

This article was downloaded by: [Bodie, Graham]

On: 2 February 2011

Access details: Access Details: [subscription number 933028017]

Publisher Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Communication Research Reports

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t714579429>

A Comparison of Four Measurement Models for the Watson-Barker Listening Test (WBLT)-Form C

Graham D. Bodie^a; Debra Worthington^b; Margaret Fitch-Hauser^b

^a Communication Studies, Louisiana State University, ^b Department of Communication, Journalism at Auburn University,

Online publication date: 02 February 2011

To cite this Article Bodie, Graham D. , Worthington, Debra and Fitch-Hauser, Margaret(2011) 'A Comparison of Four Measurement Models for the Watson-Barker Listening Test (WBLT)-Form C', Communication Research Reports, 28: 1, 32 – 42

To link to this Article: DOI: 10.1080/08824096.2011.540547

URL: <http://dx.doi.org/10.1080/08824096.2011.540547>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

A Comparison of Four Measurement Models for the Watson–Barker Listening Test (WBLT)–Form C

Graham D. Bodie, Debra Worthington, & Margaret Fitch-Hauser

This article compares 4 measurement models for the Watson–Barker Listening Test (WBLT)–Form C and constitutes the first confirmatory test of this listening comprehension measure. Results show that data does not conform to (a) a 5-factor correlated model, (b) a second-order model, or (c) a unidimensional model; and no model was sufficiently better than (d) the independence model. Exploratory analyses provide additional evidence that items are largely unrelated to one another. Given these findings, the use of the WBLT–Form C in assessments of listening comprehension is not recommended. The discussion explores what these findings imply for the conceptualization and measurement of listening and for potential revisions of the WBLT.

Keywords: Assessment; Confirmatory Factor Analysis; Listening Comprehension; Listening Test; Measurement

Since early publications bemoaning the lack of attention communication educators and researchers afford listening (e.g., Adams, 1938; Wiksell, 1946), our field has made significant strides in theorizing and conducting empirical research about this

Graham D. Bodie (PhD, Purdue University, 2008) is Assistant Professor of Communication Studies at Louisiana State University. Debra Worthington (PhD, Kansas University, 1994) is Associate Professor in the Department of Communication and Journalism at Auburn University. Margaret Fitch-Hauser (PhD, University of Oklahoma, 1982) is Associate Professor and Chair of the Department of Communication and Journalism at Auburn University. A previous version of this manuscript was presented at the annual convention of the National Communication Association, Chicago, IL. *Correspondence:* Graham D. Bodie, Department of Communication Studies, Louisiana State University, 136 Coates Hall, Baton Rouge, LA 70803; E-mail: gbodie@lsu.edu.

important communicative function (for reviews, see Bodie, Worthington, Imhof, & Cooper, 2008; Wolvin, 2010). Although several strands of research exist, the primary focus of listening research has been to discover what constitutes good listening (Bostrom, in press), and most of this research has attempted to construct and provide validation evidence for tests of listening comprehension (see Bodie & Fitch-Hauser, 2010). Although several specific tests have been developed (Buck, 2001; Rhodes, Watson, & Barker, 1990), communication scholars have most frequently utilized the Watson–Barker Listening Test (WBLT; Watson, Barker, Roberts, & Roberts, 2001).

Two current versions of the WBLT (Forms C and D) are commercially available to aid researchers, practitioners, and educators in assessing the following listening competencies: (a) evaluating message content; (b) understanding meaning in conversations; (c) understanding and remembering information in lectures; (d) evaluating emotional meanings in messages; and (e) following instructions and directions. For each version, participants are asked to watch a videotape of dialogues and monologues (including lectures) divided into these five sections; after each section, a series of eight questions are posed where participants record their answers on a standard sheet.

