Assessing the Usage Ratings of an Automated Language Intervention

Thesis

Presented in Partial Fulfillment of the Requirements for the Degree Master of Science in the Graduate School of The Ohio State University

By

Elizabeth Anne Linke

Graduate Program in Human Development and Family Science

The Ohio State University

2012

Thesis Committee:

Howard Goldstein, PhD., Advisor

Cynthia Buettner, PhD
Abstract

This study applied a 10-question usage rating survey to ratings of an automated language intervention procedure for teaching vocabulary and comprehension skills to preschoolers. A total of 116 pre-service and in-service teacher participants rated the automated language intervention on the constructs of acceptability, understanding, feasibility, and systems support. The psychometric integrity of a 10-question treatment acceptability instrument, based on the Usage Rating Profile (URP-I; Chafouleas et al., 2009), was examined and confirmed through exploratory factor analysis. Overall, results indicated high usage ratings for the intervention. Pre-service and in-service participants differed in mean ratings in the area of acceptability, which may support a concern in the current treatment acceptability literature over using pre-service teachers as respondents. Findings indicated that a shorter treatment does not compromise its psychometric integrity, and teaching personnel generally approved of this classroom-based, automated language intervention, despite slightly differing in acceptability ratings by years of teaching experience.

*Keywords:* treatment acceptability, factor analysis, automated language intervention, pre-service teachers, in-service teachers
Dedication

Dedicated to the research team at The Ohio State University Center for Response to Intervention in Early Childhood
Acknowledgments

This research would not be possible without the contributions of Dr. Howard Goldstein. Thank you for your invaluable assistance with research and writing. Additionally, I would like to thank the research team at the Center for Response to Intervention in Early Childhood for their assistance. Finally, I wish to acknowledge Dr. Cynthia Buettner for her contributions during the thesis proposal and defense.

Partial support for this research was provided through Cooperative Agreement R324C080011 from the U.S. Department of Education, Institute of Education Sciences, awarded to The Ohio State University.
Vita

2005.................................................................Ursuline Academy

2009.................................................................B.S. Elementary Education, Vanderbilt University

2009.................................................................B.S. Human and Organizational Development, Vanderbilt University

2011-2012 ..........................................................Graduate Teaching Associate, Department of Human Development and Family Science, The Ohio State University

Fields of Study

Major Field: Human Development and Family Science
Table of Contents

Abstract ................................................................................................................................. ii

Fields of Study ................................................................................................................... iv

Vita.................................................................................................................................... v

List of Tables ...................................................................................................................... viii

Chapter 1: Review of Literature ..................................................................................... 1
  Treatment Acceptability .................................................................................................... 2
  History of Survey Development ...................................................................................... 5

Chapter 2: Method ............................................................................................................. 11
  Participants ....................................................................................................................... 11
  Procedure ......................................................................................................................... 12
  Data Entry ....................................................................................................................... 13

Chapter 3: Results ............................................................................................................ 14
  Exploratory Factor Analysis ............................................................................................ 14
  Usage Ratings ................................................................................................................. 15
  Pre-Service and In-Service Personnel Ratings ............................................................... 15
Chapter 4: Discussion .......................................................... 18
References............................................................................. 25
Appendix A: In-Service Survey Instrument................................. 28
Appendix B: Pre-Service Survey Instrument................................. 29
List of Tables

Table 1. Pattern Coefficient Matrix ............................................................................. 16
Table 2. Usage Ratings ................................................................................................. 17
Chapter 1: Review of Literature

In response to recent government mandates, such as No Child Left Behind (No Child Left Behind [NCLB], 2002) and the National Reading Panel (NRP, 2000), education programs, including behavioral and academic interventions, have become increasingly research-based (Feurer, Towne, & Shavelson, 2002; Slavin, 2002). During this period, research has supported the efficacy of a variety of intervention strategies in controlled, laboratory settings (Slavin). However, teachers and paraprofessionals, not researchers, generally implement education interventions in school settings, not outpatient facilities (Gresham, 2004). School-based interventions face unique logistical and structural challenges, particularly when implemented by classroom teachers and paraprofessionals (Gresham). As a result, a specific area of education research focuses on increasing fidelity of implementation of school-based interventions.

Despite demonstrated efficacy in research settings, a strong intervention is not possible with ineffective or inaccurate implementation (Lentz, Allen, & Ehrhardt, 1996). Fidelity of implementation appears to correlate with attitudes toward the intervention (e.g., Allinder & Oats, 1997; Eckert, Hintze, & Shapiro, 1999; Witt & Elliott, 1985), an area of research known as consumer satisfaction (Hawkins, 1991). Therefore, when relying on consumers for implementation, it is necessary to assess attitudes toward the intervention (Elliott & Treuting, 1991). Because teachers are generally responsible for
implementation of school-based interventions (Lentz, Allen, & Ehrhardt), the development of quality tools for assessing consumer satisfaction is an important part of education research.

