Yes You Can: The Effects of a Module to Teach Preference Assessments and Least-to-Most Prompting Procedures

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By
Marnie Nicole Shapiro
Graduate Program in Educational Studies

The Ohio State University
2016

Dissertation Committee:
Dr. Nancy Neef, Advisor
Dr. Moira Konrad
Dr. Sheila Morgan
Dr. Stephanie Peterson
Abstract

Applied behavior analysis (ABA) has been recognized as the gold standard treatment of choice for children with autism. Therefore, training staff to implement behavioral technology is of paramount importance because the efficacy of behavior change plans depends on staff implementing them with high fidelity. Most training procedures have commonly included two or more of the following components: instructions, modeling, rehearsal, and feedback. However, expert-facilitated training is costly in terms of time and resources. This study evaluated the effects of a self-instructional module to train six teachers to implement a paired-stimulus preference assessment and least-to-most prompting procedures. The findings from a multiple probe design across participants and skills revealed that six participants achieved mastery with the introduction of the module and results generalized to clients with whom they were working. Additionally, participants indicated overall satisfaction with the goals, procedures, and outcomes of the study, which suggests that the use of self-instructional packages are a socially valid method for providing efficient, effective training to staff in the absence of an expert trainer.

Keywords: staff training, self-instructional packages, preference assessments, least-to-most prompting
Acknowledgments

The process of earning a doctorate and writing a dissertation is long and arduous and it is certainly not done singlehandedly. I would like to thank Ziwei and Gigi for their unwavering support throughout this process. You stood by me through both highs and lows. No matter where life takes us, and no matter how far we are from each other – the memories of our friendship will remain in my heart forever. I would like to express my sincerest gratitude to my dissertation committee, Dr. Konrad, Dr. Morgan, and Dr. Peterson. Dr. Konrad, you have given me some excellent advice—“try to be in the moment”—that I will continue to follow as I make my way through life. I am very grateful for all that you have done for me at OSU. Dr. Morgan, thank you for your support and commitment throughout this project and beyond. Dr. Peterson, I am honored that you agreed to be a committee member on my dissertation, despite your incredibly busy schedule; I thank you for your time and comprehensive feedback during this process.

Most importantly, I would like to thank my advisor and dissertation chair, Dr. Neef. In all honesty, it is a formidable task to condense a 3-year history into a brief acknowledgement page. Part of the difficulty lies in selecting the right stories to share, but perhaps the more difficult part is selecting the right words to describe a history with a person whom I am both privileged and honored to consider a mentor. You have served as
an excellent model of the type of scholar, professor, and scientist that I hope to emulate. You have been instrumental in my growth and development at OSU. You have unconditionally supported me throughout my entire journey and provided me with both personal and professional support along the way. For all of this, I will always have a deep sense of gratitude and respect.
Vita

2009.................................................................B.A. Undergraduate, California State University, Northridge

2013.................................................................M.S. Clinical Psychology, California State University, Northridge

2013 to present ..............................................Graduate Research Assistant, Department of Educational Studies, The Ohio State University

Publications


Shapiro, M., Kazemi, E., & Weiner, B. (2013). Whose fault is it anyway: How do
doi:10.1007/s11218-012-9200-8

Field of Study

Major Field: Educational Studies
Area of Emphasis: Special Education & Applied Behavior Analysis
# Table of Contents

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

Acknowledgments ........................................................................................................................ iii

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

Table of Contents ........................................................................................................................ vii

List of Tables ................................................................................................................................ xii

List of Figures ............................................................................................................................. xiii

Chapter 1: Introduction ....................................................................................................................... 1

  The Importance of Training ........................................................................................................ 2

  Current Training Trends ............................................................................................................. 5

  Previous Reviews in Training ..................................................................................................... 5

  Purpose Statement ..................................................................................................................... 7

Chapter 2: Literature Review .......................................................................................................... 10

  Literature Review Methodology ............................................................................................... 10

  Coding and Reliability .............................................................................................................. 12

  Participants and Setting ............................................................................................................. 12

  Interventions Targeted for Training .......................................................................................... 13
PS assessment ...................................................................................... 61

Scoring and interpretation materials for the PS assessment ............... 62

Response Measurement ........................................................................ 62

Participant target response for LTM procedure ............................ 62

Participant target response for the PA assessment .......................... 63

Participant target response for graphing student performance ...... 64

Participant target response for the PA scoring and interpretation .... 65

Duration spent reviewing the manual ................................................ 65

IOA Agreement and Measurement Fidelity .................................... 65

Correct implementation ................................................................. 65

Correct graphing of the LTM procedure .......................................... 66

Correct assessment summaries and interpretation ........................ 66

Experimental Design ....................................................................... 67

Procedure ....................................................................................... 68

Baseline ......................................................................................... 68

Written instructions for the LTM procedure ................................. 68

Written instructions for the PS assessment .................................... 69
Appendix H: Procedural Fidelity Checklist…………………………………………………..113

Appendix I: Social Validity Questionnaires………………………………………………..115
List of Tables

Table 1. A list of all articles included in the literature review and the results for some of the main variables under investigation………………………………………………..43

Table 2. Mean percentage of correct trials for specific target responses across participants for the LTM procedure…………………………………………………………………...43

Table 3. Mean percentage of correct trials for the specific target responses across participants for the PS assessment……………………………………………………….45

Table 4. Participants' responses to the social validity questionnaire for the LTM procedures………………………………………………………………………………………………64

Table 5. Participants' responses to the social validity questionnaire for the PS assessment………………………………………………………………………………………….65
List of Figures

Figure 1: LTM assessment implementation graphs……………………………………..59

Figure 2: PS assessment implementation graphs………………………………………..61
Chapter 1: Introduction

The burgeoning demand to access behavioral services necessitates the development of effective, efficient training procedures to teach novice staff accurate implementation of widely used behavioral technology. Consequently, researchers have developed evidence-based training procedures that incorporated antecedent- (e.g., instructions, video modeling) and consequent-based (e.g., performance feedback) strategies to teach staff the requisite skills to implement behavior change plans (Higbee, Akers, & Brodhead, 2014; McCulloch & Noonan, 2013; Moore & Fisher, 2007; Pollard, Petscher, & Bailey, 2006; Rosales, Stone, & Rehfeldt, 2009; Roscoe & Fisher, 2008; Severtson & Carr, 2012; Vladescu, Paden, & Kodak, 2012).

Although there is a body of literature focused on the development of effective training strategies to employ with staff, researchers have found that at the conclusion of the training many individuals still have not acquired the target skills that the training procedures were intended to teach (Clark, Cushing, & Kennedy, 2004; McCulloch & Noonan, 2013; Sturmey, 1998). From a clinical standpoint, concerns regarding the adequacy of staff training are warranted because of the potential deleterious effects improper implementation can have on consumers as well as staff (Hansford, Zilber, LaRue, & Weiss, 2010). Despite the importance of training, to our knowledge, there are relatively few articles in which the authors have reviewed the training literature and
outlined the various strategies used during staff training (e.g., Brock & Carter, 2013; Rispoli, Neely, Lang, & Ganz, 2011; Snyder et al., 2012).

**The Importance of Training**

Treatment fidelity is defined as implementing a behavior change plan as originally designed and intended (Gresham & Gansle, 1993; Gresham, MacMillan, Beebe-Frankenberger, & Bocian, 2000). The focus of treatment fidelity is on how consistently and accurately individuals implement each component of the existing intervention. Monitoring treatment fidelity provides vital information regarding *how* an intervention is being implemented, which greatly facilitates assessment of its effectiveness. In clinical research, experimenters operate under highly controlled settings, and behavioral interventions are likely to be implemented with high fidelity. However, in applied settings (e.g., a classroom) this fact may not always hold true (Wilder, Atwell, & Wine, 2006). Unfortunately, compromised treatment integrity of instructional strategies may result in variable or even poor student outcomes (Carroll, Kodak, & Fisher, 2013; Grow et al., 2009; St. Peter Pipkin, Vollmer, & Sloman, 2010; Wood, Umbreit, Liaupsin, & Gresham, 2007; Wilder et al., 2006). Thus, developing efficacious interventions for children with developmental disabilities is only the first step; we also need to ensure that staff are implementing the intervention with high fidelity.

There are several potential reasons why novice staff may have difficulty implementing an intervention as intended. First, entry-level staff may have minimal training and limited experience. The pitfalls surrounding minimal formal training are further compounded by the fact that children who engage in challenging behaviors may
be more difficult to teach (Symes, Remington, Brown, & Hastings, 2006). Without oversight and feedback regarding treatment fidelity, staff may modify interventions to mitigate the occurrence of problem behaviors in the moment. Moreover, in the presence of a challenging situation, staff may rely on their personal learning history, trial and error, and on-the-job learning experiences that have reduced the momentary occurrence of problem behavior in the past. However, such strategies may actually serve to maintain or exacerbate the problem behavior over time. Unfortunately, researchers have found that behavioral interventions in the human service field are often implemented with varying degrees of procedural integrity (Stahmer et al., 2015).

Second, if staff are not trained adequately (e.g., training in artificial settings that do not conform to the demands of the treatment environments) this may further intensify problems with fidelity. Regrettably, however, a high demand for applied behavior analysis (ABA)-services coupled with high staff turnover rates (Kazemi, Shapiro, & Kavner, 2015) and budgetary limitations may preclude companies from offering extensive supervised training (Jacobson & Mulick, 2000). Given that the efficacy of behavioral interventions hinges on a clinician’s ability to implement the intended intervention accurately, identifying effective methods to train staff to implement, generalize, and maintain newly acquired skills is paramount.

**Current Training Trends**

Often behavior analysts are responsible for training direct staff (e.g., students or teachers) to implement behavior change plans (Macurik, O’Kane, Malanga, & Reid,
Although it is unclear what the key elements of effective training are, approaches that have been effective in establishing and maintaining high procedural fidelity often include some form of verbal instructions coupled with modeling, rehearsal, and feedback. For example, Lavie and Sturmey (2002) used a multiple baseline design across participants to investigate the efficacy of behavioral skills training (BST) to teach three novice staff to conduct a paired-stimulus (PS) preference assessment. Sarkoff and Sturmey (2004) used a similar procedure to teach special education teachers to implement discrete-trial-teaching (DTT) sessions. BST consists of four main components (instructions, modeling, rehearsal, and feedback) most often used in tandem during staff training. Although there are procedural variations across studies in the application of BST, the efficacy of the training protocol has been well established (e.g., Garland, Vasquez, & Pearl, 2012; Iwata et al., 2000).

Despite the effectiveness of these training packages, a valid concern with the training technology is the length of time required to train all pertinent staff working at a company or school (Fleming, Oliver, & Bolton, 1996; Phillips, 1998). For example, in a study conducted by Lavie and Sturmey (2002), the duration of training was 80 min. Moreover, the use of in-vivo modeling or feedback (e.g., Hogan, Knez, & Kahng, 2015; Robinson, 2011) required a professional trainer be present for at least a portion of the training sessions even though many staff do not have access to an onsite trainer (Graff & Karsten, 2014). To combat these concerns, researchers have examined the acquisition of skills through the use of a self-instructional training package (Fazzio & Martin, 2011; Graff & Karsten, 2012; Severtson & Carr, 2012; Thiessen et al., 2009; Thompson et al., 2008).
Overall, results from these studies are promising and represent an important step forward in finding cost-effective methods of training that can be used to reduce the reliance on an on-site supervisor.

**Previous Reviews on Training**

There is a growing interest in exploring and characterizing key components of training packages aimed at teaching staff implementation of behavioral interventions. Although the following articles differ in scope from the present review, they provide preliminary analyses about the training strategies used with individuals who provide direct services to children with developmental disabilities.

One of the main goals in Brock and Carter (2013) was to identify professional development (i.e., training) strategies that facilitated paraprofessionals’ accurate use of educational practices. In a review of 13 studies, the authors found that researchers most often used multi-faceted training programs that incorporated modeling and corrective feedback. Researchers, unfortunately, collected follow-up data (i.e., maintenance of skills) in only half of the studies (7/13 or 53.8%); however, results did indicate that paraprofessionals’ performance maintained at mastery levels when training was discontinued. The authors also found that training was most often provided by an individual who was a member of the research team, and the duration of training (reported in 8/13 or 61.5%) ranged from 20 min to 6 hr. Finally, the authors analyzed the data to determine if training resulted in improved implementation fidelity and found an experimental effect in 29 of 35 studies (82.9%).
In a similar review, Snyder et al. (2012) used a framework from the National Professional Development Center on Inclusion (2008) to conduct a systematic review of the empirical early childhood professional development (PD) literature. The authors used the codes of “who, what, and how” to determine (a) the setting of the intervention (e.g., early childhood education, Head Start), the target consumers (e.g., children at risk or with disabilities); (b) the content of the PD implemented (e.g., social-emotional, literacy); and (c) the frequently occurring categories of PD (e.g., coaching, staff development). The authors found that a majority of the participants reported working with children with disabilities, most interventions centered on social-emotional or pre-academic content, and coaching/performance feedback was the most frequently mentioned strategy.

Rispoli et al. (2011) conducted a review of 12 studies aimed at summarizing methods to teach paraprofessionals how to implement interventions with children with autism. The authors found that performance feedback was provided in 9 of 12 studies (75%). However, in only two of nine studies did the researchers provide feedback in isolation (i.e., independent of a larger training package). Similarly, the authors found that researchers included vocal or written instructions during the intervention in 9 of 13 studies (75%), but in only five of nine (55.6%) did they provide a rationale for implementing the target intervention. The authors also collected generalization and follow-up data. Similar to the findings of Brock and Carter (2013), researchers included information on generalization across behaviors, settings, or consumers in less than half of the studies (3/12 or 25%; in two of three skills generalized), and in four studies (33.3%) researchers assessed maintenance of skills with positive results. Researchers also
provided information on the duration of training in only 3 of 12 studies (25%), which ranged from 20 min to 8 hr. Finally, the authors found that seven of the studies demonstrated positive outcomes (i.e., all paraprofessionals achieved mastery); the remaining five studies produced mixed outcomes in that positive gains were not apparent for all paraprofessionals. Although some of the authors (Brock & Carter, 2013; Rispoli et al., 2011) in the aforementioned reviews provided information regarding the number of studies in which an experimental effect was evident, none of the reviewers specified the exact training elements that resulted in positive or negative outcome. Most training components were part of a larger training package; therefore, the efficacy of each component was not evaluated in isolation. However, there is literature to support that the most effective components include feedback, and to a lesser extent, modeling (Ward-Horner & Sturmey, 2012).

**Purpose Statement**

Results from the literature review indicated that researchers have focused primarily on the efficacy of training strategies that included a combination of antecedent-based (e.g., instructions) and consequence-based (e.g., feedback) approaches facilitated by a professional with expertise in staff training. Despite advances in the effectiveness of expert-facilitated training procedures, the number of highly qualified trainers pales in comparison to the number of staff who are in need of competency-based training in order to serve individuals with developmental disabilities. Recent data indicate that 6.4 million children are receiving special education services, yet there are only 430,000 special education teachers working in school settings (National Center for Education Statistics,
2014) and only 22,000 certified behavior analysts in the United States (Behavior Analyst Certification Board, 2016). To the author’s knowledge, data regarding the number of behavior analysts working in school settings are not available; however, given the aforementioned statics regarding special education teachers, the number must be considerably less than the number of special education teachers who could benefit from evidence-based training procedures.

Although behavior analysts are not always responsible for directly working with teachers or other staff, they are in charge of training the numerous college students who may eventually work with staff (e.g., special education teachers) who may be in need of training. Special education teachers may also be responsible for training other staff; however, many report that they need further training because university-based training programs do not adequately prepare them to implement classroom management strategies and behavioral interventions (Stough, Montague, Landmark, & Williams-Diehm, 2015). Given the limited time and resources to train the consistent influx of newly hired staff or the students in charge of training professionals working in-home or school settings, human service organizations and universities may benefit from self-instructional packages that may reduce the reliance on expert staff trainers.