To date, only a handful of studies (Johnson & Long, 2007; Worthington, Fitch-Hauser, Cook, & Powers, 2009) have utilized one of these newer versions with Form C being the most popular. Although the test authors (Watson et al., 2001) reported submitting Form C to a “rigorous” revision process, explicit evidence of validity—particularly, evidence of construct validity—is lacking. Indeed, much of the research addressing validity evidence for the WBLT tends to be qualified in some way. For example, validity tests described by the WBLT authors (Watson et al., 2001) often applies to earlier forms of the test, which was 10 items longer and presented in an audio rather than video format (e.g., see Roberts, 1985) or addresses testing issues (e.g., audio vs. video test modes and oral vs. written administrations; Roberts & Vinson, 1993). When addressing construct validity, the test authors primarily described research examining the relationship between the WBLT and other listening tests (e.g., KCLT and STEP) whose own validity is questionable (Fitch-Hauser & Hughes, 1987). Moreover, in each of the studies that have utilized Form C, the researchers have assumed it conforms to the theoretical measurement model described by the test authors. Thus, we are left to question whether the test actually measures what it purports to measure, namely the five components of listening comprehension. Statistically, this model can be tested using confirmatory factor analytic techniques, and that is the purpose of this study.

Not only is a confirmatory test of Form C necessary because it has yet to be demonstrated, but there is also reason to question the factor structure of the scale. Indeed, previous research reporting on earlier versions of the WBLT have found: (a) the proposed five-factor structure has yet to be replicated; (b) factors extracted using exploratory methods account for relatively little item variance; and (c) the convergent and discriminate validity of these forms are questionable (see Bodie & Fitch-Hauser, 2010). Given that no past research has empirically verified the WBLT–Form C measurement model, and given the past work calling into question the five-factor structure using earlier forms, this study reports a confirmatory test of the proposed

measurement model. These data provide much needed assessment of the construct validity of a popular and largely accepted test of listening ability.

Method

Participants

Of the 208 participants that completed all of the WBLT–Form C questions, 87% self-identified as Caucasian, 11% as African American, and the remaining participants identified themselves as belonging to some other ethnic group (e.g., American Indian, Asian, Hispanic, or multiethnic). The average birth year of the participants was 1985.5 ($SD = 1.45$; range = 1980–1988).

Procedure

Data were collected as part of a larger study examining listening comprehension and other listening and communication variables. Upon arrival, participants were provided an informed consent statement. In the one hr-long session, they viewed videotapes containing the WBLT–Form C material and completed all 40 questions, as well as several other measures not relevant to the primary purpose of this study.

Instrument

Watson et al. (2001) reported that the WBLT–Form C measures five facets of listening comprehension: interpreting message content, understanding meaning in conversations, remembering lecture information, interpreting emotional meaning, and following instructions. Following the video presentation of the stimulus material, 40 survey items designed to test each area of comprehension are presented (8 items for each factor). Participant responses are scored as correct or incorrect (see Table 1 for descriptive statistics).

Results

Confirmatory factor analytic procedures (maximum likelihood estimation) were used to assess the fit of the proposed WBLT–Form C measurement model. Based on recommendations by Hu, Bentler, and Kano (1992), the study was sufficiently powered to assess model fit and provide parameter estimates. Given the dichotomous nature of the data, additional constraints were imposed to enable model identification. Specifically, the variance of the five latent constructs representing the five putative listening components were constrained (in addition to the regression weights of the error terms and a single regression weight of a single item from each subscale).

Model 1: Five Interrelated Factors

The first model tested included the five components as latent constructs, each with eight observed variables (the items per question) and associated error terms. Fit statistics for this model (see Table 2) suggested that it was statistically equivalent

Table 1 Number of Right and Wrong Answers for all Subscales and the WBLT Total Score

WBLT subscale	Mode	<i>Mdn</i>	<i>M</i>	<i>SD</i>	Minimum	Maximum	Minimum % correct	Maximum % correct
Evaluating message content	6	6	5.68	1.21	2	8	25	100
Understanding meaning in conversations	6	6	5.46	1.29	0	8	25	100
Understanding and remembering lectures	5	5	4.66	1.55	1	8	25	100
Evaluating emotional meaning in messages	6	6	6.05	1.22	2	8	25	100
Following instructions and directions	6	6	5.71	1.40	1	8	25	100
Total score	29	28	27.55	3.92	14	38	35	95

Number of correct answers								
0	1	2	3	4	5	6	7	8
Evaluating message content	0	0	3	8	19	57	72	46
Understanding meaning in conversations	1	0	2	8	38	55	67	31
Understanding and remembering lectures	0	3	17	30	46	53	37	21
Evaluating emotional meaning in messages	0	0	2	7	11	37	75	63
Following instructions and directions	0	2	2	10	27	41	64	53

Note. WBLT = Watson-Barker Listening Test.