Treatment Acceptability

Consumer satisfaction is one aspect of the larger construct of social validity, which Wolf (1978) defined as the social desirability, usefulness, and importance of treatment programs. Wolf envisioned social validity consisting of three levels: the social significance of the treatment goals, the social appropriateness of treatment procedures, and the social importance of treatment effects. Social validity is considered an important construct of intervention research due to the theoretical link between acceptability and intervention usage and outcomes (Tarnowski & Simonian, 1992).


Treatment acceptability research is traditionally focused on the second level of social validity: the social appropriateness of treatment procedures (Kazdin, 1980). Kazdin defined treatment acceptability as judgments about treatment procedures by potential consumers. Consumer satisfaction is a subset of treatment acceptability research, which focuses on personal opinions of consumers or potential consumers (Hawkins, 1991). Assessment of treatment acceptability and consumer satisfaction serves multiple purposes in education research. One goal is to improve dialogue between researchers and consumers (teachers, paraprofessionals) (Schwartz, 1996). Assessing treatment acceptability also may prevent or mediate treatment termination, noncompliance, and lack of improvement (Tarnowski & Simonian, 1992). Finally, assessing treatment
acceptability can suggest necessary adjustments to treatment and predict undesired treatment effects (Hawkins, 1991).

Treatment acceptability research has been limited largely to behavioral interventions, presented in analogue situations through hypothetical case studies, as opposed to clinical situations in which consumers implement an intervention before assessing acceptability (Carter, 2007). Further, few studies have assessed academic interventions (Allinder & Oats, 1997). In school settings, treatment acceptability is generally measured through responses to consumer questionnaires about intervention services (Lentz, Allen, & Ehrhardt, 1996). Although analogue research methods present potential ecological validity issues, these methods have provided researchers access to larger, more representative populations for assessing treatment acceptability of behavioral interventions (Carter).

Through the last few decades, researchers have identified a number of factors that help to determine what interventions are likely to be implemented readily and with good fidelity. For example, problem severity appears to strongly influence treatment acceptability (Reimers, Wacker, & Koepl, 1987). Consumers judge interventions more acceptable when the interventions address more severe behaviors (Allinder & Oats, 1997). In addition, the time, materials, and resources required for implementation affects treatment acceptability (Gresham, 1989; Witt, Elliott, & Martens, 1984). Generally, consumers more readily accept and implement interventions that require minimal time, materials, and resources for effective implementation (Allinder & Oats). When consumers perceive interventions to be effective, they are more likely to express high
treatment acceptability (Gresham). Finally, consumers rate interventions with minimal potential side effects more acceptable (Kazdin, 1981). Consumer qualities, such as motivation and philosophical or professional orientations also can affect treatment acceptability ratings of behavior interventions (Gresham; Kazdin & Cole, 1981).

Although the treatment acceptability literature has focused primarily on behavioral interventions, a small number of recent studies have examined acceptability of academic interventions that focus on increasing student performance in areas such as language arts, reading, and mathematics (e.g., de Mesquita & Zollman, 1995). In their research, de Mesquita and Zollman assessed teacher views of a mathematics intervention and identified similar influences on treatment acceptability of academic interventions as on treatment acceptability of behavioral interventions: problem severity, implementation time, required resources and materials, and perceived and/or actual effectiveness of treatment. In addition, de Mesquita and Zollman identified teacher involvement in the development of academic interventions as a key component of treatment acceptability in school settings. Further, teachers judge school-based interventions by availability of necessary personnel and resources, expense of the intervention, and personal and school beliefs.

Only one study was found that examined the treatment acceptability of alternative language interventions (Turan, Ostrosky, Halle, & Destefano, 2004). Preschool teachers tended to view naturalistic (classroom-based) language interventions more acceptable than therapeutic (pull-out) interventions. Acceptability of contrasting language interventions appeared to be influenced greatly by the approaches teachers have studied,
observed, and implemented and teachers’ experiences with specific language intervention strategies. Ratings of acceptability also were higher at lower grade levels and for more severe language delays (Turan et al.).