In order to further explore this line of research, this paper consists of a literature review of current training trends in behavior analysis (Chapter 2), a study exploring the effects of self-instructional modules to teach staff to implement a paired-stimulus preference assessment and least-to-most prompting procedures (Chapter 3), the outcomes
of the proposed study (Chapter 4), and a general discussion summarizing the findings and describing future research directions (Chapter 5).
Chapter 2: Literature Review

Despite the wide-spread contributions of different training approaches, to the author’s knowledge, there are only a handful of studies summarizing this body of literature (see Brock & Carter, 2013; Rispoli et al., 2011; Snyder et al., 2012). Although the existing reviews provide valuable information regarding the characteristics of effective training programs, in order to fill in the gap in the literature, a broader analysis of current trends in training is needed. Therefore, in this review, my main objectives are to (a) outline participant characteristics and setting; (b) outline the behavior analytic procedures targeted for training; (c) outline the most common training components and medium of delivery; (d) outline the time required to train all relevant staff; (e) describe training outcomes; (f) determine the generality of the outcomes; and (g) describe the degree to which the training was acceptable by staff.

Literature Review Methodology

Relevant studies were identified through a systematic search of Eric, Education Research, and PsychINFO. The terms staff or teacher or paraprofessional train* or staff or teacher or paraprofessional coach*, were combined with behavior* intervention or applied behavior analytic service. This initial search generated a total 3,003 articles after the duplicates were removed. Articles were then excluded based on their titles ($n = 2,751$), which yielded a total of 252 articles. The abstracts of the articles ($n = 252$) that
appeared to meet our predetermined inclusionary and exclusionary were examined further for eligibility.

Articles were included in the present review if the following criteria were met: (a) they were published in a peer-reviewed journal; (b) the study was conducted in the United States; (c) the aim was to train college students or staff working with school-aged children with developmental disabilities; (d) the authors indicated in the method section that all participants had no prior training implementing the target intervention; (e) the specific training strategies were reported; (f) the data were reported at the level of staff performance; and (g) the publication year was between 2000 and 2015. The latter criterion was used to reflect the most recent trends in the literature. Excluded articles were those that met the following criteria: (a) the training was focused on children or parents; (b) review articles; (c) dissertations; (d) qualitative studies focused on staff’s beliefs; and (e) articles that employed group or single-subject designs (e.g., AB designs) that were inadequate in demonstrating experimental control. Group design studies were excluded because of the difficulties associated with comparing group to individual outcomes, whereas weak single-subject designs were excluded because of the lack of conclusive evidence of an experimental effect.

I conducted forward and backward searches to identify additional training articles (n = 5). Finally, I developed a list of journals that included published articles that met the inclusionary criteria, and then identified articles between 2000 and 2015 through a hand search of Behavior Analysis in Practice, Behavior Modification, Behavioral Interventions, Education and Training in Autism and Developmental Disabilities, Focus.
Coding and Reliability

There was substantial variation along several dimensions (e.g., participant characteristics, training components, duration) among the studies that met the inclusionary criteria. We then categorized, quantified, and critically evaluated each relevant study along the following dimensions.

Participants and Setting

Participants. Data were collected on participant age, education level and background, and position. We categorized participants’ highest level of education as a high school diploma, some college experience, a bachelor’s degree, or a master’s degree. Educational background was categorized as psychology, a field relevant to education (e.g., special or elementary education, speech and language services, communication and deaf education, school psychology), or an unrelated field (e.g., English history, biology, linguistics). Participants’ positions were categorized as instructional assistants or trainees (e.g., paraprofessionals), teachers (e.g., special or general education), behavior technicians (i.e., ABA tutors), or students. The latter category allowed me to review training programs for novice students who are representative of direct staff who may be expected to implement behavior change programs.
**Setting.** The settings in which the training program occurred was coded. I categorized settings as a child’s classroom, a conference room or private office, a clinic, or a university campus.

**Interventions Targeted for Training**

Data were collected on the different types of interventions targeted for training (e.g., token economy systems, discrete trial teaching).

**Training Components and Medium of Delivery**

**Training component.** Data were collected on the different types of training components used to teach individuals how to implement a variety of procedures. Components were categorized along the following domains: (a) instructions (i.e., clear, concise verbal descriptions of the procedures); (b) modeling (i.e., demonstration of how a procedure should be implemented); (c) rehearsal (i.e., practice of the target skills until mastery is achieved); and (d) feedback (i.e., verbal praise or corrective information specifying the accuracy of the target skills).

**Medium of delivery.** The four aforementioned domains were then further categorized based on the medium in which the training component was delivered.

**Instructions.** Instructions included vocal or written instructions typically given during baseline or initial training phases. Examples of this approach included providing a group of participants with a description and the purpose of the assessment (Wallace, Mintz-Resudek, & Tarbox, 2004), providing a single participant with a brief summary
outlining the key points of the intervention (Roscoe et al., 2006), or providing staff with a copy of a training manual coupled with a brief quiz (Severtson & Carr, 2012).

**Modeling.** I categorized modeling as video or with the experimenter. Examples of this approach included the experimenter modeling correct and incorrect procedures (Hogan, Knez, & Kahng, 2015; Garland et al., 2012), or video modeling. Video modeling was further categorized into videos with (Rosales, Gongola, & Homlitas, 2015; Weldy, Rapp, & Capocasa, 2014) or without (Iwata et al., 2000) embedded instructions.

**Rehearsal.** Rehearsal was categorized as with a role-played client (Graff & Karsten, 2012; Moore & Fisher, 2007) experimenter or in-vivo with the actual client (Petscher & Bailey, 2006). In line with our operational definition, rehearsal was classified as occurring when participants were practicing with a role-played client or a client during baseline or intervention phases. Participants were not considered to be “practicing” if they had already achieved mastery in a previous condition or the intervention had been terminated.

**Feedback.** Initially, feedback was categorized into vocal, written, or video. I then further classified the dimensions of feedback along temporal domains (i.e., before the session, during the session, or immediately after the session). Examples of this approach included the experimenter providing positive and corrective feedback during role-play (Nigro-Bruzzi & Sturmey, 2010), or the experimenter providing video feedback based on the trainee’s performance following the session (Robinson, 2011).
Duration of Training

Data on the total time required to complete the training were collected.

Training Outcomes (i.e., Effectiveness)

A classification system similar to that of Alvero, Bucklin, and Austin (2000) and Balcazar, Hopkins, and Suarez (1985) was used to categorize training effectiveness. Minor revisions were made that included slight alterations to the operational definitions per category: “consistent effects” was changed to “positive outcomes.” The data were then visually inspected to determine if a meaningful change in behavior occurred and the extent to which the change could be attributed to the manipulation of the independent variable. Variability, level, and trend were also taken into account (Cooper, Heron, & Heward, 2007).

Positive outcomes. The outcomes were categorize as positive when the intervention produced unilateral, desired changes in performance (participants achieved mastery and experimental control was demonstrated). These gains must have occurred across all settings, participants, and behaviors being evaluated.

Mixed outcomes. The outcomes from the training program were categorized as mixed when the intervention produced meaningful (i.e., participants achieved mastery and experimental control was demonstrated) changes for some, but not all, participants.
No effects. The outcomes from the training program were categorized as having no effects when all participants failed to achieve mastery during the intervention phase, and experimental control was weak.

Generality

Maintenance. Data were collected on the use of maintenance probes and categorized as having occurred one week, one month, or more than one month following the termination of the intervention.

Generalization. Data were collected on the use of generalization probes and categorized as across consumers, settings, and/or skills following the termination of the intervention.

Social Validity

Data were collected on social validity measures using the framework summarized succinctly by Wolf (1978). Specifically, we looked at (a) the goals of the procedures to determine if the participants indicated that they were important and relevant to all stakeholders; (b) the training techniques to determine if they were acceptable and viable when implementing in a school or community setting; and (c) the level of satisfaction with the training outcomes.

Interobserver Agreement

I made the initial determination as to whether a study met the inclusionary criteria; a second trained individual repeated the systematic search process by applying the same criteria. If a disagreement occurred, the inclusionary and exclusionary criteria were discussed until there was a consensus, and the study was then reviewed again. Once
an agreement between coders was reached, all identified articles were then divided and coded across seven categories: (1) participant and setting characteristics; (2) target skill; (3) baseline and intervention training components and medium of delivery; (4) training duration; (5) training outcomes; (6) generality; and (7) social validity.

A second reader independently analyzed 100% of the articles to evaluate the reliability of the findings. A score sheet was developed by the first author based on the aforementioned categorizations and agreement was then assessed by an item-by-item comparison. Agreement was calculated separately for each of the categories by dividing the number of agreements between the two raters by the number of agreements and disagreements and multiplying by 100. Mean agreement across all categories was 100%

Results and Discussion

In the end, the search generated a total of 24 articles between the years of 2000 to 2015. A total of 12 journals published at least one training study; however, a majority of the articles were published in JABA (n = 12 or 50%). Although many articles published in various journals appeared to meet our initial inclusionary criteria (n = 84), upon further review of the method section, it became apparent that authors did not specify whether staff had previous experience with the target intervention.

Participants and Settings

Participants. A total of 120 adult participants were included in the studies that met our criteria; of these, 37 (30.8%) were classified as students, and 29 (24.2%) were classified as instructional assistants (i.e., paraprofessionals). The remaining 54
participants (45%) we classified as special or general education teachers (29 or 24.2%),
or behavior technicians (25 or 20.8%). When information was available (for 75/120
individuals or 62.5%), participants’ age ranged from 19 years (Pollard, Higbee, Akers, &
Brodhead, 2014) to 60 years (Robinson, 2011), and a majority (69/75 or 92%) were female. For the remaining 45 participants (37.5%), identifying information was not given.

Of the 17 studies (71%) reporting participants’ education level, 44 of 81
participants (54.3%) earned a bachelor’s degree, 19 (23.5%) had some undergraduate
experience, 11 (13.6%) earned a high school diploma, four (4.9%) had some master’s
experience, and three (3.7%) earned a master’s degree. Educational background was not
always reported for all participants who held a degree (for 52/109 individuals or 47.7%),
but a majority were working on or earned a bachelor’s degree in psychology (31/52 or
59.6%). Approximately 17 participants (32.7%) earned their degree in a field relevant to
education, whereas four participants (7.7%) earned a degree in an unrelated field. Some
of the participants had attended seminars or workshops on the basic principles of
behavior analysis (e.g., Feldman & Matos, 2013; Ward-Horner & Sturmey, 2012) and
some were students currently enrolled in ABA coursework (e.g., Iwata et al., 2000). It
should be noted that limited information was provided on educational background. This
is important information to include because it is possible that training may not be
adequate for individuals with limited to no educational and clinical experiences in this
field. Conversely, different training strategies may be needed for individuals who have
been working in the field for an extensive amount of time, but who may also have a long
history of implementing a specific procedure incorrectly. Such factors may influence the
ease by which individuals acquire the target skills and subsequent fidelity of implementation (Rispoli et al., 2011; Weldy et al., 2014). Moreover, provision of this data can aide in the interpretation and generalization of results (Rispoli et al., 2011; Weldy et al., 2014).

**Setting.** The settings were diverse, but most training interventions were conducted in the child’s classroom (12/24 or 50%). Other settings included a conference room or private office (5 or 20.8%), a clinic (4 or 16.7%), or on a university campus (3 or 12.5%).

**The Interventions Targeted for Training**

Across the 24 studies, individuals were trained in DTT (9/24 or 37.5%), SPAs (5 or 20.8%), FAs (4 or 16.7%), mand training (2 or 8.3%), pivotal response training (2 or 8.3%), aspects of a behavior intervention plan (1 or 4.2%), and token economy systems (1 or 4.2%).

**Training Components and Medium of Delivery**

**Baseline.** All researchers reported that participants had no experience conducting the targeted intervention prior to onset of training. In one study (4.2%), preceding the collection of baseline data, the experimenter held an 8 hr workshop during which students were provided with information regarding how to implement DTT (Downs, Downs, & Rau, 2008). Specifically, the training consisted of instructions, in-vivo modeling by the experimenter, and practice with corrective feedback. Although the training led to fairly good outcomes, none of the participants achieved mastery with the workshop alone.
Individual coaching sessions during which the experimenter provided vocal and written feedback were needed for all participants to implement DTT with high fidelity. In 6 of 24 studies (25%), baseline measures included asking participants to engage in their normal activities or to implement an unknown intervention (Feldman & Matos, 2013; McCulloch & Noonan, 2013; Petscher & Bailey, 2006; Robinson, 2011; Roscoe et al., 2006; Weldy et al., 2014) in order to approximate an all too likely situation where staff are asked to conduct behavioral interventions or assessments without formal instruction, assistance, or feedback.

In the remaining 17 studies (70.8%), participants were asked to implement the target intervention (e.g., mand training, DTT, stimulus preference assessments) after reading a written protocol of the procedures (e.g., Gilligan, Luiselli, & Pace, 2007; Graff & Karsten, 2012; Nigro-Bruzzi & Sturmey, 2010). In 2 of 17 studies (11.8%), written instructions given during baseline were facilitated by the experimenter (i.e., read aloud or reviewed as the participants asked questions; Gilligan et al., 2007; Hogan et al., 2015). For example, Nigro-Bruzzi and Sturmey (2010) gave participants written instructions that included the definition of a mand, how to conduct the training, and how to collect data on child performance. The experimenter then read the contents aloud and concluded by asking the participants if they had any questions. In the remaining 15 of 17 studies (88.2%) that included instructions at baseline, reviewing the material occurred independent of the experimenter (e.g., Moore & Fisher, 2007; Rosales et al., 2015; Roscoe & Fisher, 2008). Rosales et al. (2015) gave teachers the relevant sections from peer-reviewed articles outlining implementation of a stimulus preference assessment
(SPA), as well as a list of the correct target responses. If teachers asked questions, they were informed that their questions would be answered at a later time and to refer to their written instructions.

Overall, results from instruction-based methods during baseline simulations produced variable results across participants and conditions. However, none of the participants achieved mastery with written or vocal instructions alone, which is in line with the findings that instruction-based methods alone typically do not produce the desired effect (e.g., Iwata et al., 2000; Roscoe et al., 2006; Vlădescu et al., 2012). There was one exception: Nigro-Bruzzi and Sturmey (2010) noted that during mand training, one participant showed a steady increasing trend during baseline assessments; therefore, the authors did not implement training with this participant.

Although no studies included modeling during baseline assessments, in 16 of 24 studies (66.7%) rehearsal during baseline occurred with a role-played client. Severtson and Carr (2012) asked novice instructors to conduct DTT sessions with a role-played client during baseline. Instructors were given a data sheet to record the client’s responses and 10 min to read a one page description of a discrimination training program. Participants were then asked to conduct DTT with the client and received no feedback on their performance. Conversely, in 8 of 24 studies (33.3%) participants were asked to implement the intervention with a child with a developmental disability. Hogan et al. (2015) taught staff how to implement components of two students’ behavior intervention plans (BIP). During baseline assessment, staff had 10 min to review a written copy of the student’s BIP, and the experimenter reviewed each component. Staff were then asked to
“implement the BIP to the best of their ability” (p. 248). Staff were not allowed to ask questions. The only feedback given on their performance was a mark of incorrect on their data sheet that corresponded to the missed component. Finally, only one study used feedback during baseline assessments. In Downs et al. (2008), the authors asked individuals to observe each instructor throughout their daily work shift and record the use of DTT as satisfactory or unsatisfactory across all component skills. Although the authors provided general positive feedback (e.g., “thanks for working so hard”) and answered any outstanding questions, no additional specific information on individual performance was given.

**Intervention.** The schedule of training across a majority of the studies was fairly consistent in that, initially, written material was given to the participants and reviewed by the experimenter. Instructions were typically followed by practice with a role-played or actual client coupled with approving (e.g., “Good job”) and corrective (i.e., information on the accuracy of target skills) feedback by the experimenter contingent on performance. In all except three studies (Iwata et al. 2000; Wallace et al., 2004; Weldy et al. 2014), training sessions appeared to be delivered in a one-to-one context (21/24 or 87.5%). Across all 24 studies, an inter-observer agreement of at least 80% for at least 20% of all sessions was reported, which meets minimum standards (Kennedy, 2005). However, treatment fidelity of the independent variable was assessed in only 10 of 24 studies (41.7%) and was reported to be 90% or greater for at least 30% of the sessions.

