Reliability and Factor Structure of Budner's Tolerance for Ambiguity Scale

Arlin J. Benjamin, Jr.
Student Outcome Assessment Center
California State University–Fullerton, Fullerton, CA 92634

Ronald E. Riggio
Kravis Leadership Institute
Claremont McKenna College, Claremont, CA 91711

Bronston T. Mayes
Student Outcome Assessment Center
California State University–Fullerton, Fullerton, CA 92634

Budner's (1962) tolerance for ambiguity scale is a well-known and widely used measure of ambiguity tolerance. Its reliability and factor structure were examined in the present study. Four hundred thirty-six undergraduate students completed Budner's scale as part of a student assessment center program. The results of two confirmatory factor analyses failed to substantiate the plausibility of Budner's proposed single-factor model of tolerance for ambiguity, or the four-factor model reported in Furnham (1994). In addition, the Budner scale was shown to have low internal reliability (α = .44) in this sample. The low reliability estimates for this measure, coupled with the apparent lack of a replicable factor structure, suggest that the Budner scale (at least in its present form) is a poor measure of tolerance for ambiguity.

It has been over thirty years since Budner's (1962) measure of tolerance for ambiguity was first published. Budner defined tolerance for ambiguity as "the tendency to perceive ambiguous situations as desirable," whereas intolerance for ambiguity was defined as "the tendency to perceive...ambiguous situations as sources of threat" (p. 29).

An ambiguous situation is one in which the individual is provided with information that is too complex, inadequate, or apparently contradic-

Authors' Notes: This research was supported, in part, by a grant from the U. S. Department of Education, Title VIII Research Grant #84.055 to the second and third authors.

We wish to thank Dr. Carolyn Kubiak, Monica Aguirre, and Chris Belloli for their assistance in conducting this study.

Correspondence concerning this article should be addressed to the first author, who is now at Department of Psychology, University of Missouri–Columbia, Columbia, MO 65211.

Journal of Social Behavior and Personality, 1996, Vol. 11, No. 3, 625–632.

tory (Norton, 1975). Since individuals are bound to be faced with ambiguous situations in everyday life, a reliable measure of tolerance for ambiguity has the potential for practical application. There is some evidence to suggest, for example, that medical students who are high in ambiguity tolerance gravitate toward specialties which are relatively unstructured, compared to medical students who are low in ambiguity tolerance (Budner, 1962). Furthermore, Bray and Grant (1966) suggested that tolerance for ambiguity, when combined with resistance to stress, was related to salary progress in three of the seven samples they examined.

Budner has noted that the construct of ambiguity tolerance is fairly complex. In fact, each of the 16 items in Budner's tolerance for ambiguity scale is designed to indicate one of four different kinds of perceived threat (phenomenological denial, phenomenological submission, operative denial, and operative submission), and one of three different kinds of ambiguous situation (novelty, complexity, and insolubility).

Furnham (1994) tested the factor structure of several measures of tolerance for ambiguity. Budner's (1962) scale was found to measure four distinct factors: predictability (e.g., "What we are used to is always preferable to what is unfamiliar"), variety and originality (e.g., "Often the most stimulating and interesting people are those who don't mind being different and original"), clarity (e.g., "A good job is one where what is to be done and how it is to be done are always clear"), and regularity (e.g., "People who fit their lives to a schedule probably miss most of the joy of living"). These four factors were found to account for over half of the variance.

Although Budner's (1962) scale is one of the better known and more widely used measures of tolerance for ambiguity (Furnham, 1994), reliability estimates for the measure tend to be inconsistent. When Budner tested his scale on 17 different samples, he reported alpha reliabilities ranging from .39 to .62 (with a mean alpha of .49), indicating that the measure was low in internal consistency. More recently, Sobal and DeForge (1992) reported alphas of .63 and .64 when the measure was administered to their two samples. In addition, Furnham (1994) found that Budner's scale had a lower reliability estimate ($\alpha = .59$) than two other measures of tolerance for ambiguity. Furnham reported an alpha of .89 for Norton's (1975) tolerance for ambiguity scale, and an alpha of .78 for Rydell and Rosen's (1966; see also MacDonald, 1970) measure of ambiguity tolerance.