**History of Survey Development**

Subjective evaluation, generally through surveys and interviews, is the chief method of evaluating consumer satisfaction of behavioral and academic interventions (Carter, 2007). Early studies in the area of treatment acceptability focused mainly on acceptability of specific interventions and comparisons of types of interventions (Carter). One of the first surveys, the *Treatment Evaluation Inventory* (TEI; Kazdin, 1980) evaluated consumer views of acceptability of various types of behavioral interventions. Data analysis from the original study indicated that all items loaded on a unitary factor of acceptability (Kazdin). Since that time, researchers have revised and refined the TEI to focus on other factors, including effectiveness and treatment cost (Reimers & Wacker, 1989) and problem severity and understanding (Reimers, Wacker, & Cooper, 1991).

Witt and Martens (1983) extended treatment acceptability research to educational settings with the development of the *Intervention Rating Profile* (IRP). The primary goal in developing the IRP was to make practitioners aware of treatments acceptable to teachers. Data analysis demonstrated that all items loaded on one factor of acceptability. In 1985, researchers developed the *Intervention Rating Profile-15* (IRP-15) to manage the time-intensiveness of the original IRP and to increase item loading on the acceptability factor (Martens, Witt, Elliott, & Darveaux, 1985). The *Children’s Intervention Rating Profile* (CIRP) modified the IRP-15 to allow children to assess treatment acceptability of
educational interventions (Witt & Elliott, 1985). The CIRP, like the IRP and IRP-15, loaded all items on the unitary factor of acceptability. In 1992, Tarnowski and Simonian abbreviated the IRP-15 to eight items, creating the *Abbreviated Acceptability Rating Profile* (AARP). Results from this study demonstrated that a consumer satisfaction survey could be improved in terms of simplicity, reliability, and reduced administration time and not compromise psychometric integrity (Tarnowski & Simonian).

Researchers have expressed concern about using a single factor of acceptability to determine consumer satisfaction (Carter, 2007). To address this deficiency, Von Brock and Elliott (1987) developed the *Behavior Intervention Rating Scale* (BIRS). Like the Intervention Rating Profile, the BIRS was designed to measure treatment acceptability in school settings. Unlike previous consumer satisfaction rating scales, however, the BIRS measured treatment acceptability and treatment effectiveness as factors in teacher views of interventions (Von Brock & Elliott). The BIRS was a revision and extension of the Intervention Rating Profile-15, adding questions designed to address treatment effectiveness. Data analysis initially revealed three factors: treatment acceptability, treatment effectiveness, and time of effectiveness (Von Brock & Elliott). In 1991, Elliott and Treuting, investigated the construct validity of the BIRS as an instrument to measure teachers’ perceptions of treatment acceptability and perceived treatment effectiveness. Data analysis largely supported the construct validity of the BIRS.

With the development of instruments designed to measure aspects of consumer satisfaction beyond acceptability, researchers identified the need for an instrument that would provide practitioners usable information about teacher perceptions of school-based
interventions (Chafouleas, Briesch, Riley-Tillman, & McCoach, 2009). The *Usage Rating Profile* (URP-I) is a 35-question survey designed to measure four factors of consumer satisfaction: acceptability, understanding, systems support, and feasibility (Chafouleas et al.). Chafouleas et al. argued that moving beyond treatment acceptability, as the sole factor in consumer satisfaction research, would increase the likelihood of successful implementation of school-based interventions. Investigators have argued that the factors of acceptability, understanding, systems support, and feasibility offer multiple dimensions for understanding intervention usage. Using principal axis factoring with an oblique rotation resulted in a factor structure in which the four factors accounted for 49% of the common variance in the items (Chafouleas et al.).

Currently, researchers generally agree that the construct of consumer satisfaction is well represented by the factors identified by Chafouleas et al. (2009): acceptability, understanding, systems support, and feasibility. However, Chafouleas et al. cautioned against their use with pre-service teachers as research subjects, concerned that there were fundamental differences between pre-service teachers and more experienced teachers, particularly in the areas of feasibility and systems support. As a large percentage of consumer satisfaction studies enroll pre-service participants as potential consumers, this possible discrepancy needs to be evaluated directly.

This review highlights improvements in the tools for evaluating treatment acceptability. However, there are three areas that require further investigation. First, there is a need for brevity in consumer satisfaction measures. To address the need for brevity, we examined the items representing the factors of acceptability, understanding, systems
support, and feasibility, as identified by Chafouleas et al. (2009) in the URP-I, and developed a 10-question survey. The primary goal of this study was to determine if the four factors of acceptability, understanding, feasibility, and systems support, were still derived with a briefer questionnaire. We hypothesize that the developed treatment acceptability instrument will display a factor structure similar to that of the URP-I, thus demonstrating its validity as a brief, easy-to-use measure of treatment acceptability.