**Instructions.** Overall, instructions during training mirrored those given during baseline assessments except the process was most often facilitated by the experimenter.
Although the exact nature of the instructions was not always clear, all written and/or vocal instructions included detailed descriptions of the target intervention that highlighted key components. An important point to consider when disseminating vocal and written instructions is the content of the material. It was difficult to discern if researchers provided a rationale for implementing the target intervention. However, in order to gain “buy-in” or support to adopt and implement an intervention with high fidelity, it may be important to provide participants with this information (Turnbull, 2002).

Instructions were frequently used during initial training sessions (19/24 or 79.2%). In 9 of 19 studies (47.4%) that used instructions during training, the authors administered a written test before or during the intervention phase to ensure that participants had knowledge of the procedures prior to implementing the target intervention with a client. Salem et al. (2009) asked students to review a 37-page self-instructional manual for teaching DTT. The manual included, among other training components, discussion questions and mastery tests. For each mastery test, participants were required to answer at least 50% of the study guide questions randomly selected by the authors. A score of below 50% resulted in the participants reviewing the sections of the manual that corresponded to the incorrect questions. Participants were then asked to retake the test, but were required to answer only the questions that they had previously missed.

In 4 of 19 studies (21.1%) the authors used primarily instruction-based methods alone resulting in various outcomes. Graff and Karsten (2012) provided teachers with a self-instructional manual to teach implementation of SPAs. The manual included step-by-
step instructions (with limited technical language), pictures, diagrams, and a comprehensive data sheet. By the end of training, all 11 teachers met the mastery criteria for implementation, scoring, and interpreting outcomes with a role-played client. The skills also generalized to a client with whom they were working. In the other three studies, the authors taught students or staff to conduct DTT sessions using a self-instructional package, which included discussion questions, mastery tests, a practice component, and a brief 17 min video (Salem et al., 2009; Thiessen et al., 2009; Thomson et al., 2012). Thomson et al. (2012) found that only two of eight tutors achieved mastery after reviewing the manual alone; two additional tutors achieved mastery after viewing the 17 min video. Unfortunately the four remaining tutors were required to begin work at the intervention facility and additional training could not occur, even though they had failed to achieve mastery on at least one target skill.

The discrepancies regarding the findings of written self-instructional packages may be due to the complexity of the trained target skills (i.e., DTT training consisted of approximately 21 target skills, whereas SPAs consisted of approximately seven to nine). There were also pronounced differences (e.g., number of pages, the inclusion of self-practice activities, study guide questions, and mastery tests) between the packages used for each target intervention. Results from these studies give us insight into what is needed for written self-instructional packages to be effective; key components include instructions coupled with a brief video model, or instructions that contain pictures and diagrams; discussion questions and mastery tests; and potentially, self-practice activities.
In the remaining 15 of 19 studies (78.9%), modeling and feedback to increase implementation accuracy were frequently used following the initial training sessions with instructions (10/19 or 52.6%), followed by feedback alone (3 or 15.8%), and modeling alone (2 or 10.5%). For example, Wallace et al. (2004) trained educators to implement a functional analysis (FA). At intervention, staff attended a workshop during which they were initially given the description and purpose of each condition. Next, staff watched a videotaped demonstration of each condition and role-played how to implement them with one another. If staff failed to conduct the assessment at 90% or greater correct trials, contingent feedback was given immediately after the session. Only one of three participants in this study required the addition of specific feedback because of reoccurring errors in the demand condition.

**Modeling.** Modeling was another component frequently used across training sessions (14/24 or 58.3%). Modeling often took the form of a video model (11/14 or 78.6%) with (6/11 or 54.5%) or without (5/11 or 45.5%) a voice over or in-text script. Weldy et al. (2014) trained behavioral staff to conduct two types of SPAs. During the video training phase, all staff, in a group format, viewed a PowerPoint presentation that included detailed information about the procedures and a corresponding video model outlining each step of the task analysis for implementation. After each step had been reviewed separately, the authors concluded the session by showing three sessions of each assessment in its entirety. The authors found that all participants met the mastery criterion for implementation, and these results were replicated by Rosales et al. (2015).
In a similar study focusing on FAs, Moore and Fisher (2007) compared the relative efficacy of instructions and complete or partial video modeling. The two videos differed across the range and number of therapist behaviors depicted. The authors found that compared to a lecture or a video that contained only 50% of potential therapist target behaviors, a video model containing multiple examples of each potential behavior was more effective at teaching staff how to implement the different conditions in the FA. Thus, multiple exemplars spanning the full range of behaviors a therapist will be required to emit (a complete video model) appears to be a more effective medium than a partial video model. In the remaining 3 of 14 studies (21.4%), modeling occurred in-vivo and usually took the form of the experimenter modeling correct implementation during feedback and rehearsal sessions (Garland et al., 2012; Hogan et al., 2015; Robinson, 2011).

In the 14 of 24 (58.3%) studies that incorporated modeling into their training package, 10 of 14 (71.4%) consisted of modeling, instructions and feedback. Some idiosyncrasies were present across studies. For example, in three studies the authors delivered a brief session of feedback for only one participant in each study because he or she failed to achieve mastery (Moore & Fisher, 2007; Pollard et al., 2014; Rosales et al., 2015). Thus, feedback was used as an ancillary component only if participants did not achieve mastery levels of performance following video or video module-based training. In the remaining four studies (23.1%), modeling occurred either with instructions (Vladescu et al., 2012; Weldy et al., 2013), or in conjunction with feedback (Robinson, 2011; Ward-Horner & Sturmey, 2012). In the latter study, the authors conducted a
component analysis of BST to evaluate the individual effects of all components when teaching direct-care staff how to implement an FA. The authors found that feedback, and to a lesser extent, modeling were the most effective components of training, whereas instruction and rehearsal were ineffective. This finding is consistent with other researchers who have noted that written protocols and rehearsal alone do not always result in participants achieving mastery (e.g., Severtson & Carr, 2012; Thomson et al., 2012). Most often in training reviews, including this one, multiple training procedures were used in combination and it was impossible to identify the effects of individual training components on staff performance. Therefore, Ward-Horner and Sturmey (2010) recommend the need for conducting component analyses to help identify and eliminate unnecessary procedures.

Overall, results from the aforementioned studies indicated that one potentially cost-effective approach to training could be to use video modeling procedures initially in a group format and include a brief session of feedback for participants who did not learn accurate implementation of the target intervention with the video alone. As discussed in Ward-Horner and Sturmey (2010), even if half the participants achieve mastery with the video model alone, this training model would greatly reduce the need of an onsite trainer or supervisor and could cut the time required to train newly hired employees substantially.

**Rehearsal.** In 12 of 24 studies (50%) the experimenter or members of the research team served as the role-played client during rehearsal (in 1/12 participants rehearsed with the experimenter and one another; Wallace et al., 2004). In five studies (20.8%) rehearsal
occurred exclusively with clients (i.e., children with developmental disabilities), whereas
in three studies (12.5%), participants rehearsed with both a role-played client and actual
client (Garland et al., 2012; Gilligan et al., 2007; Hogan et al., 2015). For example,
Garland et al. (2012) used a TeachLivETM platform (i.e., a virtual classroom laboratory)
to train graduate students to implement DTT. During the intervention phase, the
experimenter analyzed students’ performance in the “virtual classroom.” For incorrect
responses, the experimenter reviewed the operational definition of each step, modeled
correct implementation, and provided feedback as the students rehearsed. In the
remaining 4 of 24 studies (16.7%), aside from rehearsing with a role-played client, the
authors asked participants to engage in self-practice activities. Specifically, participants
were asked to imagine that they were teaching a child with autism and to use the DTT
protocol to rate their implementation (Pollard et al., 2014; Salem et al., 2009; Thiessen et
al., 2009; Thomson et al., 2012).

The use of role-played clients during rehearsal has both its strengths and
limitations. First, role-playing avoids exposing children to pre-training mistakes (e.g.,
accidentally reinforcing or evoking problem behaviors, creating prompt dependency), and
people may find it unethical to work with children directly without training (Garland et
al., 2012). Second, a confederate playing the role of a child ensures that each participant
undergoes a similar training experience because behaviors targeted for change are
typically included in simulated scripts so as to guarantee the opportunity to practice the
target skills. Third, participants may feel more at ease making errors with a role-played
rather than actual client. Drawbacks, however, include potentially low generalization
from simulated experiences to working with children with developmental disabilities in a classroom or residential setting, and the absence of idiosyncratic behaviors on the part of the child that have not been accounted for in the simulations. Therefore, it is important to incorporate various environmental elements that will be found in the intervention setting to facilitate generalization of the target skills.

**Feedback.** Some form of feedback was used in 18 of 24 studies (75%). Most often, feedback was given by the experimenter after the training session (7/18 or 38.9%), followed by immediately prior to training based on the participants’ last performance (5 or 27.8%), and in-vivo as participants rehearsed with a client (4 or 22.2%). In two studies (11.1%), the authors either delivered feedback in-vivo and prior to the next session (Severtson & Carr, 2012), or in-vivo and after the session (Downs et al., 2008). Reid, O’Kane, and Macurik (2011) suggested that traditional methods of giving feedback (i.e., after the completion of the session) may require an increased amount of response effort on the part of the trainer or supervisor. Alternatively, the use immediate feedback during observation sessions may increase the frequency or precision with which feedback is delivered and can help the trainee identify and analyze incorrect responses in the moment instead of waiting until the end of the session (Reid et al., 2011). Although immediate feedback is considered to be best practice,

Feedback often took the form of vocal alone (7/18 or 38.9%), or vocal with written feedback (5 or 27.8%). Hogan et al. (2015) gave participants a copy of their graphs from the previous assessments and reviewed each component score. At the end of the observation period, the authors provided vocal feedback (although the exact content
of that feedback was not specified) and modeled, if necessary, the correct procedures. Some experimenters used a scoring sheet for a particular intervention as a guide to provide feedback (e.g., Pollard et al., 2014), whereas others provided vocal corrective and reinforcing feedback coupled with rating scales (i.e., satisfactory, unsatisfactory) at the end of the work shift (Downs et al., 2008).

In 4 of 18 studies (22.2%), feedback consisted of video and vocal (Iwata et al., 2000; Robinson, 2011) or video, vocal, and written forms (Roscoe et al., 2006; 2008). Prior to each session, Roscoe et al. (2008) gave participants a copy of their data sheet from the preceding session and reviewed a video of their performance. The experimenter then provided feedback on the target skills by indicating whether a skill depicted on the video and data sheet was implemented correctly. In the remaining two studies (11.1%), feedback took the form of self-monitoring alone (McCulloch & Noonan, 2013), or the authors used a pager to alert participants of a missed opportunity to engage in the target behavior and provided feedback on participants’ self-monitoring sheets (Petscher & Bailey, 2006). In the 18 of 24 studies (75%) that incorporated feedback into their training packages, 10 of 18 (55.6%) consisted of instructions, modeling, and feedback; three (16.7%) included feedback alone; three (16.7%) included instruction and feedback; and two (11%) included modeling and feedback. Previous researchers have found that feedback is one of the most effective methods to teach individuals accurate implementation of a variety of behavioral strategies (Alvero et al., 2000; O’Reilly & Renzaglia, 1994). However, the authors evaluated the effects of feedback in isolation in only four studies; most often, feedback was a part of a comprehensive training package.
**Duration of Training and Outcomes**

The duration of training referred to the amount of time participants underwent training and, for the most part, was competency-based in that each trainee met a predetermined criteria before training was considered complete (Parsons, Rollyson, & Reid, 2012). In some studies, however, duration was provided, but all participants did not achieve mastery once training concluded (e.g., McCulloch & Noonan, 2013; Salem et al., 2009). The duration of training varied substantially across studies and appeared to be related to the complexity of the trained skills. In eight of nine studies (88.9%) focused on implementation of DTT, information regarding the duration of training was provided and ranged from approximately 40 min (Gilligan et al., 2007) to 4.4 hr (Thomson et al., 2012). Positive outcomes were found in eight of nine studies (88.9%) and mixed in one of nine studies (Salem et al., 2009; 11.1%). In the two studies where the outcomes were positive (Severtson & Carr, 2012, Thiessen et al., 2009), the authors evaluated the efficacy of a self-instructional package; no performance feedback or video modeling were given initially for incorrect responding. Although all participants did eventually achieve mastery after viewing a video or the delivery of feedback, performance gains were not always maintained during generalization or maintenance probes. In three of five studies (60%) focused on implementation of SPAs, information regarding the duration of training was provided and ranged from approximately 30 min (e.g., Graff & Karsten) to 60 min (Weldy et al., 2014) for each assessment. Positive outcomes were found in all five studies. In three of four studies (75%), the authors provided information regarding the
duration of training for FAs, which ranged from approximately 15 min (Moore & Fisher, 2007) to 2 hr (Iwata et al., 2000). Positive outcomes were found in all studies.

The duration of mand and pivotal response training (PRT) was included in all four studies and was similar for both ranging from 2 hr (Nigro-Bruzzi & Sturmey, 2010) to 5 hr (McCulloch & Noonan, 2013) for mand training, and 2 hr (Feldman & Matos, 2013) to 6 hr (Robinson, 2011) for PRT. Positive outcomes were found in three of four (75%) of the studies (Feldman & Matos, 2013; Nigro-Bruzzi & Sturmey, 2010; Robinson, 2011). Similar to DTT training, although all participants achieved mastery during training, performance gains were not always maintained during generalization and maintenance probes (Robinson, 2011). Finally, information regarding the duration of training was provided in the article focused on teaching staff to implement aspects of a behavior plan (45 min to 2.5 hr; Hogan et al., 2015); however, no information was provided for the article focused on implantation of a token economy system. Across both studies, positive outcomes were noted.

Generality

Of the 24 studies included in the review, nine studies (37.5%) conducted both maintenance and generalization probes, five studies (20.8%) conducted only generalization probes, and four studies (16.7%) conducted only maintenance probes. The remaining six studies (25%) did not assess the generality of the findings. Researchers collected maintenance data anywhere between 1 week (Graff and Karsten) to 12 weeks (Wallace et al., 2004) after the termination of the intervention. In 5 of 13 studies (38.5%)
in which maintenance data were collected, at least one participant did not maintain
criterion levels of performance (McCulloch & Noonan, 2013; Petscher & Bailey, 2006;
Robinson, 2011; Rosales et al., 2015; Severtson & Carr, 2012). In regard to
generalization, probes were conducted in 14 of 24 studies (58.3%) most often across
children alone (5/14 or 35.7%), across settings and children (3 or 20%), across skills
alone (3 or 20%), across skills and children (2 or 14.3%), or across settings alone (1 or
7.1%). In 8 out of 14 studies (57.1%) in which generalization data were collected, similar
to the maintenance phase, at least one participant did not maintain criterion levels of
performance.

Social Validity

The authors collected social validity data in 10 of 24 studies (41.7%) and
primarily used Likert-type scales to evaluate outcomes after the intervention had been
implemented. Overall, participants reported that the goals were important, and the
procedures and outcomes were effective. In studies that employed multiple training
strategies (e.g., instructions, modeling, and feedback), participants’ reports of the
components they believed to be most effective aligned with the components they liked
the most (e.g., Ward-Horner & Sturmey, 2012). Participants also reported that by the end
of training, they understood the importance of the target intervention and would continue
to use the procedures throughout the school year (e.g., Feldman & Matos, 2013). In
Feldman and Matos (2013), participants’ responses indicated a preference for in-vivo
training when compared to workshops that they had attended in the past because of the
opportunity to receive direct feedback specific to the child with whom they were working.
within the intervention setting. However, in Garland et al. (2012), participants reported that they were more comfortable working with a role-played client because they would not “damage” their client as they practiced a new strategy.