Table 2 Fit Statistics for Theoretical Model 1 and the Competing Independence Model

Model	CMIN	df	p	RMSEA	LO 90	HI 90	AIC	CFI
1	953.66	735	<.001	.038	.031	.045	1,203.66	.000
2	895.94	736	<.001	.032	.024	.040	1,143.94	.126
3	1,431.92	745	<.001	.067	.062	.072	1,661.92	.000
Independence model	1,002.90	820	<.001	.033	.025	.040	1,720.00	.000

Note. CMIN = Chi-square statistic; RMSEA = root mean square error of approximation; LO 90 = Lower bound estimate for RMSEA, 90% confidence interval; HI 90 = Upper bound estimate for RMSEA, 90% confidence interval; AIC = Akaike Information Criterion; CFI = comparative fit index.

to the independence model, a model where the items are not meaningfully related to each other, $\Delta\chi^2 = 49.24$, $p > .10$. Ultimately, this suggests that the items are not related to each other in the manner as specified by the WBLT and, furthermore, that items are largely independent.

Model 2: Second-Order Factor Model

The theoretical Form C model is a second-order model whereby the first-order factors (the subscales) all share variability due to a construct called listening comprehension. In fact, when the test designers suggest tallying the subscales to form a total score, the measurement model behind this suggestion is this second-order model. Thus, this model was tested to see if imposing a second-order latent construct can improve the model fit over that found for Model 1. Model 2 was also only a slight improvement over the independence model (see Table 2).

Model Three: Unidimensional Factor Structure

The final model tested was a one-factor model whereby all 40 WBLT–Form C items loaded on one latent construct called listening comprehension. Support for this model would suggest that Form C is measuring a unitary listening comprehension skill. As seen in Table 2, data does not conform to this model any better than to the previously tested models.

As seen in Table 2, there is a rather large discrepancy between comparative fit index (CFI) and root mean square error of approximation (RMSEA) values for each model, which stems from the fact that they are derived from different formulae (Rigdon, 1996). The high degrees of freedom for each model may help explain why RMSEA falls within acceptable values while the large number of standardized residual covariances above five in absolute value helps explain lack of fit based on CFI. Because all extreme model residuals were negative, each model tested over-specifies relationships between several items.

Additional Analysis

Given the results of the confirmatory analyses, we conducted an exploratory factor analysis (maximum likelihood and varimax rotation) to see what these data suggested

Table 3 Pattern of Maximum Likelihood Factor Loadings, Rotated Matrix

Factor	Part 1 (1–8)	Part 2 (9–16)	Part 3 (17–24)	Part 4 (25–32)	Part 5 (33–40)
1	WB1	WB9			
2				WB30	
3					WB33
4					WB38
5	WB8				
6			WB21		
7			WB17		
8	WB4				
9		WB12		WB25	
10				WB27	
11			WB20		
12	WB2	WB10			
13		WB15			
14			WB24		

Note. WB = Watson Barker. The correlation between WB1 and WB9 (Factor 1) is .12 ($p = .10$). The correlation between WB12 and WB25 (Factor 9) is .21 ($p = .003$). The correlation between WB2 and WB10 (Factor 12) is .23 ($p < .001$).

about Form C. First, we requested a Kaiser–Meyer–Olkin measure of sampling adequacy to ascertain the degree to which our data were appropriate for factor analysis. That measure was .511, which, by the standards set by Kaiser (1974), is a “miserable” amount of common variance among the items.

The resultant factor analysis supported this conclusion. First, replicating research with earlier forms of the WBLT (e.g., Fitch-Hauser & Hughes, 1987), the factors extracted were unable to explain a large portion of the item variance. Indeed, one would have to interpret 13 factors to explain a cumulative 50% of the item variance. Although listening comprehension may be composed of 13 subcomponents, the current model specified by Form C only attempts to measure 5. Second, upon inspection of the factor matrices, no more than two items loaded on any one factor (see Table 3). The correlation matrix upon which the factor analysis was based shows that the

Table 4 Average Inter-Item Correlations for the Watson–Barker Listening Test–Form C Subscales

Subscale	Average r	α
Evaluating message content	.04	.23
Understanding meaning in conversations	.02	.13
Understanding and remembering lectures	.05	.26
Evaluating emotional meaning in messages	.04	.21
Following instructions and directions	.06	.32

average inter-item relationship was .03, with a maximum bivariate correlation of .35. Table 4 shows the average inter-item correlations for each putative subscale and the resultant Cronbach's alpha estimates of internal consistency. By both measures, Form C fails to show evidence of internal consistency with these data.