Second, there is a need to extend the treatment acceptability literature to academic interventions. This study focuses on an academic intervention designed to address language delays in early childhood. The intervention, still under investigation, is designed by the Center for Response to Intervention in Early Childhood (CRTIEC) (Spencer et al., 2012). The language intervention focuses on increasing vocabulary in four year olds through a series of prerecorded books with embedded vocabulary lessons. Children listen to the books and respond to automated questions, instructions, and definitions.

We desire to assess acceptability for this intervention, which, due to its automated nature, may evoke more varied opinions of acceptability than would more “typical” language interventions. We hypothesize that overall acceptability will be relatively high, due to documented efficacy of the intervention (Gresham, 1989; Spencer et al., 2012), minimal side effects (Kazdin, 1981), and minimal time needed for implementation (Allinder & Oats, 1997). Further, because the intervention is classroom-based and intended for preschool, as opposed to older children, we expect high levels of overall acceptability ratings, based on the findings of Turan et al. (2004). However, a considerable amount of resources, including books, headphones, and automated
recordings of the stories, are required, which could negatively affect acceptability ratings (Allinder & Oats). It is also conceivable that individual factors could differ in positive and negative ratings. For example, the automated nature of the intervention minimizes the amount of time teachers must spend on implementation, perhaps increasing feasibility and systems support ratings (Allinder & Oats). The area of understanding, however, could be negatively impacted by lack of motivation (Gresham) or professional orientation (de Mesquita & Zollman, 1995; Kazdin & Cole, 1981), if participants are uncomfortable or less knowledgeable with technology and automated interventions.

Third, the use of pre-service personnel as potential consumers of school-based interventions remains a concern in the treatment acceptability literature. Thus, this paper will compare acceptability ratings of pre-service and in-service personnel to determine if fundamental differences exist. Turan et al. (2004) concluded that acceptability of language interventions was greatly influenced by the approaches teachers have studied and implemented, as well as their experience with language intervention strategies. We hypothesize that the in-service participants in this study have had more experiences with language interventions than pre-service participants. Therefore, we predict differences in overall acceptability ratings between the two groups, though it is difficult to predict the direction of differences. For example, if in-service participants have had positive experiences with language interventions, their acceptability ratings may be higher than pre-service participants. The reverse could also be true, however, if in-service participants do not agree with the automated approach to vocabulary learning or the embedded nature of the intervention.
In summary, this study addresses three critical questions in the area of treatment acceptability of school-based interventions:

1. Does a 10-item survey derived from the URP-I yield the same four-factor structure of treatment acceptability as the longer instrument?

2. How do individuals introduced to an automated language intervention rate the procedures with respect to acceptability, understanding, feasibility, and systems support?

3. Do pre-service and in-service personnel differ in their mean ratings of these four dimensions of treatment acceptability?
Chapter 2: Method

Participants

Participants in this study included 94 undergraduate students enrolled in a curriculum development course at Ohio State University and 22 pre-kindergarten teachers working in Columbus City Schools.

Researchers recruited undergraduate participants from two sections of a curriculum development course with permission of the instructors. All potential participants viewed a demonstration and introduction to the language intervention and had access to an informed consent document, stating their rights as participants in this research. Consenting students were asked to complete the survey (see Appendix B) with regard to the information presented; 94 out of 104 undergraduate students consented to the study and completed surveys. The majority of students (70%) indicated an intention to pursue a career in education.

Pre-kindergarten teachers were recruited through efforts by the Center for Response to Intervention in Early Childhood (CRTIEC) at The Ohio State University. CRTIEC researchers selected teachers for their study based on school recommendations by an administrator in the Columbus City Schools. Teachers did not self-select for the CRTIEC study. Teacher participants in our study were recruited during a training session for the 2012-2013 CRTIEC study. Consenting teachers completed the consumer satisfaction survey (see Appendix A) after a day of training presented by research
scientists and research assistants working on the CRTIEC study at Ohio State University. 22 out of 22 teachers consented to the study and completed the survey. In both cases, participants completed a consumer satisfaction survey before implementation of the intervention.

In all, 116 out of 126 (92%) surveys were returned and usable in factor analysis.

**Procedure**

During one session of an undergraduate education course at The Ohio State University, researchers outlined the procedures, goals, materials, and initial research outcomes of the CRTIEC Tier 2 Vocabulary and Comprehension intervention. They also presented a recorded demonstration of the embedded vocabulary and comprehension instruction. Researchers prepared a standard presentation and allowed time for clarifying questions as needed. The training lasted approximately 60 minutes.