There were several problems with the existing methods of collecting data. First, social validity data were assessed in only 42% of the studies. The willingness of community members to continue to embrace and adopt a specific training method depends upon the continuous collection and analyses of social validity data. Therefore, as recommended by Schwartz and Baer (1991), the social validity of the goals, procedures, and anticipated effects should be known prior to the commencement of the intervention and assessed throughout. Second, current methods of collecting data were quite subjective, and the results may have been confounded by the fact that participants’ behavior may have been under the control of social contingencies provided by the experimenter (Schwartz & Baer, 1991). Similarly, the authors used rating scales that primarily measured participants’ verbal behavior, which is concerning given the difficulties surrounding accurate reports. Moreover, individuals’ self-reported “satisfaction” is relative. To some individuals, satisfaction may mean that they were extremely satisfied, whereas others may think that the training was just merely acceptable. Third, most Likert-type measures used have unknown psychometric properties (Carr, Austin, Britton, Kellum, & Bailey, 1999). Although existing measures of obtaining social validity data are a good first step, we now need to look at alternative direct measures (e.g., behavior correlates of satisfaction, consumers’ choice of an intervention following completion of two or more interventions, the continued use of the
intervention overtime) for collecting this vital information (Schwartz & Baer, 1991). See Table 1 for results for some of the main variables under investigation.
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Participants</th>
<th>Target Skill</th>
<th>Training Duration</th>
<th>Baseline Training</th>
<th>Intervention Training</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Downs et al. (2008)</td>
<td>6 undergrad. students btw. 20 to 22 yrs.; psych. majors</td>
<td>DTT</td>
<td>At least 8 hr</td>
<td>*T, M, R, FB</td>
<td>R: w/child FB: in-vivo vocal FB + written checklists &amp; ratings after sessions</td>
<td>Positive</td>
</tr>
<tr>
<td>Feldman &amp; Matos (2013)</td>
<td>3 paras. btw. 20 to 50 yrs.; one is a linguistics major; others earned a high school degree; employed at a public school supporting children with autism</td>
<td>PRT</td>
<td>2 hr. to 3 hr</td>
<td>Engaged in their normal activity</td>
<td>T: reviewed a manual R: w/child EB: in-vivo vocal</td>
<td>Positive</td>
</tr>
<tr>
<td>Garland et al. (2000)</td>
<td>4 grad. students btw. 23 to 54 yrs.; employed in K-12 public school systems</td>
<td>DTT</td>
<td>≈ 1.5 hr</td>
<td>I</td>
<td>T: list of operational def. M: model by ex. R: w/child + embedded in the FB condition w/ex. EB: written &amp; vocal instructional FB after sessions</td>
<td>Positive</td>
</tr>
<tr>
<td>Gilligan et al. (2007)</td>
<td>3 paras. btw. 21 to 23 yrs.; psych. &amp; elementary comm. majors; employed in K-12 public school system</td>
<td>DTT</td>
<td>≈ 40 min</td>
<td>I</td>
<td>R: w/child + embedded in the FB condition w/ex. FB: vocal FB preceding each session</td>
<td>Positive</td>
</tr>
</tbody>
</table>

Continued

Table 1. A List of all Articles Included in the Review and the Results for Some of the Main Variables Under Investigation
Table 1 continued

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Participants</th>
<th>Targeted Skill</th>
<th>Training Duration</th>
<th>Baseline Training</th>
<th>Intervention Training</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graff &amp; Karsten (2012)</td>
<td>11 teachers w/ a BA or MA degree; employed in a behavior-analytic school</td>
<td>SPA</td>
<td>≈ 30 min per assessment</td>
<td>I</td>
<td>I: written self-instructional manual with data sheet, pictures, and diagrams</td>
<td>Positive</td>
</tr>
<tr>
<td>Hogan et al. (2015)</td>
<td>4 teachers btw. 2.5 to 24 yrs.; employed at a nonpublic day school for children with autism and other developmental disabilities</td>
<td>BIP</td>
<td>45 min to 2.5 hr.</td>
<td>I</td>
<td>I: written copy of BIP + graph of last session, M: model by ex., R: w/ex., FB: vocal FB at end of session</td>
<td>Positive</td>
</tr>
<tr>
<td>Iwata et al. (2000)</td>
<td>11 undergrad. students in psych.</td>
<td>FA</td>
<td>2 hr.</td>
<td>I</td>
<td>I: outlines of conditions that were vocally reviewed by the ex.</td>
<td>Positive</td>
</tr>
<tr>
<td>McCulloch &amp; Noonan (2013)</td>
<td>3 para. btw. 26 to 46 yrs.; high school and BA degrees; employed at a public elementary school</td>
<td>Mand</td>
<td>3 hr. to 5 hr.</td>
<td>Engaged in their normal activity</td>
<td>I: text to accompany a video model, M: videos w/ voice over scripts</td>
<td>Mixed</td>
</tr>
</tbody>
</table>

Continued
Table 1 continued

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Participants</th>
<th>Targeted Skill</th>
<th>Training Duration</th>
<th>Baseline Training</th>
<th>Intervention Training</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moore &amp; Fisher (2007)</td>
<td>3 trainees with BA degrees in psych.</td>
<td>FA</td>
<td>≈ 15 min to 30 min</td>
<td>I</td>
<td>I: lecture training&lt;br&gt;&lt;br&gt;&lt;br&gt;M: partial &amp; complete video model&lt;br&gt;&lt;br&gt;&lt;br&gt;R: w/ex.&lt;br&gt;&lt;br&gt;&lt;br&gt;&lt;br&gt;&lt;br&gt;(FB): a single session of FB before next session</td>
<td>Positive</td>
</tr>
<tr>
<td>Nigro-Bruzzi &amp; Sturmey (2010)</td>
<td>6 teachers (i.e., 3 special education; 3 speech therapists)</td>
<td>Mand</td>
<td>≈ 3 hr.</td>
<td>I</td>
<td>I: written instructions reviewed by ex.&lt;br&gt;&lt;br&gt;&lt;br&gt;M: video model&lt;br&gt;&lt;br&gt;&lt;br&gt;R: w/ex.&lt;br&gt;&lt;br&gt;&lt;br&gt;F: in-vivo vocal FB</td>
<td>Positive</td>
</tr>
<tr>
<td>Petscher &amp; Bailey (2006)</td>
<td>3 instructional assistants; employed in a self-contained classroom</td>
<td>Token economy</td>
<td>NR</td>
<td>Engaged in their normal activity</td>
<td>I: vocal instructions of goals and the procedures&lt;br&gt;&lt;br&gt;&lt;br&gt;F: in-vivo a pager alerted participants to a missed opportunity to engage in target behavior + a self-monitoring form w/ accuracy FB</td>
<td>Positive</td>
</tr>
</tbody>
</table>

Continued
Table 1 continued

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Participants</th>
<th>Targeted Skill</th>
<th>Training Duration</th>
<th>Baseline Training</th>
<th>Intervention Training</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pollard et al. (2014)</td>
<td>4 students btw. 19 to 21 yrs.; deaf ed., English, history ed. &amp; comm. majors</td>
<td>DTT</td>
<td>≈ 2 hr.</td>
<td>I</td>
<td>I: interactive module w/tests M: embedded videos of correct &amp; incorrect implementation R: w/ex &amp; child + self-practice (FB): scoring sheet to provide FB after the session</td>
<td>Positive</td>
</tr>
<tr>
<td>Robinson (2011)</td>
<td>4 paras. btw. 18 to 60 yrs.; high school degree or undergrad. psych. &amp; biology majors; employed in a self-contained classroom</td>
<td>PRT</td>
<td>Average ≈ 1.5 hr.</td>
<td>Engaged in their normal activity</td>
<td>M: model by ex. FB: video FB during which PRT techniques were discussed at the end of the session</td>
<td>Positive</td>
</tr>
<tr>
<td>Rosalls et al. (2015)</td>
<td>3 teachers; BA degree in unrelated field or MA degree in speech &amp; language pathology</td>
<td>SPA</td>
<td>NR</td>
<td>I</td>
<td>I: written instructions embedded in a video M: video model (FB): a single session of FB after the session</td>
<td>Positive</td>
</tr>
<tr>
<td>Roscoe &amp; Fisher (2008)</td>
<td>3 behavior technicians</td>
<td>SPA</td>
<td>≈ 1 hr.</td>
<td>I</td>
<td>R: w/ex. FB: before session watched videotape &amp; given data sheet from last session</td>
<td>Positive</td>
</tr>
</tbody>
</table>

Continued
Table 1 continued

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Participants</th>
<th>Targeted Skill</th>
<th>Training Duration</th>
<th>Baseline Training</th>
<th>Intervention Training</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roscoe et al. (2006)</td>
<td>4 undergrad. trainees</td>
<td>SPA</td>
<td>NR</td>
<td>Told to conduct SPA</td>
<td>1: brief summaries of the assessments R: w/ex. FB: before session watched videotape &amp; given data sheet from last session</td>
<td>Positive</td>
</tr>
<tr>
<td>Salem et al. (2009)</td>
<td>4 undergrad. students</td>
<td>DTT</td>
<td>4.47 hr</td>
<td>I</td>
<td>I: primarily written self-instructional manual w/ tests + short video clip R: w/ex. + self-practice</td>
<td>Mixed</td>
</tr>
<tr>
<td>Severson &amp; Carr (2012)</td>
<td>6 instructors btw. 22 to 33 yrs.; employed in a school for developmental disabilities</td>
<td>DTT</td>
<td>1.5 hr to 3 hr</td>
<td>I</td>
<td>I: written self-instructional manual w/ tests M: video w/embedded text &amp; voiceover script R: w/ex. + self-practice FB: reviewed errors in preceding session + in-vivo practice w/ FB</td>
<td>Positive</td>
</tr>
<tr>
<td>Thiessen et al. (2009)</td>
<td>4 undergrad. students</td>
<td>DTT</td>
<td>4.34 hr</td>
<td>I</td>
<td>I: self-instructional manual w/tests R: w/ex. + self-practice</td>
<td>Positive</td>
</tr>
</tbody>
</table>

Continued
Table 1 continued

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Participants</th>
<th>Targeted Skill</th>
<th>Training Duration</th>
<th>Baseline Training</th>
<th>Intervention Training</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thomson et al. (2012)</td>
<td>8 ABA tutors; employed at an ABA program for children with autism</td>
<td>DTT</td>
<td>4.4 hr</td>
<td>I</td>
<td>I: mainly written self-instructional manual w/ tests + short video clip R: w/ex. + self-practice</td>
<td>Positive</td>
</tr>
<tr>
<td>Vladeasuch et al. (2012)</td>
<td>3 trainees; worked in an early intervention clinic</td>
<td>DTT</td>
<td>NR</td>
<td>I</td>
<td>I: written protocol R: w/ex. M: video model w/voiceover script</td>
<td>Positive</td>
</tr>
<tr>
<td>Wallace et al. (2004)</td>
<td>3 teachers (i.e., special and general education) &amp; a school psychologist</td>
<td>FA</td>
<td>3 hr</td>
<td>I</td>
<td>I: written description &amp; purpose of each condition M: video demonstration of each condition R: w/ex. + trainee (FB): vocal corrective FB immediately after session</td>
<td>Positive</td>
</tr>
<tr>
<td>Ward-Horner &amp; Sturmey (2012)</td>
<td>3 teachers with either a high school BA, or MA degree; reading or special education major</td>
<td>FA</td>
<td>NR</td>
<td>I</td>
<td>I: written description of purpose and procedure M: videotape of conditions R: w/ex. FB: vocal &amp; written FB prior to each session based on preceding session</td>
<td>Positive</td>
</tr>
</tbody>
</table>

Note. I = instructions; M = modeling; R = rehearsal; FB = feedback; NR = not recorded; w/ex. = with the experimenter
**Highlights from the Main Findings**

I wished to identify common training methods used with individuals who usually serve as frontline staff to children with developmental disabilities, but who have no previous experience with the target intervention. In this literature review, I summarized 24 studies in which the authors used a combination of antecedent- and consequent-based strategies to teach novice individuals’ accurate implantation of behavioral technology. The findings indicated that a wide variety of participants took part in the reported studies. Most often, participants included students, teachers, and paraprofessionals with various educational and professional experiences. Educational background was not always reported for all participants who held a degree, but a majority of the participants held a bachelor’s degree in psychology.

Training packages most often included instructions describing the procedures or goals of the intervention coupled with video modeling, rehearsal with a role-played client, and corrective and reinforcing feedback delivered by the experimenter after the session. Some researchers did find aspects of training programs to differ in their efficacy (Roscoe et al., 2006; Moore & Fisher, 2007; Ward-Horner & Sturmey, 2012). Most often, instructions alone were insufficient at bringing participants to mastery, whereas modeling and feedback were found to be the more effective components. The duration of training was noted in a majority of the studies—ranging anywhere from 30 min to about 8 hr—and appeared to depend on the complexity of the trained skills (e.g., SPA training lasted around 30 min, whereas DTT training most often averaged around 3 hr to 4 hr).
In terms of effectiveness, the outcomes of the training packages were relatively similar, with all but two reporting positive or mostly positive outcomes. Inter-observer agreement of at least 80% for at least 20% of all sessions was reported across all the studies; however, fidelity of the independent variable was only assessed in a little less than half. Less than 40% of the studies reported collecting both maintenance and generalization data and about 40% of studies reported collecting either generalization or maintenance data; authors typically found that at least one participant did not maintain criterion levels of performance (e.g., McCulloch & Noonan, 2013; Petscher & Bailey, 2006). Unfortunately, less than half of the studies reported collecting social validity data; researchers most often used Likert-type scales to evaluate outcomes after the intervention had been implemented.

Overall, these findings replicate previous research that has found that (a) multi-faceted training programs that incorporated instruction, modeling, and corrective feedback were often used to teach staff behavioral strategies (Brock & Carter, 2013; Rispoli et al., 2011); (b) a majority of the studies did not report follow-up data to assess the generality of the findings (Brock & Carter, 2013; Rispoli et al., 2011); (c) a majority of the studies reported positive outcomes (Brock & Carter, 2013; Rispoli et al., 2011); (d) the experimenter or members of the research team provided the training (Brock & Carter, 2013); and (e) the duration of training ranged from about 20 min to 8 hr (Brock & Carter, 2013; Rispoli et al., 2011). The aforementioned reviews provided valuable information regarding the characteristics of effective training programs with paraprofessionals. However, this review extends the literature by using a wider lens to examine current
trends in training when applied to novice students, paraprofessionals, teachers, and
behavior technicians who use various behavior analytic procedures. This review also
focused on the social validity of the training procedures, which directly impacts fidelity
and continued use of the training procedures, as well as the duration of training in relation
to the complexity of the trained skills. Finally, this review examined the fidelity of the
individual conducting the training. Without this vital information, we cannot conclude
whether a specific training program was effective or ineffective due to various levels of
fidelity or adequate or inadequate due to the components in the training package.

Limitations

Conclusion based on these results are limited because of several factors that
should be taken into account during interpretation. First, the search terms may have been
too narrow in scope. I may have overlooked some studies because of the terms behavior*
interventions or applied behavior analytic services and the fact that our sample was
limited to individuals (e.g., students, paraprofessionals) who had no prior experience with
the target intervention, but who worked with children with developmental disabilities.
Initially, 84 articles were identified for potential inclusion, and the literature base may be
described as limited because of the small number of included studies (n = 24) in this
review. However, anecdotally, studies in our final review appeared to be representative of
the larger training literature aimed at teaching staff implementation of behavioral
technology.
Second, the extent to which we could analyze and describe the components and effects of the different training packages hinged upon the authors’ description of the procedures. In some studies, however, the descriptions of the training strategies were brief and not at all technological making it difficult, at times, to analyze and summarize their results. Along the same lines, it was difficult to assess whether the training had been carried out as intended because many studies did not report treatment integrity data for the independent variable.

Third, the duration of training was not always explicitly stated; therefore, mean duration of training was sometimes challenging, if not impossible, to assess. Finally, although a definitive method of determining the outcomes of a study (e.g., all individuals had to achieve mastery and performance needed to maintain during follow-up probes) was included, evaluation of the study design (i.e., experimental control) was subjective. Interobserver agreement for experimental control was 100% in this review; however, it is possible that other individuals would have differing opinions about the presence of an experimental effect.

**Recommendations for Future Research and Practice**

Several areas for future research arose as I summarized the literature on training. First, as previously mentioned, the overall effects of individual training components were primarily unknown because a majority of the studies successfully used a combination of multiple training procedures simultaneously (e.g., Hogan et al., 2015; Iwata et al., 2000; Nigro-Bruzzi & Sturmey, 2010). Consequently, it was very difficult to isolate the
variable(s) responsible for the greatest change in behavior and draw conclusions about their relative effectiveness. In the future, researchers should conduct both component and comparative analyses to identify critical training strategies associated with ease of acquisition (see Ward-Horner & Sturmey, 2012). From a clinical stand point, a systematic evaluation of the multiple components could help practitioners develop training programs comprised of elements that are both effective and efficient. Along the same lines, many of the studies had varying training durations. Determining the minimum amount of time and sessions needed to increase treatment fidelity may be beneficial because it will increase the efficiency of staff training. Moreover, it may contribute to the development of an efficient training model that is suitable for widespread use across entire schools or districts. From a practical stand point, this would increase the time teachers spend working directly with students in their classrooms or developing lesson plans by reducing the time teachers spend in training.

