, one from each first-order factor.

The exclusion of these items did not substantially lower the reliabilities of the first-order factors (alpha = .86, .88, .83, and .85). The test of internal consistency revealed only one deviation (.11) greater or equal to the critical value. The vast majority of the remaining deviations were substantially lower than the critical value (80% $\leq$ .05) and the mean absolute deviation was .04. The parallelism test with the outside factor revealed that all 100 deviations were less than the critical value. The mean absolute deviation was .02. The parallelism test between the first-order factors identified 5 of 150 (3.3%) of the deviations were greater or equal to the critical value. Of these, all were less than .15. The mean absolute deviation was .04.[3]

The reliability of the second-order factor also remained acceptable (alpha = .90)

TABLE 1
INTERNAL CONSISTENCY AND PARALLELISM DEVIATION MATRICES FOR THE SECOND-ORDER FACTOR

	PRCA Subscales			Outside Scale Item				
Subscale	Group	Meeting	Dyad	1	2	3	4	5
Group				.00	−.02	−.02	−.02	−.02
Meeting	−.01			−.04	−.06	−.06	−.06	−.07
Dyad	.02	−.01		.11	.09	.09	.09	.09
Public Speaking	−.01	.02	−.02	−.01	−.02	−.02	−.02	−.02

and the factor met both the internal consistency (mean absolute deviation = .02) and the parallelism (mean absolute deviation = .05) criteria, with one deviation larger than .10. The deviations from the test of parallelism between the dyad subscale and the outside factor, however, were uniformly larger than desirable. These deviations are presented in Table 1.

While, strictly speaking, not all the criteria for the second-order factor model were met, the data closely approximated this model. In all, only 7 of the 336 (2.1%) observed correlations examined deviated from the correlations predicted by the model by .10 or more. Of these, all deviations were .14 or less. The average deviation across tests was .03, substantially below .10. Therefore, given the close approximation of the data to this model, the consistency of this model with prior conceptual and empirical work, and the inconsistency of the data with rival measurement models, there is little reason not to conclude that the PRCA-24 forms a second-order factor model.

Replications

While the results presented above were derived from a large and geographically diverse sample, they still represent only one sample. Since generalizability is of great importance in the advancement of a measurement model, we attempted to replicate the results with three additional data sets.

The first sample (N = 1483) included incoming freshmen at a medium sized eastern university who completed the PRCA-24 during freshman orientation in 1982. The data were collected as part of a longitudinal study of the impact of CA on student retention (McCroskey, Booth-Butterfield, & Payne, 1989). This sample was 56% male and 44% female. The gender breakdown and dropout rate closely approximated that of all freshman students reported by the Office of Admissions and Records at the host institution. One limitation of this data set was that no multiple-item outside measures for testing parallelism were administered.

The second data set consisted of students (N = 142) enrolled in communication courses at a large midwestern university. In addition to the PRCA-24, participants completed measures of lie acceptability and trait suspicion (Levine & McCornack, 1989). Thus, this data set allowed for testing parallelism with two outside factors.

The third data set included students (N = 872) at the University of Puerto Rico as part of a study of intercultural differences related to CA (McCroskey, Fayer, & Richmond, 1985). Only students who reported Spanish to be their first language were included in the study. The students completed the PRCA-24 twice, once for their feelings when speaking in Spanish and the other for their feelings when speaking in English. Since virtually every Puerto Rican student studies English (as a second language) every year from the first grade through college, most are capable of communicating at or above a moderate level of proficiency in English. Again, no outside measures were administered.

The results from the first replication were consistent with the second-order factor model. As with the initial results, the same four items proved problematic and were deleted. Only one deviation (.11) exceeded .10 in the test of internal consistency of the first-order factors and mean absolute deviation was .03. None of the deviations from the tests of parallelism with the first-order factors (mean absolute deviation = .03) or the test of internal consistency for the second-order factor exceeded this value.

The results of the second replication were also consistent with the second-order factor model (with the same four items deleted). Since sampling error became an issue with the substantially smaller sample size, significance tests were employed to determine the fit of the model. The number of statistically significant deviations was below that expected by chance (3.6%) and none of the Chi Square values were significant.