It is also worth noting that Budner's scale did not correlate highly with the other measures of tolerance for ambiguity examined in Furnham's study. The lower correlations between the Budner scale and the other measures of ambiguity tolerance may have been due in part to its rela-

tively low reliability. Test-retest reliability, on the other hand, tends to be somewhat higher. Budner, for example, reported a test-retest reliability of .85 after two weeks when administered to one of the samples in his study. However, Sobal and DeForge (1992) reported a somewhat lower test-retest reliability of .64 when they used the Budner scale.

The internal consistency of a scale is usually considered sufficiently high if estimates of coefficient alpha are above .70, and a recent meta-analysis shows that coefficient alphas of at least .70 are reported in the vast majority of published and unpublished research (Peterson, 1994). Budner (1962) argues that the low internal consistency of his measure is due to the complex nature of the concept of tolerance for ambiguity. Budner contends that the more complex a construct and its corresponding measure are, the lower the reliability estimate will be as a result.

Although a high internal reliability estimate does not guarantee that a scale is unidimensional, it is suggestive of unidimensionality (see, e.g., Cortina, 1993). Low internal consistency, on the other hand, may indicate either that the scale is multidimensional or that the scale is poorly constructed. It is worth noting that when a scale is used as a predictor (or independent) variable, its reliability will influence the degree to which the standardized and unstandardized regression weights are biased (Pedhazur, 1982). Lower reliability is associated with increased bias in the corresponding regression weights, thus reducing the researcher's ability to make accurate predictions or inferences based on the data.

The purpose of the present study is twofold. First, an attempt will be made to determine if either a single-factor model of ambiguity tolerance (Budner, 1962) or the four-factor model described by Furnham (1994) adequately fits the data from our sample. Given the history of low reliability estimates, and the apparent complexity of the tolerance for ambiguity construct, we expect that a single-factor model will not adequately fit the data. What remains to be seen is if Furnham's four-factor model will provide a good fit to the data. Second, the reliability of Budner's (1962) tolerance for ambiguity measure will be examined. We expect the internal consistency of the measure to be marginally acceptable, at best.

METHOD

Participants

The sample in the present study consisted of 436 undergraduate business students from California State University, Fullerton, who had volunteered to participate in a Student Outcome Assessment Center's (SOAC) testing activities (Aguirre, Riggio, Mayes, & Kubiak, 1995). The sample included 162 White or Anglo students, 144 Asian students,

47 Hispanic or Mexican-American students, and 83 students from other ethnic backgrounds. Participants ranged in age from 19 to 65, with an average age of 25. There were 192 men, 195 women, and 49 participants who declined to provide gender information.

Questionnaire and Procedure

The 16-item tolerance for ambiguity scale developed by Budner (1962) was administered as part of a battery of psychological measures. All responses were based on a seven-point scale, from 1 (Strongly Disagree) to 7 (Strongly Agree). The seven-point scale is identical to the one used by Budner (1962). All participants completed the questionnaire while their particular testing session was in progress.

RESULTS

To test the factor structure of the scale, two confirmatory factor analyses using LISREL VII and maximum likelihood estimation (ML) were conducted. Approximately half of the sample (216 cases) was randomly selected for each of these analyses. A covariance matrix was used as input for these analyses. The analyses tested the goodness of fit of both a single-factor model of tolerance for ambiguity (Budner, 1962), and Furnham's (1994) four-factor model. For the single-factor model, we specified that all of the test items would be related to a general factor of tolerance for ambiguity. For the four-factor model, we specified the same relationship between the individual test items and the factors as that reported by Furnham (1994). A model is considered to adequately fit the data if one measure of goodness of fit, χ^2, is low and the probability of obtaining a χ^2 of that size is greater than .05, and if the goodness of fit indicator (GFI) and adjusted goodness of fit indicator (AGFI) are above .90. The results of the factor analyses showed that neither model provided a good fit to the data (see Table 1). An examination of the factor loadings (lambdas) in the single-factor model indicated that a number of the test items failed to load on a general tolerance for ambiguity factor. Similarly, an examination of the four-factor model showed that ten items failed to load on their specified factors (see Table 2).