Second, the authors assessed generality or collected social validity data in less than half of the studies. Even if the initial results indicated that the training strategy used was effective, definitive conclusions could not be drawn about the program because it was impossible to assess if performance maintained at criterion levels once the training was discontinued or whether skills generalized to novel situations. If performance gains do not maintain and generalize, then training could be deemed as useless by a practitioner or organization providing behavior analytic services. Without a deliberate focus on generality, practitioners may discontinue use of the target intervention. Similarly, if the practitioners do not view the training as important or effective, due to any number of
reasons, the likelihood of its use will decrease. Social validity measures improve the probability of successful adoption, as they give us a better understanding of the key points that lead to companies or school districts embracing and adopting a certain training model. Given that the social validity data collected relied on self-report, which clouds interpretation because of potential social contingencies, future research should focus on developing objective measures of social validity that may lead to training packages that consumers prefer.

Third, many of the studies did not include the delivery of vocal or written instructions that contained a rationale for the targeted intervention. Researchers and expert trainers should consider providing a rationale in an attempt to establish “buy-in” on the part of the participant before training. In conjunction with staff, evaluation of the functional relationship between teacher training and student outcomes should be explored to further establish a collaborative relationship. It is also important to include the collection of treatment fidelity for all behavior change plans because application of the procedures in the applied setting is at risk for low fidelity due to the setting (Wilder et al., 2006), the complexity of the training protocols (e.g., Codding, Feinberg, Dunn, & Pace, 2005) and the limited amount of time and resources (e.g., Mautone et al., 2009). Additionally, training should be viewed as an ongoing process and frequent follow-up probes should be conducted because of the deleterious effects of factors associated with practitioner drift and low treatment fidelity.

Fourth, a majority of the training packages required that an expert trainer (e.g., supervisor) be present for at least a portion of the training program for staff to achieve
mastery; however, many individuals do not have ongoing access to an onsite trainer (Graff & Karsten, 2012). In this review, a few strategies (e.g., a written self-instructional package with questions, tests, pictures; online modules; and video models with embedded text and voice over script) were identified that maximized a trainer’s efficiency in-field, but capitalized on ease of dissemination. Such strategies have considerable potential as low cost, effective training tools (e.g., Garff & Karsten; Pollard et al., 2014; Salem et al., 2009; Thiessen et al., 2009; Weldy et al., 2014). Even if only half the participants achieve mastery with these methods of training, it would greatly cut the time required to train newly hired employees. The considerable efficiency of self-instructional manuals for teaching individuals to conduct basic behavioral procedures renders such training strategies highly useful and warrants further research of their effectiveness in order to refine and translate research findings into diverse applied settings.
Chapter 3: Research Paper

This chapter includes a stand-alone research paper containing a brief literature review, a description of the methods, a summary of the results, and a discussion outlining the implications of the findings.

Applied behavior analysis (ABA) has been recognized as the gold standard treatment of choice for children with autism (Centers for Disease Control and Prevention, 2009). Therefore, training staff to implement behavioral technology is of paramount importance because the efficacy of behavior change plans depends on staff implementing them with high fidelity. Staff training procedures have commonly included two or more of the following components: instructions, modeling, rehearsal, and feedback (Downs, Downs & Rau, 2008; Lavie & Sturmey, 2002; Petscher & Bailey, 2006; Rosales, Stone, & Rehfeldt, 2009; Roscoe, Fisher, Glover, & Volkert, 2006). These procedures have been effective; however, the participants typically achieved mastery only after lengthy training sessions. For example, in McCulloch and Noonan (2013), the duration of training to teach a mand procedure ranged from 3 to 5 hr. Likewise, in Salem et al. (2009), the duration of training to teach DTT averaged 4.47 hr. In both studies, however, only some participants achieved mastery by the conclusion of the intervention. In these situations, access to expert trainers (e.g., behavior analysts) who work with staff who provide services to the
numerous children with special needs is of paramount importance; however, the exact number of behavior analysts working in schools or private agencies is unclear.

Recent data from the National Center for Education Statistics (2014) suggest that 6.4 million children are receiving special education services, but there are only 22,000 certified behavior analysts (Behavior Analyst Certification Board, 2016) who could potentially supervise staff providing these services. Although behavior analysts are not always responsible for directly working with teachers or other staff, they are in charge of training the numerous college students who may eventually work with staff who are in need of training (Behavior Analyst Certification Board, 2016). In addition, special education teachers, who may also be responsible for training staff, report that their university-based coursework and practicum sites do not adequately prepare them for implementing behavioral interventions in-field (Stough, Montague, Landmark, & Williams-Diehm, 2015). Considerable time and resources are allocated to training; therefore, practicing behavior analysts, staff, and students in a behavior analytic program may benefit from efficient, effective methods of training that could reduce reliance on expert-facilitated training.

**Antecedent-Based Strategies**

Researchers have examined the acquisition of skills through the use of antecedent-based strategies that minimize the necessity for expert trainers (e.g., Neef, Parrish, Egel, & Sloan, 1986; Neef, Trachtenberg, Loeb, & Sterner, 1991; Rosales, Gongola, & Homlitas, 2015; Thiessen et al., 2009; Thomson et al. 2012; Weldy & Rapp,
example, video modeling is an effective evidence-based antecedent strategy that has been used to teach implementation of preference assessments as well as a variety of other skills (Catania, Almeida, Liu-Constant, & DiGennaro Reed, 2009; Rosales, Gongola, & Homlitas, 2015; Weldy & Rapp, 2014; Wightman et al., 2012). However, there are limitations associated with the creation of professional videos. First, the viability of video-based interventions may be hindered if the necessary technology (e.g., television, video player) is not readily available (Graff & Karsten, 2012). This is especially true if the organization providing services does not have the funds to allocate resources to creating high-quality videos. Second, developing the videos may be time intensive (e.g., 40 hr to create a video) depending on the availability of experts with software-technical capabilities (Nosik & Williams, 2011). Finally, it may not be practical to take video models into the intervention setting if an individual is in need of supplemental prompts when working with a client (Graff & Karsten, 2012). An alternative antecedent-based strategy that may offer benefits similar to video models, but addresses some of the aforementioned limitations (e.g., ease of availability when working with a client, cost and time to create the videos, and fewer technological demands), is the use of written self-instructional manuals.

Neef, Parrish, Egel, and Sloan (1986) were among the first researchers to evaluate the effects of self-instructional manuals on respite care workers implementation of various respite care skills (e.g., management of problem behavior, bedtime and toileting routines, and medical emergencies). The authors conducted four experiments in which
they examined (1) the relative effects of three different self-instructional manual formats; (2) the effects of the manual as a whole versus a workshop training approach; (3) a simplified training assessment manual capitalizing on available resources and practicality of implementation; and (4) whether respite workers could effectively implement the designated programs.

In Experiment 1, the manuals consisted of concepts supplemented with written examples, pictures alone, or written examples coupled with pictures. To create the manuals, the researchers used pictographs, cartoons, and illustrations. The written information was geared to a seventh-grade reading level and covered four major content areas. The basic format for each content area consisted of behavioral objectives, management strategies that included a rationale for the strategies, and quizzes (including a remedial quiz) after which correct answers were given. The researchers found that the performance of six trainees improved substantially following the presentation of the manuals, all three formats demonstrated similar effects on performance, and results generalized when working with actual clients.

In Experiment 2, the authors found that although there were no significant differences between the manual format versus the workshop format, the manual was rated more favorably than the workshop. Furthermore, a cost-benefit analysis revealed that the cost for the manual was only $7.45 per manual, whereas the rate for the workshop was $9.25 per hr. for the trainer’s time, in addition to the other costs associated with the workshop (e.g., $20 for the overhead and $27 for the videotape). The authors also noted that the distribution of a self-instructional manual affords individuals the chance for
multiple reviews and remedial training could occur if necessary. Alternatively, the
difficulties associated with scheduling and rescheduling workshops, as well the
continuous arrangement of a series of workshops outweigh the benefits of its use. In
Experiments 3 and 4, the authors found that with the written manual, all participants
eventually surpassed the criterion of 85% correct responses and during simulation probes,
participants achieved mastery in all selected skill areas. Although the authors
demonstrated that the manual was effective for teaching a variety of skills, since that
time, little research has been conducted on the use of self-instructional manuals.

More recently, self-instructional technology has greatly expanded because of
various technological advances some of which include the improvement of software tools
to create the manuals and the integration of multimedia. All of these factors have made it
easier to create and disseminate self-instructional manuals. Fazzio and Martin (2007)
created a discrete-trial teaching manual to teach discrete trial training (DTT) to
individuals who work with children with autism. The manual consisted of 21-pages
outlining 19 DTT components. Both Fazio et al. (2009) and Arnal et al. (2007) then
tested the efficacy of this manual. Across the two studies, students took an average of 2.4
hr to master the manual. However, for the four students in Arnal et al., mean performance
at baseline on a 19-item checklist for conducting DTT was 44%. With the introduction of
the manual, mean performance increased to only 67%. Similar results were found in
Fazzio et al., in that mean performance at baseline was 34% and increased to 61% after
reviewing the manual. To achieve mastery, however, the authors showed a video on
DTT, or role-played and provided feedback. Fazzio and Martin (2007) revised the self-
instructional manual based on feedback they received from their participants. In addition to several textual changes, the researchers added a practice component in each section that prompted the participant to stop and rehearse; they also added a brief instructional video to accompany the manual.

Salem et al. (2009) then evaluated the revised 37-page manual with the video. At baseline, four students were given 10 min to read a 1-page summary for teaching a matching task. At intervention, students were given a 37-page manual, study questions, mastery tests, a video demonstration, and self-practice activities. On each mastery test, the students were required to correctly answer at least 50% of the questions randomly selected by the authors. If they did not answer at least 50% correct, students were asked to review the sections of the manual that corresponded to the incorrect answers. The students were then asked to retake the mastery test and were required to answer only the questions that they previously answered incorrectly. The students also watched a 17 min video that included a review of the studied material as well as demonstrations of trials for a matching task. During this time, students were asked to follow along with the manual. There were four self-practice activities (one for each chapter). Students were asked to imagine that they were teaching one of the four tasks to a child with autism, to role play with their imaginary client while using the DTT checklist, and then to rate themselves on role-played items.

The authors found that half of the participants (i.e., 2 of 4) reached the mastery criterion, but all participants significantly improved after baseline (i.e., performance improvement of about 33%). The two students who achieved mastery were then asked to
conduct generalization probes with a 4-year old child diagnosed with autism. The authors found that students averaged 74.4% correct responses when working with the child, which was a decrease from the intervention phase (80% correct responses) with the confederate. The differences in performance, however, may be explained by the fact that, unlike the confederate, the child did not follow prescribed responses on a script.

Similarly, Thomson et al. (2012) found that less than half of the participants (i.e., 3 of 8) met the mastery criterion after reviewing Fazzio and Martin (2007) alone, but the remaining five participants achieved mastery after reviewing a 17-minute video of an expert modeling the teaching of a task. Although these results are promising, their findings were not very robust.

In 2012, Graff and Karsten evaluated the effects of a self-instructional manual to train staff to implement, score, and interpret the outcomes from a paired stimulus (PS) and multiple stimulus without replacement (MSWO) assessment. Participants included 11 teachers, without prior experience with implementing preference assessments, employed at a school for children with autism and related developmental disabilities. At baseline, the authors gave all teachers written instructions from the method sections of Fisher et al. (1992), for the PS assessment, and DeLeon and Iwata (1996), for the MSWO assessment; teachers had 30 min to review the material. Teachers then implemented a stimulus preference assessment (SPA) based on the written instructions alone. At the end of the assessment, teachers calculated selection percentages for all of their collected data, as well as identified the item they would use to teach the client a new skill. The authors then exposed the teachers to two different training procedures: (1) enhanced written
instructions consisting of step-by-step instructions (that contained limited technical jargon), diagrams, pictures, and comprehensive data sheets, or (2) written instructions and data sheets consisting of the same written instructions from baseline with the added data sheet contained in the enhanced written instruction manual. The latter training procedure afforded the authors the opportunity to analyze the effects of the data sheets alone on accurate implementation.

At the end of training, the authors gave teachers hypothetical data and asked them to calculate selection percentages and identify the stimulus that they would use to teach a new skill. In sum, 5 of 11 teachers advanced from written instructions (baseline), to written instructions and data sheet, to enhanced written instructions, to generalization probes with clients from the school setting. The remaining six teachers advanced from written instructions (baseline), to enhanced written instructions, to generalization probes. The authors found that with the written instructions and written instructions plus data sheets, none of the teachers met the mastery criterion (i.e., 90% or greater correct trials across two consecutive sessions). Conversely, with the use of enhanced written instructions, all 11 teachers met the mastery criteria for implementation, scoring, and interpretation, and the skills generalized to clients (i.e., children with disabilities receiving services on site).

Shapiro, Kazemi, Pogosjana, Rios and Mendoza (in press) replicated and extended Graff and Karsten (2012) with undergraduate students. The authors found that none of the participants achieved mastery with the manual developed by Graff and Karsten (2012). The authors then made modifications to their self-instructional manual
based on observations of participants’ errors (e.g., incorrect placement or removal of stimuli) during the prior sessions. Minor modifications to the manual included bolding information about the target responses or increasing the font size of the text. More substantial modifications included the addition of pictures and diagrams (e.g., specified that 1 ft was approximately the vertical length of a piece of letter-size paper or provided multiple examples for calculating selection percentages). The authors found that 6 of 8 participants achieved mastery with the modified manual; two participants needed feedback to meet the mastery criterion.

An influx of newly hired staff requires frequent training from an expert, but constraints such as lack of time or limited financial resources are barriers to receiving effective initial training. Thus, the considerable efficiency of self-instructional manuals for teaching individuals to conduct basic behavioral procedures renders such training strategies highly useful and warrants further research of their effectiveness in order to refine and translate research findings into diverse applied settings. Despite the strengths of the studies listed above, there were differential effects associated with the use of manuals to teach different skills, such as DTT, preference assessments, and respite care responsibilities.

Furthermore, studies contained limitations that could be addressed. First, some researchers did not conduct baseline and intervention generalization probes with actual clients (e.g., Fazzio & Martin, 2007; Graff & Karsten, 2012; Salem et al., 2009; Shapiro et al., 2016), and all researchers did not conduct long-term maintenance probes. Most probes occurred immediately following training or 1 week post intervention (e.g., Fazzio
Second, the use of students in a Behavior Modification course (e.g., Arnal et al., 2007; Fazzio & Martin, 2007; Salem et al., 2009) may not be representative of staff who are in need of training. Third, the need of an experimenter to facilitate the completion and scoring of mastery tests or practice exercises (e.g., Arnal et al., 2007; Fazzio & Martin, 2007; Neef et al., 1986; Salem et al., 2009) precludes examination of the manual’s effectiveness when completed independently, thereby limiting the “self-instructional” component of the intervention.

Fourth, in many studies participants required supplemental procedures, such as modeling or feedback, to achieve mastery (e.g., Arnal et al., 2007; Fazzio & Martin, 2007; Neef et al., 1986; Salem et al., 2009). Fifth, there was a lack of data to indicate the amount of time participants spent reviewing the manuals (e.g., Graff & Karsten, 2012; Neef et al., 1986; Shapiro et al., 2016), or whether the manuals were reviewed during implementation (e.g., Fazzio & Martin, 2007; Graff & Karsten, 2012; Shapiro et al., 2016). Finally, researches have yet to create a written manual for teaching implementation of least-to-most (LTM) procedures and graphing of client data, nor have researches assessed whether acquired skills could generalize to teaching novel skills to clients.

I wished to build on the body of existing research on self-instructional manuals to evaluate (a) the effects of a user-friendly PowerPoint module that did not contain mastery tests or self-practice exercises on newly hired teachers’ acquisition of a PS assessment and LTM procedures; (b) the extent to which acquired skills maintained overtime (e.g., 1
week and 1 mo probes) and generalized to working with actual clients; (c) whether participants could use the LTM procedures to teach a novel skill to an actual client, (d) the total time spent reviewing the training modules and whether the written self-instructional manual was reviewed during implementation; and (e) the extent to which the participants’ found the modules to be effective for acquiring a new skill.