Discussion

This article sought evidence that the WBLT–Form C validly measures five components of listening comprehension. Although this measure of listening comprehension is reported to have undergone a substantial process of revision, until now researchers were left to assume that the new scale conforms to the specified measurement model. Based on prior work, this assumption was likely ill advised. Indeed, research using alternate forms of the WBLT (Forms A and B) consistently showed instability in empirically generated factor structures across studies and samples (e.g., compare Fitch-Hauser & Hughes, 1987 with Villaume & Weaver, 1996). Overall, our results provide empirical grounds to suggest that Form C should not be used as an assessment of listening comprehension. Not only did the first two confirmatory analyses suggest that the theoretical model assumed by Form C does not explain the covariance structure of the collected data, but results from the confirmatory test of the unidimensional model, as well as results from the exploratory analysis, showed that these 40 items are not strongly related to each other. A core tenet of scale construction and subsequent efforts to provide validity evidence is that items are at least moderately correlated with one another (DeVellis, 2003). So, why are items not more highly correlated?

One reason for the low correlations among items may be found in the reliance on multiple choice questions, scored as right or wrong. Perhaps dichotomous scoring does not fully reflect listening ability, with the valid use of dichotomous scoring likely dependent on context. For instance, the section “understanding meaning” assumes that there is always only one correct meaning of a given utterance. Although there are certainly cases where this may be true (e.g., if a friend says, “I’ll pick you up at 7:00 p.m. at the North entrance to Coates Hall”), in many interactions meaning can be as varied as the number of attendant listeners. Research across the academic landscape suggests that deriving meaning from conversation is more complex than picking out a single, correct meaning (for a review, see Edwards, *in press*). Consequently, right–wrong scoring may misrepresent the multitude of meanings that may be viable alternatives. Indeed, a person who is able to generate multiple alternative meanings from a given utterance may be a more proficient or competent listener (see Bursleson, *in press*).

A second reason may stem from the fact that participants are answering questions based on more information than the test authors intended. For instance, research on listening comprehension within second-language learning shows that “listeners vary in their use of, and their ability to process and utilize the nonverbal components of spoken texts to create meaning from the texts” even with information that was not intended to measure this skill (Wagner, 2008, p. 238). In Form C, evaluating meaning from nonverbal cues is only represented within one of the subscales, yet this skill may help determine scores on most if not all of the five competency areas. Of course, in

most real-world listening situations, people utilize verbal and nonverbal information so, on the surface, this may not be problematic. With respect to constructing listening tests that supposedly measure distinct listening sub-skills, however, the inability to separate those skills is highly problematic. As the WBLT is revised and in other future attempts to develop appropriate measures of listening comprehension, test authors should be mindful to write new items and/or create new vignettes that seek to separate listening into meaningful constituent parts without conflating those parts.

Before scholars engage in the laborious task of generating new test items or creating new videos, however, perhaps basic questions regarding what components make up listening comprehension should be asked again (Bodie et al., 2008). Although the WBLT attempts to assess competence in five areas, there is little rationale for these specific areas and the consequent exclusion of others. For instance, the ability to listen in a way that allows ones interlocutor to feel better in times of stress or the ability to listen in ways that allow for appropriate conflict resolution are not assessed, yet these two types of listening are likely important to interpersonal functioning, relational satisfaction, and well-being. Indeed, a variety of listening “skills” are likely important and, thus, should be assessed in any comprehensive measure of listening comprehension. Unfortunately, although listening scholars are in agreement that comprehension is multidimensional, the various sub-skills important for listening comprehension are not universally accepted (Bostrom, in press).