A cross-validation with the other half of the sample (220 cases) yielded similar results (see Table 3). Once again, the goodness of fit indicators suggested that both the single-factor and four-factor models did not adequately fit the data. An examination of the factor loadings in the single-factor model indicated that a number of the test items failed to load on a general tolerance for ambiguity factor. Similarly, an examination of the four-factor model showed that ten items failed to load on their specified factors (see Table 4).

TABLE 1 Goodness of Fit for Single-Factor and Four-Factor Models of Tolerance for Ambiguity

Model	*df*	χ^2	*Probability*	*GFI*	*AGFI*
Single-Factor (Budner, 1962)	104	265.33	$p < .001$	.86	.81
Four-Factor (Furnham, 1994)	104	299.51	$p < .001$	.85	.80

TABLE 2 Factor Loadings for Single-Factor and Four-Factor Models of Tolerance for Ambiguity

	Single Factor	*Four Factors*			
Item	*Tolerance for Ambiguity*	*Predict-ability*	*Variety and Originality*	*Clarity*	*Regularity*
Definite answer	.38	.28			
Like to live in foreign country	.09	.05			
Problems are solvable	.24				.58
Scheduled lives not enjoyable	-.20				-.36
Good jobs always clear	.50			.11	
Complex problems are fun	.09				-.22
Simple problems are easier	.33			.36	
Original people more interesting	-.06		.36		
Familiar is preferable	.45	.44			
Yes or no answers shallow	-.37		.08		
Few surprises	.43	.48			
Insufficient information	.01		.20		
Prefer familiar people	.16			-.15	
Vague tasks are opportunities	.06		.23		
Similar values desirable	.58	.64			
Good teacher challenges you	.09		.82		

TABLE 3 Goodness of Fit for Single-Factor and Four-Factor Models of Tolerance for Ambiguity: Cross-Validation

Model	*df*	χ^2	*Probability*	*GFI*	*AGFI*
Single-Factor (Budner, 1962)	104	246.95	$p < .001$	.86	.82
Four-Factor (Furnham, 1994)	104	243.69	$p < .001$	.89	.85

TABLE 4 Factor Loadings for Single-Factor and Four-Factor Models of Tolerance for Ambiguity: Cross-Validation

	Single Factor	*Four Factors*			
Item	*Tolerance for Ambiguity*	*Predict-ability*	*Variety and Originality*	*Clarity*	*Regularity*
Definite answer	.56	-.35			
Like to live in foreign country	.16	-.20			
Problems are solvable	-.03				.37
Scheduled lives not enjoyable	.04				-.05
Good jobs always clear	.20			.60	
Complex problems are fun	.26				-.33
Simple problems are easier	.06			.45	
Original people more interesting	.18		-.34		
Familiar is preferable	.27	-.17			
Yes or no answers shallow	-.07		-.15		
Few surprises	.54	-.76			
Insufficient information	-.02		-.24		
Prefer familiar people	.22			.40	
Vague tasks are opportunities	-.05		-.19		
Similar values desirable	.51	-.54			
Good teacher challenges you	.23		-.44		

The above factor analyses were replicated with only the native English-speaking students from the sample. The results for both the single-factor and four-factor models were similar to those found in the previous analyses. Once again, neither model adequately fit the data, indicating that the lack of goodness of fit of the models is not an artifact of language. Subsequent exploratory factor analyses on both halves of the present sample failed to generate a stable set of factors.