Method

Participants

Participants included six newly hired (i.e., within the last six months) female teachers who provided educational services to students with developmental disabilities. The teachers, between the ages of 22 to 29 years old, all earned a bachelor’s degree and two of six held a state issued moderate-to-intensive teaching license. The teachers received initial on-the-job training on basic principles of behavior analysis and data collection procedures. However, the teachers had no coursework in behavior analysis or experience observing or conducting SPAs or LTM procedures. Teachers were recruited by disseminating an IRB-approved email to the director of the facility. The email indicated that teachers could attend an SPA and LTM training session and elect to participate in a research study designed to improve training procedures. The teachers did not receive any funding as compensation for their participation.

Setting and Materials

All training sessions were conducted in small rooms located at the private school. For the LTM procedures, participants were given the items needed to teach a multi-step
self-care skill of washing hands, including hand soap and paper towels. The room was furnished with desks, chairs, and a sink. For the PS assessment, participants were given the items necessary to conduct and interpret the outcomes, including paper, writing utensils, a calculator, and a pool of eight stimuli (licorice, gummy bears, cookies, popcorn, silly putty, stickers, bubbles, and an action figure). The items in the room were arranged such that the role-played client always sat across from the participant.

**Role-played client assessment scripts.** The program research coordinator, with previous experience conducting SPAs and using LTM procedures, was trained to behave as the role-played client throughout the study. The role-played client had access to the simulated scripts during the experiment, which were kept out of view from the participants.

**LTM prompting.** Four scripts were developed based on the published article by Ault and Griffen (2013) and one of four was randomly assigned to each session. The sequence of client responses were varied and specified the exact sequence of steps the role-played client was required to follow across ten trials (i.e., one session). The client was required to emit ten responses to complete the behavior chain of washing hands: (a) turn on the water; (b) wet hands for 3 to 5 s; (c) place one hand on the pump dispenser; (d) place the other hand below the spout; (e) push the pump one time; (f) rub hands together for 5 to 10 s; (g) rinse hands for 5 to 10 s; (h) turn off the water; (i) take a paper towel; and (j) dry hands with the paper towel. Some responses required the role-played client to emit the aforementioned responses without a prompt (2 of 10 trials), whereas
others required either a vocal (3 of 10 trials), model (2 of 10 trials), or physical (3 of 10 trials) prompt to complete the sequence.

**Scoring and graphing for the LTM procedures.** At the end of baseline and written instruction phases and during maintenance probes, participants were given hypothetical data and were asked to score and graph students’ performance. The experimenter created four different sets of hypothetical data and counterbalanced the order of presentation across participants. The data sheets included information about students’ unprompted and prompted responses across five sessions. Also included was a place for participants to track the client’s performance using a weighted point system (e.g., an independent level—place a 3 on the data sheet; a full physical guidance prompt—place a 0 on the data sheet). Based on their calculations, participants were required to create hand drawn graphs that reflected the client’s overall performance.

**PS assessment.** The same scripts developed by Graff and Karsten (2012) were used and one of four scripts was randomly assigned to each session. Across all scripts, I varied the sequence of client responses and specified the exact trial in which the role-played client was to emit prescribed responses (i.e., atypical and typical responses) across 10 trials (i.e., one session). Typical responses (5 of 10 trials) included selecting one stimulus within 5 s of the participant placing both stimuli on the table and instructing the client to select a stimulus. Atypical responses (i.e., the other 5 trials) included the client (a) choosing both stimuli 2 to 3 s after the participant instructed the client to select a stimulus (1 of 10 trials); (b) choosing both stimuli in quick succession immediately after the delivery of a vocal prompt to select a stimulus (1 of 10 trials); (c) choosing a stimulus
that did not appear in the predetermined pool of stimuli (1 of 10 trials); or (d) engaging in a behavior other than selecting any stimuli (2 of 10 trials).

**Assessment summary and interpretation materials for the PS assessment.** At the end of baseline and written instruction phases and during maintenance probes, participants were given hypothetical data and were asked to calculate selection percentages and interpret assessment outcomes. The experimenter created three different hypothetical sets of data using the data sheets included in the self-instructional manual developed by Graff and Karsten (2012) and counterbalanced the order of presentation across participants. The data sheets included a place for the participants to calculate selection percentages from 28 trials and identify the stimuli they would use to teach a new skill from the same eight stimuli presented during the assessment.

**Response Measurement**

**Participant target responses for the LTM procedures.** The mastery criterion was implementation of the procedure with 90% or greater correct trials across two consecutive sessions. On each trial, observers scored seven specific target responses as correct or incorrect. Specifically, observers recorded whether participants correctly emitted the following target responses during each trial: (a) vocal direction, observers scored a correct response if the participant delivered the direction to wash hands (e.g., “Go and wash your hands.”); (b) vocal prompt, observers scored a correct response if the participant allowed the client 3 to 5 s to respond before delivering a vocal prompt for the target step; (c) vocal and model prompt, observers scored a correct response if the participant allowed the client 3 to 5 s to respond before delivering a vocal prompt for the
target step coupled with in-vivo modeling; (d) vocal and physical prompt, observers scored a correct response if the participant allowed the client 3 to 5 s to respond before delivering a vocal prompt for the target step coupled with a hand-on-hand physical prompt; (e) reinforcer delivery, observers scored a correct response if the participant delivered social praise contingent upon independent or prompted completion of the target steps; (f) data collection, observers scored a correct response if the participant recorded the client’s response as independent or the highest prompt level provided; and (g) scoring and graphing of student behavior, observers scored a correct response if the participant correctly scored the students’ data using a weighted point system and then graphed the students’ unprompted and prompted responses.

**Participant target responses for the PS assessment.** The mastery criterion was implementation of the preference assessment with 90% or greater correct trials across two consecutive sessions. Each PS assessment included 10 trials; on each trial, observers scored six specific target responses as correct or incorrect. Participant target responses were defined in a manner identical to that described in Graff and Karsten (2012). Specifically, observers recorded whether participants accurately emitted the following target responses during each trial: (a) stimulus presentation, observers scored a correct response if two stimuli were placed on the table in front of the client; (b) stimulus position, observers scored a correct response if the stimuli were placed approximately 1 ft (i.e., 6 in. to 1 ft ½ in.) in front of the client and approximately 1 ft from one another; (c) postselection response, observers scored a correct response if the non-selected stimulus (i.e., the item that was never selected by the client) was removed before recording the
client’s selection; (d) response blocking, observers scored a correct response if participants moved their hands forward towards the client’s hands to block the selection of multiple stimuli; (e) trial termination, observers scored a correct response if participants terminated the trial (i.e., removed all stimuli from the table) when no client response occurred within 4 to 7 s after the delivery of a prompt to select a stimulus; and (f) and scoring and interpretation, observers scored a correct response if the participants correctly calculated selection percentages and identified the high-preference item.

I calculated percentage of correct trials for implementation by dividing the number of trials during which the participant correctly emitted all target responses specific to implementation by the total number of trials implemented and multiplying by 100. If a participant implemented any one of the target responses incorrectly, the entire trial was scored as incorrect. For example, if during the LTM procedure the participant implemented two of three target responses for that trial correctly, but failed to provide reinforcement for task completion, the entire trial would be scored as incorrect. Percentage of correct trials for a specific target response was calculated by dividing the number of correctly implemented target responses by the overall number of opportunities to engage in the target response during each session and multiplying by 100.

**Participant target responses for scoring and graphing student performance.**

A correct response was recorded if the participant accurately scored and graphed the students’ performance. A correct response required the participant to (a) correctly score the student’s data using a weighted point system; (b) correctly label the horizontal axis (i.e., sessions); (c) correctly label the vertical axis (i.e., points earned); (d) correctly label
the condition labels (e.g., baseline, LTM intervention); (e) correctly add the phase change line; and (f) correctly graph multiple data paths that corresponded to the unprompted and prompted data on the hypothetical data sheet.

**Participant target responses for the PS assessment summary and interpretation.** A correct response was recorded if the participant accurately calculated selection percentages from a set of hypothetical data and correctly identified the high preference item (i.e., items selected on 80% or greater of the trials) that should be used to teach the hypothetical client a new skill.

**Duration spent reviewing the module.** Data were collected on the amount of time participants reviewed the modules during training.

**The number of times participants referred to the written self-instructional manual.** The printed written self-instructional manual was identical to the module and was provided to the participants in case they were in need of supplemental prompts when implementing the procedures. Data were collected on the number of times participants referred to the written self-instructional manual during implementation.

**Interobserver Agreement and Measurement Fidelity**

**Correct implementation.** All sessions were videotaped and scored by the experimenter in vivo; a second graduate research assistant independently scored 33% of all sessions via videotape. For total percentage of correct trials during implementation, an agreement was defined as both observers recording the trial as correctly implemented (i.e., all target responses for that trial scored as correct). A disagreement was defined as only one observer scoring correct implementation for a trial. The experimenter divided
the number of agreements by the number of agreements plus disagreements and multiplied by 100. The mean agreement for total percentage of correct trials during implementation for the LTM procedures was 94% (range, 90% to 100%), and for PS assessment was 93% (range, 80% to 100%). I also calculated percentage of correct trials for each specific target response. An agreement was defined as both observers recording an error or correct implementation for a specific target response during a given trial. A disagreement was defined as only one observer recording a specific target response as correct for that trial. I divided the number of agreements for each target response by the number of agreements plus disagreements and multiplied by 100. Mean agreement for specific participant target responses for the LTM procedures was 95% (range, 90% to 100%) and for the PS assessment was 93% (range, 86% to 100%).

**Accuracy of scoring and graphing the LTM procedure.** I calculated accuracy of scoring and graphing for 100% of all sessions. An agreement was defined as both observers recording scoring and graphing of hypothetical data as correct or incorrect; a disagreement was defined as only one observer recording scoring and graphing as correct or incorrect. Mean agreement for graphing was 100%.

**Accuracy of assessment summaries and data interpretation.** I calculated accuracy of scoring and interpreting outcomes from the PS assessment for 100% of all sessions. An agreement was defined as both observers recording scoring and identification of the highly preferred item as correct or incorrect; a disagreement was defined as only one observer scoring the assessment summary and interpretation as correct or incorrect. Mean agreement for participant scoring and interpretation was 100%.
The procedural fidelity of the role-played client and experimenter was also assessed. Via a video recording, the experimenter scored a trial as correct if the role-played client correctly emitted the actions prescribed on the script; a trial was scored as incorrect if the client failed to emit the prescribed responses. Integrity for implementation was calculated on a trial-by-trial basis by dividing the number of correctly implemented trials by the number of total trials and multiplying by 100. Mean integrity across all trials was 94% (range, 85% to 100%) for the LTM procedures, and 100% for the PS assessment. I calculated procedural fidelity for the experimenter on a session-by-session basis across 33% of all sessions and phases. A correct response was recorded if the experimenter correctly implemented the corresponding steps prescribed on our task analysis. Mean fidelity for the LTM procedures was 95% (range, 90% to 100%), and for the PS assessment was 100%. In regard to the amount of time spent reviewing the module during training, agreement was calculated by dividing the smaller number by the larger number and multiplying by 100; interobserver agreement was 100%. The number of times participants referred to the written self-instructional manual during implementation was also calculated by dividing the smaller number by the larger number and multiplying by 100; interobserver agreement was 100%.

**Experimental Design**

A multiple probe design across participants and skills was used to evaluate the effects of a self-instructional module on accurate implementation of a LTM procedure and PS assessment.
Procedure

**Baseline.** A probe was conducted to determine if the participants’ had the targeted skills in their repertoire. The teachers were supplied with the items necessary to conduct LTM procedures or a PS assessment. At the beginning of each session, the teachers were told to either (a) determine the client’s preference, or (b) teach the client how to complete the multi-step skill of washing hands using LTM prompting. Probes occurred with the role-played client and feedback was provided regarding when to initiate and complete a session; no feedback was provided for correct and incorrect responses.

**Written instructions LTM prompting procedures.** Participants were given an excerpt from an article written by Ault and Griffen (2013) outlining the procedures for implementing LTM prompting. In the instructions, participants were informed that “in a typical trial sequence, the teacher delivers a task direction and waits a specified response interval (i.e., 3 to 5 s) for the student to respond independently. However, if an error or no response occurs, the teacher delivers the first prompt and waits the specified response interval for the student to respond. If an error or no response occurs, the teacher delivers the next more intrusive prompt and waits the specified response interval. This sequence continues until the student responds correctly or until the teacher delivers all the prompts in the hierarchy. If the student responds correctly at any time in the sequence, either independently after the task direction or following a prompt, the teacher reinforces the response and records the independent response or the prompt level that resulted in the correct response by the client.” The hierarchy was as follows: (a) an independent level (place a 3 on the data sheet); (b) a verbal prompt; tell the client what to do (place a 2 on
the data sheet); (c) a verbal and model prompt; tell the client what to do and model how to do it (place a 1 on the data sheet); and (d) a verbal and full physical guidance prompt; tell the client what to do and help them do it (place a 0 on the data sheet). Information regarding graphing the students’ performance following the procedure was also provided.

Written instructions for the PS assessment. Participants were given the an excerpt from an article written by Fisher et al. (1992), which contained the following directions: (a) place two items on the table approximately 1 ft in front of the client and 1 ft apart; (b) provide a vocal verbal prompt for the selection response; (c) remove the non-selected item and record the selected item on the data sheet; (d) record “no response” if an item is not selected within 5 s of the prompt; and (e) block the client’s attempt to simultaneously select more than one item. Information regarding calculating selection percentages and identifying the high-preference item was also provided.

Self-instructional modules for LTM prompting procedures. The experimenter used the article from Ault and Griffen (2013) to develop the self-instructional module. The 17-page PowerPoint module included four main sections. The first section included information about the prompting hierarchy and how to collect data. The second section included examples about how and when to implement each prompt. The third section included information about reinforcing each unprompted and prompted step. The fourth section included information on how to graph students’ performance. Each section contained approximately 100 words.

Self-instructional modules for preference assessments. The experimenter used the written self-instructional manual from Shapiro et al. (2016) modified from the self-
instructional manual developed by Graff and Karsten (2012). The 8-page PowerPoint module included three main sections. The first section included information about behaviors to emit prior to the session (e.g., selecting items for the assessment, randomly assigning the items to each trial). The second section included information about conducting the assessments (e.g., stimulus position, response blocking). The third section included information pertaining to scoring the assessment (e.g., calculating preference hierarchies, interpreting data). Each section contained approximately 100 words.

Participants in the written instruction phase were given 30 min to review the written material from Ault and Griffen (2013) and Fisher et al. (1992) before initiating the first session. Participants were then asked to inform the experimenter when they were ready to begin. The role-played client provided no feedback to the participants except to indicate the initiation and completion of a session. At the end of baseline and written instruction phases and during maintenance probes for the LTM procedures and PS assessment, the experimenter gave the participants hypothetical data and asked them to score and graph the data or to calculate selection percentages and interpret assessment outcomes (i.e., identify the high preference item they would use to teach a new skill).

**Generalization and Maintenance Probes**

The experimenter conducted in-situ generalization probes during baseline and training assessments. Generalization probes occurred at the private school with each of the participants’ respective clients (i.e., three boys ranging in age from 7–9 years, diagnosed with a developmental disability). The experimenter conducted maintenance probes approximately 1 week and 1 month after the teachers met the mastery criterion.
During the 1 month maintenance probe for the LTM procedure, participants were asked to conduct the procedure with their respective clients. To examine the generality of the skill across other skills, participants were asked to use the LTM procedure to teach a novel behavioral sequence (e.g., shoe tying) to an actual client. Similar to the procedure for hand washing, the client was required to emit ten responses to complete the chain of shoe tying: (a) pick up right lace with right hand; (b) pick up left lace with left hand; (c) cross laces to make an X; (d) fold lace on right side through hole using right hand; (e) pull both laces tight; (f) make a loop with the right lace, pinch with thumb and index finger (hold it); (g) make a loop with the left lace, pinch with thumb and index finger (hold it); (h) cross laces to make an X; (i) fold lace on the right through the hole using the right hand; and (j) pull both laces tight. The 1 month maintenance probe for the PS assessment was also conducted with the participants’ clients. Prior to the assessment, participants were asked to select eight items for each of their clients based on information provided by staff and the participants’ observations when working with their clients. Teachers identified several potential reinforcers, which included food, toys, activities, and social interactions (e.g., hugs).