The results from the Puerto Rican sample, however, deviated substantially from the second-order model both for Spanish and English. Even with the four questionable items deleted, 5% of the deviations for English and 10.5% of the deviations for Spanish exceeded .10. Moreover, these deviations were not confined to specific items or specific tests.

DISCUSSION

The purpose of the present investigation was to identify and provide support for the measurement model underlying the PRCA-24, a frequently employed measure of trait communication apprehension. To this end, three rival measurement models were identified and discussed in terms of their assumptions, implications, and the evidence necessary for their support. These models included a Guttman simplex, a linear, unidimensional model, and a second-order factor model. The data best fit the second-order factor model. This finding was replicated with two additional data sets. The results of a third replication employing cross-cultural data, however, failed to replicate the second-order model.

The second-order factor model views the context-specific subscales as related but discrete subconstructs. These constructs combine to form a global construct. Thus, the subscales corresponding to each construct may be used independently as measures of CA in a given context or summed as a measure of global trait CA. The use of subscale scores, however, should be considered only when mandated by substantive considerations (such as selecting treatment modalities) since they exhibit lower reliabilities due to fewer items. This model is consistent both with prior conceptual work by McCroskey (1984a) and with a substantial body of literature supporting the validity of the PRCA-24.

In addition to the empirical support generated by this investigation, other reasons bolster our confidence in this conclusion. As previously noted, the second-order factor model is consistent with the current conceptualization of CA delineated by McCroskey (1984a). This model is also consistent with data supporting the construct validity of the PRCA-24. Finally, we believe this model can account for the apparent instability evident in the results of exploratory factor analysis techniques.

Our conclusions do not have any direct negative implications for the majority of previous investigations involving the entire PRCA-24 or any of its subscales. The conclusions may, however, prove troublesome for subsequent research insofar as the factor structure of the PRCA-24, for a given sample, should no longer be tested with traditional (exploratory) factor analysis. Rather, subsequent researchers desiring confirmation of their use of the measure should seek to confirm that the second-order factor structure fits the data from their samples.

It is also advisable to reduce the PRCA-24 to 20 items and to refrain from using items 1, 10, 17, and 24. These latter items proved problematic in each of the three data sets found to be consistent with the second-order factor model. Our results suggest that including these four items introduces error into the measurement model.

Another implication for subsequent research concerns the use of the scale with non-native English speakers or participants from other cultures. Since the measurement model did not hold for the Puerto Rican sample, our results lead us to question the validity of the scale's use in intercultural research. While the results based on one sample from one culture should not be used to damn the results of previous research using the PRCA-24 in intercultural contexts, these results do warrant caution in this area. It is quite possible that the CA construct and measure cannot be translated into the language and culture of some other groups around the world. Furthermore, this implication may well extend to the cross-cultural use of other measurement instruments validated within a single culture. In such cases, additional cross-cultural measurement work is surely warranted.

Finally, one of the primary contributions of this research is to serve as a model for the explication and testing of measurement models. The models and statistical procedures detailed here are applicable to a wide variety of measurement instruments. While the specification and confirmation of a measurement model is not sufficient to constitute the validation of a measure, it is certainly a necessary part of the validation process. We strongly believe that careful measurement work is a necessary prerequisite for valid and useful results.

END NOTES

[1]The standard errors for the deviations ranged from .008 to .01. Thus, the confidence interval around each deviation is less than ± .02. If we adopted statistical significance as a criterion, 75% of the deviations would differ significantly from zero, and the model must be rejected. This criterion is, in our opinion, misleading because it makes type II error (a false negative) probable.

If we perform correlational analyses with a sample of 9000, it is almost impossible not to obtain statistically significant effects (e.g., a correlation of .04 is significant). If we ran across such a result, we would dismiss such a small correlation as trivial even though zero is outside its confidence interval. Just as a correlation of .04 is trivial, regardless of sample size, so too is a deviation of .04 trivial.