Teacher participants completed the survey after a 2-hour training session with researchers and developers of the CRTIEC Tier 2 Vocabulary and Comprehension intervention. Similar to the undergraduate participants, the teachers learned procedures, goals, materials and initial research outcomes. This training was longer and more in depth than the undergraduate training sessions, with more time to explore the intervention materials and practice procedures. If differences arise in mean usage ratings, it is possible that differences in training sessions could be a factor.

**Item development and content validation.** Researchers developed items based on the Usage Rating Profile – Intervention (URP-I; Chafouleas, Briesch, Riley-Tillman, & McCoach, 2009). The initial study of the 35-item URP-I resulted in four factors: acceptability, understanding, feasibility, and systems support. In this study, researchers
selected questions based on high factor loadings and reworded the questions for appropriateness and use with a language intervention. The final result was a 10-question survey with questions addressing four hypothesized constructs: four questions addressing acceptability, two addressing understanding, three addressing feasibility, and one question addressing systems support.

**Data Entry**

Researchers entered pre-service and in-service raw data into SPSS, a statistics software program, treating each of the 10 questions in the survey as a separate variable. Researchers conducted exploratory factor analysis to determine if the 10 questions were appropriate for determining ratings of acceptability, understanding, feasibility, and systems support. Researchers identified mean usage ratings for 88 of 116 participants (22 of 22 in-service participants and 66 of 88 pre-service personnel – those who indicated intention to pursue a career in education), as well as for the subgroups of in-service and pre-service personnel. Further, researchers determined mean ratings of acceptability, understanding, feasibility, and systems supports for 88 pre-service and in-service participants and for each subgroup separately. To determine if mean differences existed between pre-service and in-service respondents, researchers conducted between-group t-tests on mean ratings of acceptability, understanding, feasibility, and systems support.
Chapter 3: Results

Exploratory Factor Analysis

The first goal of this study was to determine if a 10-question, treatment acceptability survey could produce reliable and valid results of consumers’ ratings of acceptability, understanding, feasibility, and systems support. To this end, we conducted an exploratory factor analysis (EFA) to determine if the results from the entire data set (in-service and pre-service participants) would factor into the four constructs identified by Chafouleas et al. (2009) in their development of the Usage Rating Profile-I.

To determine if exploratory factor analysis (EFA) was appropriate for this data, we looked at several key indicators. First, a correlation matrix of the data showed no evidence of multicollinearity; that is, no two survey items exhibited correlations greater than 0.8. Item 10 appeared minimally correlated ($r < 0.3$) with all other items. Because this was the only item representing the systems support factor, minimal correlation with other items was expected. Overall, the correlation matrix demonstrated data appropriate for EFA.

Next, we examined the results of the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy and Bartlett’s Test of Sphericity. The KMO test is used to determine if the sample size of the data is sufficient in relation to the number of survey items. The KMO test was .811, demonstrating that the patterns of correlation are relatively compact, increasing the reliability of EFA. The Bartlett’s Test determines if the correlation matrix is an identity matrix, which would make the data inappropriate for EFA. Because
Bartlett’s test was significant ($p < .05$), we can confidently conclude that the correlation matrix is not an identity matrix. Based on these indicators, we concluded that exploratory factor analysis was appropriate given our data set.

Given this conclusion, we ran an exploratory factor analysis on our survey response data. Using principal axis factoring with an oblique rotation, four factors were extracted. The pattern coefficient matrix demonstrates the dimensionality of each survey item. Results presented in Table 1 reveal that each survey item is highly loaded ($r > .67$) onto its primary factor. This indicates that four factors is a good fit for these data and that items fall into expected factors. The structure matrix indicates a moderate degree of correlation between factors, confirming our decision to use an oblique rotation. Cronbach’s Alpha (.897) indicated a high level of internal consistency for this survey, lending further support to the validity of our survey.

Usage Ratings

As can be seen in Table 2, mean usage ratings for all participants for this language intervention were fairly high. For 88 pre-service and in-service participants, the mean usage rating was 4.99 ($S = .78$), on a Likert-type scale of 1-6, with 6 being the highest rating. Examining mean ratings for each factor demonstrates slight differences: The mean ratings (with standard deviations in parentheses) for each factor were: acceptability 4.90 (.91), understanding 5.33 (.72), feasibility 4.93 (.86), and systems support 4.99 (1.03).