**Social Validity Assessment**

At the completion of the study, the experimenter provided each participant with a brief questionnaire containing nine questions, which the participant completed anonymously. Participants were asked to indicate their degree of agreement or disagreement on a 5-point rating scale (“Strongly disagree to “Strongly agree”) with questions aimed at the importance of the treatment goals, the acceptability of the training
method used, and the effectiveness of the outcomes. Participants were also asked two open ended questions pertaining to the least and most helpful parts of the self-instructional module. Example questions included “Using least-to-most prompting procedures to teach a skill is important,” An instructional module can be used when training clinicians, teachers, and/or caregivers to conduct assessments with children,” I feel confident that I can correctly implement the procedure,” “The module was easy to follow,” and “The most or least helpful part(s) of the self-instructional module was/were the following.”
Chapter 4: Results

On average, the total training time for the six participants who completed the modules was 18 min for the LTM procedure, and 12 min for the PS assessment. Participants did not refer to the written self-instructional manual during the procedures. Figure 1 depicts the percentage of trials with correct implementation for the three teachers that progressed from baseline, to the written instructions from Ault and Griffen (2013) and Fisher et al, (1992), to the module for the LTM procedures and the PS assessment. At baseline for the LTM procedures, none of the participants met the mastery criterion: Rosa \( M = 13\% \) (range, 10\% to 20\%); Sophia \( M = 30\% \) (range, 20\% to 40\%); and Star \( M = 13\% \) (range, 10\% to 20\%). When participants were provided with the written instructions, mean performance increased slightly: Rosa \( M = 37\% \) (range, 20\% to 50\%); Sophia \( M = 50\% \) (no range); and Star \( M = 30\% \) (range, 20\% to 40\%). Written instruction probes with actual clients for Rosa, Sophia, and Star were 50\%, 30\%, and 40\%, respectively. With the introduction of the module, all participants achieved mastery: Rosa \( M = 90\% \) (no range); Sophia \( M = 90\% \) (range, 80\% to 100\%); and Star \( M = 90\% \) (no range). Training probes with clients for Rosa, Sophia, and Star were 90\%, 100\%, and 90\%, respectively. At baseline for the PS assessment, none of the participants met the mastery criterion: Rosa \( M = 0\% \); Sophia \( M = 0\% \); and Star \( M = 0\% \). When we provided participants with the written instructions, mean performance increased slightly:
Rosa $M = 40\%$ (no range); Sophia $M = 30\%$ (range, 10\% to 30\%); and Star $M = 30\%$ (no range). Written instruction probes with actual clients for Rosa, Sophia, and Star were 30\%, 50\%, and 40\%, respectively. With the introduction of the module, all participants achieved mastery: Rosa $M = 95\%$ (range, 90\% to 100\%); Sophia $M = 90\%$ (range, 80\% to 100\%); and Star $M = 90\%$ (no range). Training probes with clients for Rosa, Sophia, and Star were 100\%, 100\%, and 90\%, respectively. During 1 week and 1 month maintenance probes, all participants maintained performance at mastery levels for both procedures.
Figure 1. PS assessment and LTM implementation graphs for the participants who progressed through baseline, to written instruction from Ault and Griffen (2013) and Fisher et al. (1992), to the modules. Open squares & circles = probes with clients. Open diamonds for the LTM procedure indicate generalization to a different skill. Maint. = Maintenance.
I also evaluated the participants’ graphing of unprompted and prompted responses for the LTM procedure. At the written instruction phase, none of the participants accurately graphed the data. Following the module, Rosa, Sophia, and Star correctly graphed the data. During 1 week and 1 month maintenance probes, results mirrored those found following the presentation of the module in that Rosa, Sophia, and Star maintained performance.

In regard to scoring the PS assessment and interpreting the outcomes. At the written instructions phase, none of the participants accurately calculated selection hierarchies and interpreted the outcomes. Following the module, Rosa, Sophia, and Star accurately generated preference hierarchies and identified the high preference item. During the 1 week and 1 month maintenance probes, results mirrored those found following the presentation of the model in that Rosa, Sophia, and Star maintained performance.

Figure 2 depicts the percentage of trials with correct implementation for the three teachers who progressed from baseline, to the module for the LTM procedures and PS assessment. At baseline for the LTM procedures, none of the participants met the mastery criterion: Shir $M = 23\%$ (range, 20\% to 30\%); Aisha $M = 47\%$ (range, 40\% to 50\%); and Cai $M = 54\%$ (range, 50\% to 60\%). Baseline probes with actual clients for Shir, Aisha, and Cai were 30\%, 30\%, and 50\%, respectively. With the introduction of the module, all participants met the mastery criterion: Shir $M = 87\%$ (range, 80\% to 100\%); Aisha $M = 90\%$ (range, 80 to 100\%); and Cai $M = 100\%$. Training probes with clients for Shir, Aisha, and Cai were all 100\%. 

76
At baseline for the PS assessment, none of the participants met the mastery criterion: Shir $M = 0\%$; Aisha $M = 0\%$; and Cai $M = 0\%$. Baseline probes with actual clients for Shir, Aisha, and Cai were 0%, 0%, and 0%, respectively. With the introduction of the module, all participants met the mastery criterion: Shir $M = 90\%$ (range, 90% to 100%); Aisha $M = 90\%$ (range, 90 to 100%); and Cai $M = 100\%$. Training probes with clients for Shir, Aisha, and Cai were 100%, 100%, and 100%, respectively. See Tables 2 and 3 for detailed information regarding specific target responses for each procedure. During 1 week and 1 month maintenance probes, all participants maintained performance at mastery levels for both procedures.
Figure 2. PS assessment and LTM procedures implementation graphs for the participants who progressed through baseline, to the modules. Open squares & circles = probes with clients. Open diamonds for the LTM procedure indicate generalization to a different skill. Maint. = Maintenance.
Table 2. Mean percentage of correct trials for specific target responses across participants for the LTM procedures. Note. Parentheses next to the dependent variables (DVs) = the number of opportunities participants could emit the target response each session. WI = written instructions.

<table>
<thead>
<tr>
<th>DV</th>
<th>Baseline</th>
<th>WI</th>
<th>Module</th>
<th>Baseline</th>
<th>Module</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task Direction (1)</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Reinforcement (10)</td>
<td>80</td>
<td>80</td>
<td>90</td>
<td>65</td>
<td>94</td>
</tr>
<tr>
<td>Recording data (10)</td>
<td>0</td>
<td>63</td>
<td>94</td>
<td>0</td>
<td>92</td>
</tr>
<tr>
<td>Vocal Prompt (3)</td>
<td>50</td>
<td>60</td>
<td>88</td>
<td>56</td>
<td>100</td>
</tr>
<tr>
<td>Modeling (2)</td>
<td>0</td>
<td>55</td>
<td>100</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Physical guidance (3)</td>
<td>70</td>
<td>100</td>
<td>100</td>
<td>67</td>
<td>100</td>
</tr>
</tbody>
</table>
I also evaluated the participants’ graphing of unprompted and prompted responses for the LTM procedure. At baseline, none of the participants accurately graphed the data. Following the module Shir, Aisha, and Cai, correctly graphed the data. During the 1 week and 1 month maintenance probes, results mirrored those found following the presentation of the model in that Shir, Aisha, and Cai maintained performance.
In regard to scoring the PS assessment and interpreting the outcomes, at baseline, none of the participants accurately calculated selection hierarchies and interpreted the outcomes. Following the module, Shir, Aisha, and Cai accurately generated preference hierarchies and identified the high preference item. During the 1 week and 1 month maintenance probes, results mirrored those found following the presentation of the model in that Shir, Aisha, and Cai maintained performance.

Finally, I assessed whether the six teachers who mastered the LTM prompting could use the procedure to teach a novel skill to an actual client. Generalization probes with actual clients for Rosa, Sophia, Star, Shir, Aisha, and Cai were 100%, 90%, 90%, 100%, 90%, and 100%, respectively.

See Tables 4 and 5 for detailed information regarding specific answers to the social validity questionnaire. Participants reported that the module easy to follow, and they felt confident that they could correctly implement the procedures. Participants noted that compared to the written instructions from Ault and Griffen (2013) and Fisher et al. (1992), the addition of colorful pictures and diagrams in the module facilitated acquisition of the procedures. Participants did not provide any suggestions for improving the modules.
<table>
<thead>
<tr>
<th>Questions</th>
<th>Mean rating (1 = strongly disagree, 5 = strongly agree)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using least-to-most prompting procedures to teach a skill is important.</td>
<td>4.7 (range, 4 to 5)</td>
</tr>
<tr>
<td>This procedure is an appropriate way to assist a client with learning new skills.</td>
<td>4.5 (range, 4 to 5)</td>
</tr>
<tr>
<td>An instructional module can be used when training clinicians, teachers, and/or caregivers to implement least-to-most prompting strategies with children.</td>
<td>4.7 (range, 4 to 5)</td>
</tr>
<tr>
<td>Providing an instructional module is an effective way to train clinicians, teachers, and/or caregivers to implement least-to-most prompting strategies with children.</td>
<td>4.8 (range 4 to 5)</td>
</tr>
<tr>
<td>I feel confident that I can correctly implement least-to-most prompting strategies with children.</td>
<td>4.3 (range 3 to 5)</td>
</tr>
<tr>
<td>I would recommend the use of self-instructional modules when teaching novice individuals these skills.</td>
<td>4.8 (range, 4 to 5)</td>
</tr>
<tr>
<td>The module was easy to follow.</td>
<td>4.8 (range 4 to 5)</td>
</tr>
</tbody>
</table>

Table 4. Participants’ responses to the social validity questionnaire for the LTM prompting procedures
<table>
<thead>
<tr>
<th>Questions</th>
<th>Mean rating (1 = strongly disagree, 5 = strongly agree)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using a preference assessment to identify a client’s preference is important.</td>
<td>4.5 (range, 4 to 5)</td>
</tr>
<tr>
<td>Conducting the assessment is an appropriate way to identify items that would motivate a client.</td>
<td>4.5 (range, 4 to 5)</td>
</tr>
<tr>
<td>An instructional module can be used when training clinicians, teachers, and/or caregivers to conduct assessments with children.</td>
<td>4.8 (range, 4 to 5)</td>
</tr>
<tr>
<td>Providing an instructional module is an <em>effective</em> way to train clinicians, teachers, and/or caregivers to conduct assessments with children.</td>
<td>4.8 (range, 4 to 5)</td>
</tr>
<tr>
<td>I feel confident that I can correctly conduct a paired stimulus preference assessment.</td>
<td>4.3 (range, 3 to 5)</td>
</tr>
<tr>
<td>I would recommend the use of self-instructional modules when teaching novice individuals these skills.</td>
<td>4.7 (range, 4 to 5)</td>
</tr>
<tr>
<td>The module was easy to follow.</td>
<td>4.8 (range, 4 to 5)</td>
</tr>
</tbody>
</table>

Table 5. Participants’ responses to the social validity questionnaire for the PS assessment
Chapter 5: Discussion

The main objective of this study was to build on the literature surrounding the use of self-instructional manuals (e.g., Arnal et al., 2007; Graff & Karsten, 2012; Neef, Parrish, Egel, & Sloan, 1986; Shapiro et al., 2016) and address several gaps in the current literature. For example, the lack of baseline and intervention generalization probes with actual clients, the need for an experimenter to facilitate the completion and scoring of mastery tests or practice exercises, the absence of long-term maintenance probes, the lack of information surrounding the amount of time participants spent reviewing the manuals or whether the manuals were reviewed during implementation, and the creation of a module to teach implementation, scoring, and graphing of a LTM procedure.

Specifically, I evaluated (a) the effects of a user-friendly PowerPoint module that did not include mastery tests or self-practice activities on acquisition; scoring; and interpreting outcomes, for the PS assessment, and acquisition; scoring; and graphing data, for the LTM procedures; (b) the extent to which acquired skills maintained over an extended period of time and generalized to working with clients; (c) whether participants could use the LTM procedures to teach a new behavioral chain (i.e., tying shoes); (d) the total time spent reviewing the training modules and whether the written self-instructional manual was reviewed during implementation; and (e) the extent to which the participants found the modules to be effective for acquiring a new skill and whether any changes
modifications could be made to facilitate acquisition. This study looked at the social validity of the training methods used because it is an important consideration for behavior analytic agencies when adopting a specific training method. Social validity data also help bridge the gap between research and practice.

Consistent with the results from other studies aimed at teaching correct implementation of PS assessments or DTT procedures, the findings indicated that individuals without any training could not accurately implement the procedures with written instructions from Ault and Griffen (2013), or Fisher et al. (1992). That is, none of the three teachers who progressed through baseline, to written instructions, to the PowerPoint module for the LTM prompting procedures and the PS assessment met the mastery criterion at baseline or with the addition of written information alone. However, with the introduction of the PowerPoint modules, all teachers achieved mastery and results maintained and generalized to students with whom they were working (see Graff & Karsten, 2012).

Similar results were found for the three teachers who progressed through baseline, to the PowerPoint module for the LTM procedures and the PS assessment in that all participants, in the absence of written instructions, achieved mastery after reviewing the self-instructional modules and results maintained when working with actual clients. I also found that all participants could teach a novel skill (i.e., shoe tying) to an actual client. In regard to scoring and graphing data, all participants were able to score and interpret the outcomes for the PS assessment and all could graph the data for the LTM procedures. Notably, although all participants in this study had continued access to the written
instructions obtained from the self-instructional modules, participants did not review the material once the sessions commenced. Finally, I found the modules to be efficient in that the total time spent reviewing the training modules was approximately 19 min (range, 17 m to 20 m), for the LTM procedure, and 13 min (range, 12 m to 15 m), for the PS assessment, and participants deemed the modules to be effective for teaching a new skill.

The findings in this study were similar to previous studies in that participants achieved mastery following a manual that used pictures, diagrams, and minimal technical language (Graff & Karsten, 2012; Neef, Parrish, Egel, & Sloan, 1986; Shapiro et al., 2016). Although this was the first study to evaluate the effects of modules on the acquisition of LTM procedures, the results were promising and the procedures should be replicated by other researchers to determine if similar results are achieved.

Limitations and Future Research

This study contributes further evidence that a PowerPoint module could be used to teach staff to conduct preference assessments and could be extended to teach LTM procedures. Although favorable results were found, there were some limitations that deserve mention. A first potential limitation is that stimuli identified as preferred items in the PS assessment were not corroborated with a reinforcer assessment. Researchers have demonstrated that results from comprehensive preference assessments are likely to function as reinforcer assessments (DeLeon & Iwata, 1996; Fisher et al., 1992). However, in this study, it is unclear whether the preferred stimuli would have reinforcing effects. A second potential limitation is that data were not collected on student performance.
Researchers should evaluate the long-term outcomes of the training method, including those related to student outcomes, so as to validate whether meaningful outcomes for clients were related to the fidelity of the behavioral strategies implemented. This information would further confirm the importance of teacher performance on student outcomes. Unfortunately, these two limitations were not addressed because probes with students at the school were difficult given teachers’ and students’ availability throughout the day. Probes required removing the children from the instructional setting and parents agreed to only a few assessments with their children.

A third potential limitation was that no data were collected on the time it took to create the online modules (anecdotally, it took about a half hour per module). Researchers may consider doing a cost-benefit analysis (see Neef, Parrish, Egel, & Sloan, 1986) because there is a possibility that agencies will begin to adopt these self-instructional modules to use during training. It is important to note that video modules, another effective antecedent-based training strategy, can be costly; for example, Nosik and Williams (2013) reported that the development of one 20 m comprehensive training video for DTT or backward chaining took approximately 40 hr to complete. However this limitation may be mitigated by the fact that the cost per trainee decreases over time as more personnel are trained using the same video. An additional avenue for future research when a self-instructional module alone is not effective, is either embedding reading checks (multiple choice or fill-in-the-blank questions) into the module during which the participants would receive immediate feedback or creating a written test consisting of open-ended questions that would be completed and scored immediately.
after reviewing the module. Researchers should also continue to examine the
effectiveness of self-instructional modules when disseminated to the broader community
(e.g., parents), as well as determine the conditions under which a self-instructional
modules would be beneficial. A fourth potential limitation was the modules focused only
on implementation of the procedure with high fidelity. In the future, researchers should
consider creating modules that facilitate instructional decision making. For example,
identifying when to use these procedures with a particular student or acquisition of a
particular skill. A fifth potential limitation surrounded the fact that it was unclear which
components (e.g., pictures, diagrams, nontechnical language) of the self-instructional
module were most beneficial for achieving mastery. In the future, researchers should
isolate and evaluate each component independently. Finally, I used percentage of correct
trials to calculate the fidelity of implementation. In the future, researchers may consider
using a more sensitive measure (e.g., percentage of correct steps) to calculate procedural
fidelity.