The reliability of the tolerance for ambiguity scale was somewhat lower than expected ($\alpha = .44$). One potential explanation for the low reliability estimate is that a large number of students in the present sample were not native English speakers. The reliability of the tolerance for ambiguity measure was substantially lower when administered to non-native English speakers ($\alpha = .33$) than when administered to native English speakers ($\alpha = .49$).

DISCUSSION

The results of the confirmatory factor analyses indicate that neither a single-factor model nor a four-factor model (Furnham, 1994) adequately fits the data. Perhaps such findings are not surprising, given the low reliability of the Budner scale. Although the four factors reported by Furnham are intuitively plausible, the model was simply not replicable in the present study.

It is worth noting that Budner's (1962) scale was not developed for use with non-native English speakers, which comprised a large proportion of this sample. On the other hand, the participants in both Budner's (1962) samples and Furnham's (1994) sample were primarily native English speakers. Given the composition of our sample, a failure to replicate either a single-factor model or Furnham's (1994) four-factor model could conceivably be explained by contending that the scale was only intended for native speakers of English. However, when we analyzed the data of only the native English-speaking students in our sample, we once again failed to replicate either model.

The results of the present study provide further evidence that Budner's (1962) tolerance for ambiguity scale is not a reliable measure. It appears that the measure is even less reliable when administered to non-native English speakers than it is when administered to native English speakers. The wording of many of the items may be ambiguous and confusing, leaving the meaning of these items open to a variety of different interpretations (e.g., "A good job is one where what is to be done and how it is to be done are always clear"). If subjects are misinterpreting the items, or are simply failing to understand the items, this will lead to more random response patterns, which in turn will lead to lower reliability. The results

of the confirmatory factor analyses, along with the low reliability of the measure, suggest that Budner's (1962) tolerance for ambiguity scale does not adequately measure the construct.

Furnham's (1994) recent research indicates that four of the more commonly used measures of ambiguity tolerance contain more than one factor. Such findings suggest that researchers need to take care when choosing a measure of tolerance for ambiguity for their own studies. At the very least, the findings of the present study indicate that researchers desiring a measure of ambiguity tolerance should refrain from using the Budner scale (at least in its present form) in their research.

REFERENCES

Aguirre, M., Riggio, R.E., Mayes, B.T., & Kubiak, C.R. (1995, April). *Assessing critical skills of business students with an assessment center.* Paper presented at the Western Decision Sciences Institute Annual Meeting, San Francisco, CA.

Bray, D.W., & Grant, D.L. (1966). The assessment center in the measurement of potential for business management. *Psychological Monographs: General and Applied, 80*(17, Whole No. 625).

Budner, S. (1962). Intolerance for ambiguity as a personal variable. *Journal of Personality, 30,* 29–50.

Cortina, J.M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78,* 98–104.

Furnham, A. (1994). A content, correlational and factor analytic study of four tolerance for ambiguity questionnaires. *Personality and Individual Differences, 16,* 403–410.

MacDonald, A.P., Jr. (1970). Revised scale for ambiguity tolerance: Reliability and validity. *Psychological Reports, 26,* 791–798.

Norton, R.W. (1975). Measurement of ambiguity tolerance. *Journal of Personality Assessment, 39,* 607–619.

Pedhazur, E.J. (1982). *Multiple regression in behavioral research: Explanation and prediction* (2nd ed.). Fort Worth, TX: Harcourt Brace Jovanovich.

Peterson, R.A. (1994). A meta-analysis of Cronbach's coefficient alpha. *Journal of Consumer Research, 21,* 381–391.

Rydell, S.T., & Rosen, E. (1966). Measurement and some correlates of need cognition. *Psychological Reports, 19,* 139–165.

Sobal, J., & DeForge, B.R. (1992). Reliability of Budner's intolerance for ambiguity scale in medical students. *Psychological Reports, 71,* 15–18.