Pre-Service and In-Service Personnel Ratings

As can be seen in Table 2, the mean ratings appear higher for the in-service participants than for the pre-service personnel. To determine if mean ratings differed significantly, we ran between-group t-tests for overall usage ratings, as well as mean
ratings in the four identified areas of treatment acceptability. Participants’ ratings differed significantly in only one area: acceptability, \( t(86) = 2.15, p < .05 \). In the areas of overall usage, understanding, systems support, and feasibility, differences in mean ratings by pre-service and in-service personnel were not statistically significant.

**Table 1:**

*Pattern Coefficient Matrix*

<table>
<thead>
<tr>
<th>Component</th>
<th>Acceptability</th>
<th>Understanding</th>
<th>Systems Support</th>
<th>Feasibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>I would be excited to use this intervention.</td>
<td>.695</td>
<td></td>
<td>.210</td>
<td></td>
</tr>
<tr>
<td>I understand how to use this intervention.</td>
<td></td>
<td>.862</td>
<td></td>
<td>.156</td>
</tr>
<tr>
<td>This intervention is a good way to handle language delays.</td>
<td>.860</td>
<td>.136</td>
<td></td>
<td>-.161</td>
</tr>
<tr>
<td>The directions for using this intervention are clear to me.</td>
<td></td>
<td></td>
<td></td>
<td>.928</td>
</tr>
<tr>
<td>I am motivated to try this intervention.</td>
<td>.833</td>
<td></td>
<td>-.106</td>
<td>.114</td>
</tr>
<tr>
<td>The amount of time required to use this intervention is reasonable.</td>
<td>.118</td>
<td>.205</td>
<td></td>
<td>.674</td>
</tr>
<tr>
<td>This intervention is an effective choice for addressing a variety of problems.</td>
<td>.784</td>
<td></td>
<td>-.113</td>
<td>.244</td>
</tr>
<tr>
<td>The intervention could be implemented as frequently as desired.</td>
<td></td>
<td></td>
<td></td>
<td>.816</td>
</tr>
<tr>
<td>The amount of time required for record-keeping with this intervention is reasonable.</td>
<td></td>
<td></td>
<td></td>
<td>.103</td>
</tr>
<tr>
<td>Implementation of this intervention would require support from my coworkers.</td>
<td></td>
<td></td>
<td></td>
<td>.965</td>
</tr>
</tbody>
</table>

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.
a. Rotation converged in 8 iterations.

Note: Primary factor loading in bold.
Note: Factor loadings < 0.1 not on table.
<table>
<thead>
<tr>
<th>Factor</th>
<th>All Participants (n=88) M (SD)</th>
<th>In-Service (n=22) M (SD)</th>
<th>Pre-Service (n=66) M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usage (Overall)</td>
<td>4.99 (.78)</td>
<td>5.24 (.82)</td>
<td>4.91 (.87)</td>
</tr>
<tr>
<td>Acceptability</td>
<td>4.90 (.91)</td>
<td>5.25 (.74)</td>
<td>4.78 (.93)*</td>
</tr>
<tr>
<td>Understanding</td>
<td>5.33 (.72)</td>
<td>5.36 (.79)</td>
<td>5.31 (.71)</td>
</tr>
<tr>
<td>Feasibility</td>
<td>4.93 (.86)</td>
<td>5.20 (.89)</td>
<td>4.84 (.83)</td>
</tr>
<tr>
<td>Systems Support</td>
<td>4.89 (1.03)</td>
<td>5.09 (.97)</td>
<td>4.82 (1.05)</td>
</tr>
</tbody>
</table>
This paper addressed three main questions:

1. Does a 10-item survey derived from the URP-I yield the same four-factor structure of treatment acceptability as the longer instrument?

2. How do individuals introduced to an automated language intervention rate the procedures with respect to acceptability, understanding, feasibility, and systems support?

3. Do pre-service and in-service personnel differ in their mean ratings of these four dimensions of treatment acceptability?

Based on our sample of 116 participants, we found our 10-question treatment acceptability instrument displayed the same factor structure as the 35-item Usage Rating Profile from which it was derived. The results of exploratory factor analysis supported our hypothesis in that four factors were a good fit for our data, and each survey item was highly loaded onto the factor that it was supposed to represent. Further, the high Cronbach’s alpha (.897) indicates a high level of internal consistency in our survey instrument. These findings support our attempt to create a less burdensome, more consumer-friendly survey without compromising its psychometric integrity. Ultimately, this simplifies administration of and responses to surveys of treatment acceptability. Decreasing the length of a survey without compromising its factor structure could strengthen the construct validity of the survey, in that consumers have a smaller burden of questions to consider when responding. Further, shorter surveys could increase response rates, which could strengthen internal validity of studies.
Our second research question addressed how individuals would rate an automated language intervention, given differences in administration and organization from more typical language interventions. We expected relatively high overall usage ratings from both pre-service and in-service personnel, and this hypothesis was supported by our results. Overall, participants indicated high ratings for this intervention, with a mean overall usage rating of 4.99 on a 6 point Likert-type scale. Results also demonstrated high mean ratings in individual factors, especially for the construct of acceptability, 5.33 ($S = .72$). It appears that this group of participants views an automated language intervention quite favorably.