Implications and Conclusion

An important consideration is that some procedures (e.g., functional analyses) or
behaviors (e.g., self-injurious) cannot safely or ethically be carried out or addressed
without extensive staff training and feedback. However, it is possible that for more
complex low-risk behavioral strategies, self-instructional manuals alone (Graff &
Karsten, 2012) or with brief supplemental feedback (Arnal et al., 2007; Fazzio et al.,
2009) could better facilitate acquisition of skills. Although live access to an expert trainer
may constitute an ideal training scenario, this may be impractical for agencies that are
constantly hiring new employees, especially if there are only a few certified behavior analysts on staff. Therefore, even if half of the participants achieve mastery with a manual, this would greatly reduce the need for an onsite trainer. Moreover, with the advancement of technology, it would be possible to embed a semi-professional video model into the module for participants that need extra assistance to achieve mastery. Although some trainees may require more intensive training, such as vocal feedback, the self-instructional manual allows trainers to allocate limited resources to needed areas.

Taken together, results from this study (a) offer further support for the use of self-instructional manuals; (b) represent an important step forward in finding efficient models of training; and (c) provide robust evidence that self-instructional modules can be used to effectively teach a LTM prompting procedure and a PS assessment. Given that this low cost, portable intervention eliminated the need for expert-facilitated training, further research is warranted to determine whether well-developed self-instructional modules can be adapted to a variety of low-risk applied technologies (e.g., prompt fading, data collection). This information may be beneficial for practitioners who have little experience with the implementation of a PS assessment and LTM procedures and/or are responsible for training newly hired staff how to implement these procedures. Furthermore, this information would also be advantageous to providers of behavior analytic services who have limited time and resources to train the consistent influx of newly hired staff and may appeal to many human service organizations, especially if the turnover rates are high.
References


93


* Indicates studies included in the review
<table>
<thead>
<tr>
<th></th>
<th>Task direction, waiting 3-5 s</th>
<th>VP, waiting 3-5 s</th>
<th>VP+MP, waiting 3-5 s</th>
<th>VP+PG, waiting 3-5 s</th>
<th>Reinforcing</th>
<th>Recording</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Turn on water</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Wet hands for 3-5 s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Place one hand on pump dispenser</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Place other hand below soup spout</td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Push pump 1 time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Rub hands together 5-10 s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Rinse hands 5-10 s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Turn off water</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Take paper towel</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Dry hands with paper towel</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX B

WRITTEN INSTRUCTIONS FOR THE LTM PROCEDURE
The system of least prompts (SLP) or least-to-most prompting procedure involves defining a hierarchy of prompts and then systematically delivering those prompts in order from the least amount of assistance required to the most amount of assistance until the student is able to perform the behavior independently. In a typical trial sequence, the teacher delivers a task direction and waits a specified response interval (e.g., 3-5 seconds) for the student to respond independently. If an error or no response occurs, the teacher delivers the first prompt and waits the specified response interval for the student to respond. If an error or no response occurs, the teacher delivers the next more intrusive prompt and waits the specified response interval. This sequence continues until the student responds correctly or until the teacher delivers all the prompts in the hierarchy. If the student responds correctly at any time in the sequence, either independently after the task direction or following a prompt, the teacher reinforces the response and records the independent response or the prompt level that resulted in the correct response by the student.

The teacher selects a prompt hierarchy consisting of (a) an independent level, (b) a verbal prompt, (c) a verbal + model prompt, and (d) a verbal + full physical guidance prompt. When presenting the verbal prompt, the teacher tells the student exactly what to do on a step of the task analysis (e.g., "Get your hands wet" for step two of the task analysis). When using a verbal + model prompt she tells him what to do on a step while doing the step herself (e.g., saying "Get your hands wet. Watch me." while wetting her own hands in the water). When presenting the verbal + full physical guidance prompt, she tells the student what to do while providing hand-over-hand instructional prompting to correctly complete the step.

To graph the data, the teacher is using a graphing method that assigns points on a weighted scale based on the prompt level to which the student responds correctly (see attached figure). The points are assigned so that responding to less intrusive prompts receives more points than responding to more intrusive prompts. In this way, the teacher will see an accelerating trend if the student increases his responses to less intrusive prompts as instruction continues over time. In this case, the teacher assigns 3 points for every correct independent response, 2 points for every correct response following a verbal prompt, 1 point for every correct response following a verbal + model prompt, and 0 points for every correct response following a verbal + full physical guidance prompt.
### Task Analysis and Student Responding

<table>
<thead>
<tr>
<th>Step</th>
<th>Task Analysis</th>
<th>Student Responding</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Turn on water</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>Wet hands for 3-5 s</td>
<td>V V V V V I I I</td>
</tr>
<tr>
<td>3.</td>
<td>Place one hand on pump dispenser</td>
<td>PG PG PG PG PG PG PG</td>
</tr>
<tr>
<td>4.</td>
<td>Place other hand below soap spout</td>
<td>PG PG PG PG PG PG PG M</td>
</tr>
<tr>
<td>5.</td>
<td>Push pump 1 time</td>
<td>PG PG PG M M PG M M</td>
</tr>
<tr>
<td>6.</td>
<td>Rub hands together 5-10 s</td>
<td>PG PG M M V V V V</td>
</tr>
<tr>
<td>7.</td>
<td>Rinse hands 5-10 s</td>
<td>+ M M PG PG M PG M V</td>
</tr>
<tr>
<td>8.</td>
<td>Turn off water</td>
<td>PG PG M V V M V V V</td>
</tr>
<tr>
<td>9.</td>
<td>Touch sensor on paper towel dispenser</td>
<td>PG PG PG PG PG PG PG M V</td>
</tr>
<tr>
<td>10.</td>
<td>Tear off paper towel</td>
<td>M M I I M M V V</td>
</tr>
<tr>
<td>11.</td>
<td>Dry palms of hands</td>
<td>PG PG PG PG M M M M</td>
</tr>
<tr>
<td>12.</td>
<td>Dry tops of hands</td>
<td>PG PG PG PG PG M PG M</td>
</tr>
<tr>
<td>13.</td>
<td>Throw paper towel in trash can</td>
<td></td>
</tr>
</tbody>
</table>

**Date**
- 1/2 1/3 1/4 1/5 1/6 1/7 1/10 1/11 1/12 1/13 1/14

<table>
<thead>
<tr>
<th>Session length (in min)</th>
<th>3 3 3 7 7 6 7 5 5 6 5</th>
</tr>
</thead>
<tbody>
<tr>
<td># of I points</td>
<td>3 3 0 3 3 6 6 3 6 6 6</td>
</tr>
<tr>
<td># of V points</td>
<td>2 2 2 4 8 2 8 10</td>
</tr>
<tr>
<td># of M points</td>
<td>2 2 3 3 4 5 4 5</td>
</tr>
<tr>
<td># of PG points</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>Total session points</td>
<td>3 3 0 7 7 11 13 15 13 18 21</td>
</tr>
</tbody>
</table>

**Note.** I = Independent (3 pts.); V = Verbal (2 pts.); M = Verbal + model (1 pt.); PG = Verbal + physical guidance (0 pts.); SLP = System of least prompts.
APPENDIX C

LEAST-TO-MOST PROMPTING MODULE
Example: Snapshot of first two steps

Step 1. Task direction (once only at the beginning of the session) to “wash hands”. Client must turn on water

You

Wash your hands (task direction)

No response

Client

Turn on water (verbal prompt)

Turns on water

Praise correct response

If client responds (after verbal prompt on left), praise correct responding and move on to the next step (e.g., wet hands); if not, move on to the next prompt (muscle) if the client still does not respond, move on to the next most intrusive prompt. (physical guidance). This is the key for the first step. This is how you record the data. Somewhere (40) prompt as V, which is worth 1 point.

<table>
<thead>
<tr>
<th>Day</th>
<th>Task Analysis</th>
<th>Patient Responding</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Turn on water</td>
<td>---</td>
</tr>
<tr>
<td>2</td>
<td>Wash hands for 5-10 s</td>
<td>---</td>
</tr>
<tr>
<td>3</td>
<td>Place one hand on soap dispenser</td>
<td>---</td>
</tr>
<tr>
<td>4</td>
<td>Place other hand below soap spout</td>
<td>---</td>
</tr>
<tr>
<td>5</td>
<td>Push pump 1 time</td>
<td>---</td>
</tr>
<tr>
<td>6</td>
<td>Rub hands together for 10 s</td>
<td>---</td>
</tr>
<tr>
<td>7</td>
<td>Rinse hands 5-10 s</td>
<td>---</td>
</tr>
<tr>
<td>8</td>
<td>Turn off water</td>
<td>---</td>
</tr>
<tr>
<td>9</td>
<td>Wash between paper towel</td>
<td>---</td>
</tr>
<tr>
<td>10</td>
<td>Dry hands</td>
<td>---</td>
</tr>
<tr>
<td>11</td>
<td>Dry hands</td>
<td>---</td>
</tr>
<tr>
<td>12</td>
<td>Dry hands</td>
<td>---</td>
</tr>
<tr>
<td>13</td>
<td>Turn off water in fruit can</td>
<td>---</td>
</tr>
</tbody>
</table>
Assessment Criteria for PS

Name of observer: Video tape label: 
Date: 

1 = correct 
0 = incorrect

If one of the five target behaviors is incorrect, the entire trial should be counted as incorrect. Mark NA in grey box if not applicable during a trial. Numbers with * indicate the trial should be represented.

<table>
<thead>
<tr>
<th>Trials (simulated sheet A)-PS</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of stimuli on table (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance of stimuli (1 ft)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Removed unselected stimulus before collecting data</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Blocked response</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Ended trial within 5 secs. of no response</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
</tbody>
</table>

**Blocking Correct:** Score correct if p’s hands move towards the client’s hands when the client attempted to select more than one item or an item that was not in the predetermined pool of stimuli.
<table>
<thead>
<tr>
<th>Trial</th>
<th>Left position</th>
<th>Right position</th>
<th>Item Selected</th>
<th>Directions for Simulated Client (A)-PS</th>
<th>Correct for role-played client</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>Left</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>Right</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>6</td>
<td>NR</td>
<td>No response</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>Right</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>3</td>
<td></td>
<td>Slowly both</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>Right</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>7</td>
<td></td>
<td>Left</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>1</td>
<td>NR</td>
<td>No response</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>3</td>
<td></td>
<td>Quickly both</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>4</td>
<td></td>
<td>Not on table</td>
<td></td>
</tr>
</tbody>
</table>

PS Assessment: Correct = 1; Incorrect = 0
APPENDIX F

WRITTEN INSTRUCTIONS FOR THE PS ASSESSMENT
PS Preference Assessment

Place two stimuli approximately half-way between yourself and the consumer, and 1 ft apart. Verbally prompt the consumer to select a stimulus. If the consumer selects a stimulus within 5 s of the verbal prompt, remove the stimulus not selected and record the name of the selected stimulus on your sheet. If no selection response occurs within 5 s, remove both stimuli, record “no response” for that trial, and prepare stimuli for the next trial. If the consumer attempts to select both stimuli at the same time, block the response and repeat the trial. You also need to calculate selection percentages and identify the high-preference item following the assessment.
APPENDIX G

PREFERENCE ASSESSMENT MODULE
6. If the client attempts to:
   a. Take both items at once or
   b. An item not placed on the table by you:
      - Try to gently block one or both of their hands. If the client does take both items, do not record data on that trial, but present the same items again.

7. Continue in this manner until all trials have been completed.

Hint: Block their hands by reaching your hands towards the client's hands to stop them.
APPENDIX H

PROCEDURAL FIDELITY CHECKLIST
<table>
<thead>
<tr>
<th>Direction</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduced the sim consumer</td>
<td>□ Yes □ No</td>
</tr>
<tr>
<td>Told participant that simulated client will tell them when to start and stop (before first trial only)</td>
<td>□ Yes □ No</td>
</tr>
<tr>
<td>Told participant that simulated client cannot answer questions (before first trial only)</td>
<td>□ Yes □ No</td>
</tr>
<tr>
<td>Told participant that they were to implement the prescribed procedure before each session</td>
<td>□ Yes □ No</td>
</tr>
<tr>
<td>Provided the data sheet (or allowed continued use of sheet) and other required material for that session</td>
<td>□ Yes □ No</td>
</tr>
<tr>
<td>Explained that the experimenter cannot answer any questions related to the instructions until the debriefing period</td>
<td>□ Yes □ No</td>
</tr>
<tr>
<td>Told participant that participant can refer to instructions whenever they would like (before first trial only)</td>
<td>□ Yes □ No</td>
</tr>
<tr>
<td>Gave the participant the module and recorded the duration of review</td>
<td>□ Yes □ No</td>
</tr>
<tr>
<td>Recorded the number of times participant referred to instructions during implementation</td>
<td>□ Yes □ No</td>
</tr>
<tr>
<td>Reset the room after each session</td>
<td>□ Yes □ No</td>
</tr>
<tr>
<td>Gave the participant hypothetical data and calculator</td>
<td>□ Yes □ No</td>
</tr>
<tr>
<td>• Told the participant the following information:</td>
<td>□ Yes □ No</td>
</tr>
<tr>
<td>□ The participant has 20 minutes to complete instructions</td>
<td>(to meet this criterion must meet all tasks in category)</td>
</tr>
<tr>
<td>□ Experimenter cannot answer any questions related to the instructions until the debriefing period</td>
<td></td>
</tr>
<tr>
<td>□ If the participant finishes early, they should get the experimenter</td>
<td></td>
</tr>
<tr>
<td>Reviewed the debriefing form</td>
<td>□ Yes □ No</td>
</tr>
<tr>
<td>Percentage = ________/13</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX I

SOCIAL VALIDITY QUESTIONNAIRES
Brief Questionnaire (least-to-most prompting strategies)

On a scale of 1-5 with 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, and 5 = strongly agree, please rate the following statements.

• Using least-to-most prompting procedures to teach a skill is important
  1 2 3 4 5

• This procedure is an appropriate way to assist a client with learning new skills
  1 2 3 4 5

• An instructional module can be used when training clinicians, teachers, and/or caregivers to implement least-to-most prompting strategies with children.
  1 2 3 4 5

• Providing an instructional module is an effective way to train clinicians, teachers, and/or caregivers to implement least-to-most prompting strategies with children.
  1 2 3 4 5

• I feel confident that I can correctly implement least-to-most prompting strategies with children.
  1 2 3 4 5

• I would recommend the use of self-instructional modules when teaching novice individuals these skills.
  1 2 3 4 5

• The module was easy to follow.
  1 2 3 4 5

• The most helpful part(s) of the self-instructional module was/were the following:

• The least helpful part(s) of the self-instructional module was/were the following
Brief Questionnaire (PS assessment)

On a scale of 1-5 with 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, and 5 = strongly agree, please rate the following statements.

• Using a preference assessment to identify a client’s preference is important
  1 2 3 4 5

• Conducting the assessment is an appropriate way to identify items that would motivate a client.
  1 2 3 4 5

• An instructional module can be used when training clinicians, teachers, and/or caregivers to conduct assessments with children.
  1 2 3 4 5

• Providing an instructional module is an effective way to train clinicians, teachers, and/or caregivers to conduct assessments with children.
  1 2 3 4 5

• I feel confident that I can correctly conduct a paired stimulus preference assessment.
  1 2 3 4 5

• I would recommend the use of self-instructional modules when teaching novice individuals these skills.
  1 2 3 4 5

• The module was easy to follow.
  1 2 3 4 5

• The most helpful part(s) of the self-instructional module was/were the following:

• The least helpful part(s) of the self-instructional module was/were the following