Of course, viewing listening comprehension as something that should be defined and measured in one way may not be the best approach. According to Kaplan (1964), the conceptualization and operationalization of listening comprehension reflected in the WBLT (and other measures of comprehension) amounts to treating listening as a construct—something “defined on the basis of the observables” (p. 55). The alternative is to consider this concept a theoretical term—one whose “meaning derives from the part it plays in the whole theory in which it is embedded, and from the role of theory itself” (Kaplan, 1964, p. 56). Indeed, scholars have recently argued for theorizing listening rather than treating it as a concept, allowing it to take on different meanings depending on the theoretical structure posed for its explanation (Bodie, 2010; Bostrom, in press). This leaves the assessment of competence in listening theory-dependent and the development of tests reliant on theoretically sophisticated treatments of listening competence (for a similar argument, see Wilson & Sabee, 2003).

In its current form, the WBLT seems to reflect a somewhat outdated view of listening. To be fair, listening scholars have spent considerable time attempting to generate acceptable definitions of listening (see ILA, 1995) and outline the skills that constitute competence in listening (for a review, see Brownell, 2010) with little focus on “theorizing listening” (Bodie, 2009, 2010, in press); thus, this critique is not localized to the WBLT. Although test developers typically draw from theories of memory (Janusik, 2007) or theories that outline “information sources” likely to contribute to retention (Watson & Barker, 1984; Watson et al., 2001), the assumption is still “that if scholars could develop a clear, comprehensive, and consensually agreed upon definition of [listening comprehension] and create reliable and valid measures of that concept, then we could get about the business of developing an encompassing theory” of listening

(Wilson & Sabee, 2003, p. 7). Instead, treating listening as a theoretical term shifts the focus from seeking a “universally accepted listening test” (Watson & Barker, 1984) to analyzing the role and function of listening within a particular theoretical framework.

Whether listening is treated as a construct or as a theoretical term, test developers should be mindful of how to best assess listening (Buck, 2001). Although creating new items to assess some aspects of comprehension may prove fruitful, other aspects important to listening comprehension might be more appropriately measured with existing instruments. For instance, evaluating emotional meanings in messages might be more adequately (and validly) measured using the Profile of Nonverbal Sensitivity tests or reports of emotional intelligence rather than creating new tests (see Hall & Bernieri, 2001). If operationalizations can be found for the various components of listening comprehension of interest to a researcher, the measure of listening comprehension becomes a battery of tests that have garnered evidence for validity; this is similar to the current practice of testing cognitive ability (e.g., Scholastic Aptitude Test, Stanford–Binet Intelligence Scale, and the Wechsler Adult Intelligence Scale; Strauss, Sherman, & Spreen, 2006). This approach seems most in line with treating listening as a theoretical term insofar as the measures utilized to assess listening will differ based on the conceptualization of listening under question. In addition, it is important for researchers to be constantly vigilant in testing the validity of any instruments they use. Indeed, validity is a process, not an end result. This may be particularly important for videobased tests like the WBLT—As hair and dress styles change and cognitive styles vary, the impact of how the test is delivered may have an impact on the usefulness of the test.

Conclusion

The results of this study provide important insights into the ability of researchers and instructors to use the WBLT for the assessment of listening comprehension. As researchers, we have a responsibility to fully report findings pertaining to evidence of validity for constructed tests. Full reporting helps to identify problematic elements of measures and refine them. Our study also highlights the importance of testing both new and established measures, particularly when published support for a measure is lacking. Finally, a primary goal in listening scholarship and study is to expand the understanding of listening and listening processes, and adequate measures are necessary to achieve this goal. Such research helps listening scholars to have greater confidence in the measures they use and, subsequently, the results they report and the conclusions they draw.

References

- Adams, H. M. (1938). Listening. *Quarterly Journal of Speech*, 24, 209–211.
- Bodie, G. D. (2009). Evaluating listening theory: Development and illustration of five criteria. *International Journal of Listening*, 23, 81–103.