This high overall usage rating supports research findings that acceptability is positively affected by demonstrated efficacy of the intervention (Gresham, 1989), minimal side effects (Kazdin, 1981), and minimal time and effort required for implementation (Allinder & Oats, 1997). In training sessions, researchers explained initial results of this language intervention (Spencer et al., 2012) through charts, graphs, and videos, perhaps increasing acceptability ratings. Further, researchers emphasized that the automated nature required little time or effort by teachers to implement, perhaps a factor in high feasibility and systems support ratings. Initial research of this intervention has found no negative side effects (Spencer et al., 2012), which also could have contributed to higher acceptability ratings. Additionally, because this language intervention is classroom-based and intended for preschool children, high acceptability ratings agree with the findings of Turan et al. (2004), who found that preschool teachers reported higher treatment acceptability for classroom-based interventions than did teachers of older children.
There were small, not statistically significant, differences in mean ratings for the four identified constructs of acceptability, understanding, feasibility, and systems support. We hypothesized that while feasibility and systems support may be positively impacted by the relatively lower burden of implementing an automated intervention versus a teacher-led intervention, the construct of understanding could be negatively impacted by lack of experience and/or comfort with technology and automated interventions. This did not appear to be the case, as the means of each construct were not significantly more positive or more negative than the overall usage rating mean. Overall, the high mean ratings lend credibility to automated language interventions as an important tool in classroom-based interventions.

Our third and final research question addressed a potential discrepancy in ratings between pre-service and in-service personnel. Treatment acceptability research has used pre-service personnel, particularly undergraduate and graduate students in education programs, to assess acceptability of behavioral and academic interventions, largely due to convenience factors (Chafouleas et al., 2009). Chafouleas et al. cautioned against this, voicing concerns of fundamental differences between these groups, particularly in the areas of feasibility and systems support. Further, Turan et al. (2004) found that experience with language interventions had a significant effect on acceptability ratings, a factor that could cause discrepancies between acceptability ratings of pre-service personnel and in-service personnel.

Therefore, we hypothesized a difference in acceptability ratings due to the likelihood that in-service participants have had more experience with language interventions than pre-service participants. Turan et al. (2004) found that more positive
acceptability ratings correlated with positive experiences with language interventions, and lower ratings correlated with negative experiences. Although we hypothesized a difference in acceptability ratings, we did not predict a direction of difference, due to lack of information on participants’ experiences with and views toward language interventions. Ultimately, our results demonstrated a statistically significant difference in mean acceptability ratings between pre-service and in-service participants, $t(86) = 2.15, p < .05$. In-service participants rated the intervention more acceptable ($M = 5.25, S = .74$) than pre-service participants ($M = 4.78, S = .93$). With a medium effect size of $.53$, we can conclude that teaching experience may have a moderately positive effect on acceptability ratings of language interventions.

It is important to note that despite the concerns of Chafouleas et al. (2009), our study found no significant differences in the constructs of feasibility and systems support. We also did not find significant differences in mean ratings of overall usage or in the area of understanding. This could indicate that education programs are sufficiently preparing pre-service personnel for the challenges of implementing classroom-based interventions. Thus, while experience does appear to correlate positively with mean acceptability ratings, there is no correlation between experience and mean ratings of understanding, feasibility, and systems support, all of which contribute to the overall construct of consumer satisfaction.

There are limitations to consider in these findings. First, it is possible that the significant difference in mean acceptability ratings could be due to differences in training sessions. Because the training sessions were run slightly differently, with different presenters and different lengths of time, it is possible that a longer training session for in-
service participants contributed to higher ratings of acceptability. Turan et al. (2004) concluded that knowledge of language interventions could affect acceptability ratings. Thus, if in-service participants had more knowledge of this language intervention, it could have positively affected their mean ratings of acceptability. Further research into this possibility is necessary. Beyond differences in training sessions, we do not have sufficient background information on the participants to conclude why differences occurred in the area of acceptability, but not understanding, feasibility, or systems support. Future studies should examine prior experience with interventions, language delays, and general attitudes toward interventions to shed light on this issue.