We set our criterion at .10 to guard against type II error; rejecting a model on the basis of trivial but statistically significant deviations. This value maintains a rather conservative critical value. Being able to predict 336 correlations, each to within one percent of the variance, represents a level of precision uncommon in social scientific research. The fact that our average deviation was .03 serves as a testimonial to this level of precision.

Following the same reasoning, we also opted not to use Chi Square tests as a test of overall fit. Chi Square is calculated by multiplying the sample size by the sum of squared deviations. Thus, as sample size increases, so does the Chi Square value. Since the degrees of freedom are determined by the over-identification status, not the sample size, the critical value for Chi Square does not reflect the sample size. For this reason, Chi Square is also misleading for large sample data.

[2] The wording of the four problematic items is: (1) I dislike participating in group discussions. (10) I am afraid to express myself at meetings. (17) While conversing with a new acquaintance, I feel very relaxed. (24) While giving a speech I get so nervous, I forget facts I really know.

[3] All tables of deviations and the results of tests of rival models are available, upon request, from the first author.

REFERENCES

Allen, J., Richmond, V.P., & McCroskey, J.C. (1984). Communication and the chiropractic profession: Part 1. *Journal of Chiropractic, 21,* 24–30.

Berger, B.A., Baldwin, J.J., McCroskey, J.C., & Richmond, V.P. (1983). Communication apprehension in pharmacy students: A national study. *American Journal of Pharmaceutical Education, 47,* 95–102.

Christie, R. & Geis, F. (1970). *Studies in machiavellianism.* New York: Academic Press.

Guttman, L. (1955). A generalized simplex for factor analysis. *Psychometrika, 20,* 173–192.

Hunter, J.E. & Boster, F.J. (1987). A model of compliance-gaining message selection. *Communication Monographs, 54,* 63–84.

Hunter, J.E., Cohen, S.H., & Nicol, T.S. (1982). *PACKAGE.* East Lansing, MI: Michigan State University.

Hunter, J.E. & Gerbing, D.W. (1982). Unidimensional measurement, second-order factor analysis, and causal models. *Research in Organizational Behavior, 4,* 267–320.

Hunter, J.E., Gerbing, D.W., & Boster, F.J. (1982). Machiavellian beliefs and personality: Construct invalidity of the Machiavellianism dimension. *Journal of Personality and Social Psychology, 43,* 1293–1305.

Levine, T.R. & McCornack, S.A. (1989, November). *Distinguishing between types of suspicion and measuring a predisposition towards suspicion: Two studies validating a measure of trait suspicion.* Paper presented at the annual meeting of the Speech Communication Association, San Francisco, CA.

McCroskey, J.C. (1984a). The communication apprehension perspective. In J.A. Daly and J.C. McCroskey (Eds.) *Avoiding communication: Shyness, reticence, and communication apprehension* (pp. 13–38). Beverly Hills, CA: Sage.

McCroskey, J.C. (1984b). Self report measurement. In J.A. Daly a..d J.C. McCroskey (Eds.) *Avoiding communication: Shyness, reticence, and communication apprehension* (pp. 81–94). Beverly Hills, CA: Sage.

McCroskey, J.C., Beatty, M.J., Kearney, P., & Plax, T.G. (1985). The content validity of the PRCA-24 as a measure of communication apprehension across communication contexts. *Communication Quarterly, 33,* 165–173.

McCroskey, J.C., Booth-Butterfield, S., & Payne, S.K. (1989). The impact of communication apprehension on college student retention and success. *Communication Quarterly, 37,* 100–107.

McCroskey, J.C., Fayer, J.M., & Richmond, V.P. (1985). Don't speak to me in English: Communication apprehension in Puerto Rico. *Communication Quarterly, 33,* 185–192.

Richmond, V.P. & McCroskey, J.C. (1985). *Communication: Apprehension, avoidance, and effectiveness.* Scottsdale, AZ: Gorsuch Scarisbrick, Publishers.

Seibold, D.R. & McPhee, R.D. (1980). A new analysis of Daly's "Assessment of social-communicative anxiety via self reports: A comparison of measures." *Communication Monographs, 47,* 149–154.