- Bodie, G. D. (2010). Treating listening ethically. *International Journal of Listening*, 24, 185–188.
- Bodie, G. D. (in press). The understudied nature of listening in interpersonal communication: Introduction to a special issue. *International Journal of Listening*.
- Bodie, G. D., & Fitch-Hauser, M. (2010). Quantitative research in listening: Explication and overview. In A. D. Wolvin (Ed.), *Listening and human communication in the 21st century* (pp. 46–93). Oxford, England: Blackwell.
- Bodie, G. D., Worthington, D. L., Imhof, M., & Cooper, L. (2008). What would a unified field of listening look like? A proposal linking past perspectives and future endeavors. *International Journal of Listening*, 22, 103–122.
- Bostrom, R. N. (in press). Rethinking conceptual approaches to the study of “listening.” *International Journal of Listening*.
- Brownell, J. (2010). The skills of listening-centered communication. In A. D. Wolvin (Ed.), *Listening and human communication in the 21st century* (pp. 141–157). Oxford, England: Wiley/Blackwell.
- Buck, G. (2001). *Assessing listening*. Cambridge, England: Cambridge University Press.
- Burleson, B. R. (in press). A constructivist approach to listening. *International Journal of Listening*.
- DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). Thousand Oaks, CA: Sage.
- Edwards, R. (in press). Listening and message interpretation. *International Journal of Listening*.
- Fitch-Hauser, M., & Hughes, A. (1987). A factor analytic study of four listening tests. *Journal of the International Listening Association*, 1, 129–147.
- Hall, J. A., & Bernieri, F. J. (Eds.). (2001). *Interpersonal sensitivity: Theory and measurement*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Hu, L., Bentler, P. M., & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin*, 112, 351–362.
- International Listening Association (ILA). (1995). An ILA definition of listening. *Listening Post*, 53, 4–5.
- Janusik, L. (2007). Building listening theory: The validation of the Conversational Listening Span. *Communication Studies*, 58, 139–156.
- Johnson, D. I., & Long, K. M. (2007). Student listening gains in the basic communication course: A comparison of self-report and performance-based competence measures. *International Journal of Listening*, 21, 92–101.
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39, 31–36.
- Kaplan, A. (1964). *The conduct of inquiry: Methodology for behavioral science*. San Francisco, CA: Chandler.
- Rhodes, S. C., Watson, K. W., & Barker, L. L. (1990). Listening assessment: Trends and influencing factors in the 1980s. *Journal of the International Listening Association*, 4, 62–82.
- Rigdon, E. E. (1996). CFI versus RMSEA: A comparison of two fit indexes for structural equation modeling. *Structural Equation Modeling*, 3, 369–379.
- Roberts, C. V. (1985, March). *Preliminary research employing the Watson–Barker Listening Test: A validation of the Instrument*. Paper presented at the meeting of the International Listening Association.
- Roberts, C. V., & Vinson, L. (1993). An investigation of the effect of presentation and response media on listening test scores. *International Journal of Listening*, 7, 54–73.
- Strauss, E., Sherman, E. M. S., & Spren, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). Oxford: Oxford University Press.
- Villaume, W. A., & Weaver, J. B., III. (1996). A factorial approach to establishing reliable listening measures from the WBLT and the KCLT: Full information factor analysis of dichotomous data. *International Journal of Listening*, 10, 1–20.
- Wagner, E. (2008). Video listening tests: What are they measuring? *Language Assessment Quarterly*, 5, 218–243.

- Watson, K. W., & Barker, L. L. (1984). Listening behavior: Definition and measurement. *Communication Yearbook*, 8, 178–197.
- Watson, K. W., Barker, L. L., Roberts, C. V., & Roberts, J. D. (2001). *Watson–Barker Listening Test: Video version/facilitator’s guide*. Sautee, GA: SPECTRA.
- Wiksell, W. (1946). The problem of listening. *Quarterly Journal of Speech*, 32, 505–508.
- Wilson, S. R., & Sabee, C. M. (2003). Explicating communicative competence as a theoretical term. In J. O. Greene & B. R. Burleson (Eds.), *Handbook of communication and social interaction skills* (pp. 3–50). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Wolvin, A. D. (2010). Listening theory. In A. D. Wolvin (Ed.), *Listening and human communication: 21st century perspectives*. Oxford, England: Blackwell.
- Worthington, D. L., Fitch-Hauser, M., Cook, J., & Powers, W. G. (2009, November). *Cultural differences in listening comprehension: An investigation of the Watson–Barker Listening Test*. Paper presented at the annual convention of the National Communication Association, Chicago.