Although the findings of this study are important in the field of treatment acceptability research, there are possible threats to external validity that must be addressed. First, we caution against generalizing these results without more research, as the participants in this study all teach or attend school in Columbus, Ohio. The teachers are employed in a large, urban school district, and the undergraduate students are enrolled at a large, urban public university. This could limit the generalizability of the findings, if these participants differ fundamentally from pre-service and in-service teaching personnel in other areas of the country.

The findings of this paper add to the treatment acceptability literature and also provide researchers with importance issues for future research. Because our study was conducted solely in a large city in a Midwestern state, it is important that future research in the field of language interventions focus on teachers in different areas of the country and of different racial, ethnic, and socioeconomic backgrounds, in order to learn more about attitudes toward automated and teacher-directed language interventions. Our
finding of a significant difference in mean acceptability ratings of pre-service and in-service teachers indicates a need for further research into the relationship between acceptability ratings and teaching experience. Because our participants experienced slightly different training sessions, it is also important to conduct follow-up studies in which groups participate in highly similar and highly dissimilar training sessions. Such studies could identify a correlation between intensity of training sessions and differences in acceptability ratings.

Our study also indicated a high acceptance rate for an automated language intervention, an encouraging finding in an education world increasingly focused around technology. To assess differences in usage ratings before and after implementing an intervention, we recommend administering the survey after teachers have completed their involvement in the CRTIEC study. We also recommend research that directly compares automated and teacher-directed language interventions. Such research could indicate differences in mean ratings of acceptability, understanding, feasibility, and systems support, and provide researchers insight into teachers’ views of technological advancements in classroom-based interventions. Further, we recommend investigating the relationship between acceptability ratings of an automated intervention and fidelity of implementation. Because teachers generally implement classroom-based interventions, it is critically important to obtain information on their attitudes toward language and other academic interventions.

In summary, our findings offer several contributions to the treatment acceptability literature. The results of exploratory factor analysis indicate that a 10-question, user-friendly survey instrument does not compromise its psychometric integrity. Researchers’
use of shorter surveys could lead to increases in response rates and decreases in participant fatigue, both important factors in internal validity. This study also extends the treatment acceptability literature to language interventions, specifically automated interventions. Because treatment acceptability literature has focused primarily on behavioral interventions, these results are critically important when considering how teachers view classroom-based, automated language interventions. Our results indicate high mean ratings of acceptability, understanding, feasibility, and systems support, thus increasing confidence in fidelity of implementation and positive outcomes for children. Our results also indicate slight differences acceptability ratings by pre-service and in-service teachers, which could caution researchers from using undergraduate and graduate education students as respondents in treatment acceptability studies. However, mean usage ratings did not differ significantly, which lends credibility to the use of pre-service teachers’ responses when a survey instrument has a focus beyond treatment acceptability.
References


Appendix A: In-Service Survey Instrument

Center for Response to Intervention in Early Childhood
Tier 2: Vocabulary and Comprehension

Based upon information presented concerning a Tier 2 Vocabulary and Comprehension intervention, we are seeking your feedback, as a classroom teacher preparing to implement this study. Circle the number that best describes your agreement or disagreement with each statement. Please answer every question.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Somewhat Disagree</th>
<th>Somewhat Agree</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>I would be excited to use this intervention.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>I understand how to use this intervention.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>This intervention is a good way to handle language delays.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>The directions for using this intervention are clear to me.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>I am motivated to try this intervention.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>The amount of time required to use this intervention is reasonable.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>This intervention is an effective choice for addressing a variety of problems.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>The intervention could be implemented as frequently as desired.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>The amount of time required for record-keeping with this intervention is reasonable.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Implementation of this intervention would require support from my coworkers.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>
Appendix B: Pre-Service Survey Instrument

Based upon information presented concerning a Tier 2 Vocabulary and Comprehension intervention, we are seeking your feedback about the procedures. Imagine that you are a classroom teacher considering the use of this intervention. Circle the number that best describes your agreement or disagreement with each statement. Please answer every question.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Somewhat Disagree</th>
<th>Somewhat Agree</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>I would be excited to use this intervention.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>I understand how to use this intervention.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>This intervention is a good way to handle language delays.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>The directions for using this intervention are clear to me.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>I am motivated to try this intervention.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>The amount of time required to use this intervention is reasonable.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>This intervention is an effective choice for addressing a variety of problems.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>The intervention could be implemented as frequently as desired.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>The amount of time required for record-keeping with this intervention is reasonable.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Implementation of this intervention would require support from my coworkers.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>I intend to pursue a career in education.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>