COMPETENCY-BASED SUPERVISION IN ABA

Competency-Based Clinical Supervision in Applied Behavior Analysis

by
Yulema Cruz

An Applied Dissertation Submitted to the
Abraham S. Fischler College of Education
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Education

Nova Southeastern University
July 10, 2019
Acknowledgements

This work would not have been possible without the support, guidance, and commitment of my dissertation committee. Dr. Barbara Garcia-Lavin, thank you for teaching me to engage in the exercise of academic writing for diverse audiences. I am a better writer and critical thinker for it. Dr. Dana Mills, thank you for teaching me to broaden my research horizons. Your contributions helped shape me into a better consumer of applied research beyond the field of behavior analysis. Most remarkably, Dr. Jon Bailey, thank you for teaching me the magic of ABA research. Your unwavering commitment to this work, as well as inspiration and mentorship throughout this process, have shaped the researcher I am today. I will remain a student of behavior analysis research, rooted on your teachings.

To my research team, especially Dr. Tom Welsh: this project was improved by your insight and dedication to the study of behavior. Thank you for allowing me to learn from you; I will forever be grateful for your wisdom and encouragement. To the rest of the research team: George, Kitty, Rainy, Jay, Melody, Anna, Sara, Kaitlin, Alyssa, Gabe, Dr. Thyer, Hana, Karly, Tania, and others whose names escape me at the moment: you were the village it took to make this happen. Thank you for welcoming me into yours. Jenna Kriss and Tania Duarte, your contribution to this research surpassed countless hours of data collection. Thanks in great part to your dedication, the study’s questions were empirically answered; your work helped establish a basis for the conceptualization and analysis of supervisory behavior in our field. Dr. Karly Cordova, thank you for being my beacon of light. Your bravery, commitment to ethics, outstanding values, professionalism, and compassion helped to inspire this research idea.

I owe special thanks to Dr. Tara Loughrey and Dr. Jessica Dunn, who graciously supported the completion of this study at The Victory Center for Autism and Related Disabilities. I am also grateful for the kindness and enthusiasm shared by Beatriz Kourakles from Monarch Center for Autism, as well as Alina Slaterpryce, Tiffany Lomas, Sandra Merida, and the staff from Project Brilliance. Thank you for allowing me to bring this project to life. This study was first piloted at Behavior Analysis Inc.; I offer my sincere gratitude to Dr. Steve Starin, Silvia Starin, Dr. David Garcia, Kristine Cardenas, and Raquel Aidie for trusting me to try a “promising idea”.

Nobody’s support has meant more to me than my family’s. Mother, you are the original family researcher. Thank you for always fostering a culture of professionalism and intellectual curiosity. Thanks to you, I will continue to be a student of my profession. Last but not certainly not least, I owe infinite thanks to my husband Fernando Carpinatti, without whose unconditional love and support the successful completion of this study would not have been possible. We did it!
COMPETENCY-BASED SUPERVISION IN ABA

Approval Page

This applied dissertation was submitted by Yulema Cruz under the direction of the persons listed below. It was submitted to the Abraham S. Fischler College of Education and School of Criminal Justice and approved in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Education at Nova Southeastern University.

Barbara Garcia-Lavin, Ph.D.
Committee Chair

Dana Scott Mills, Ph.D.
Committee Member

Jen S. Bailey, Ph.D., BCBA-D
Committee Member

Kimberly Durham, Psy.D.
Dean
Statement of Original Work

I declared the following:

I have read the Code of Student Conduct and Academic Responsibility as described in the Student Handbook of Nova Southeastern University. This applied dissertation represents my original work, except where I have acknowledged the ideas, words, or material of other authors.

Where another author’s ideas have been presented in this applied dissertation, I have acknowledged the author’s ideas by citing them in the required style.

Where another author’s words have been presented in this applied dissertation, I have acknowledged the author’s words by using appropriate quotation devices and citations in the required style.

I have obtained permission from the author or publisher—in accordance with the required guidelines—to include any copyrighted material (e.g., tables, figures, survey instruments, large portions of text) in this applied dissertation manuscript.

Yulema Cruz, MS, BCBA
Name

July 10, 2019
Date
COMPETENCY-BASED SUPERVISION IN ABA

Abstract

Competency-Based Clinical Supervision in Applied Behavior Analysis. Yulema Cruz, 2019: Applied Dissertation, Nova Southeastern University, Abraham S. Fischler College of Education and School of Criminal Justice. Keywords: supervision, clinical supervision, competency-based supervision, Applied Behavior Analysis.

This applied dissertation evaluated a new method that provided a systematic way of training and evaluating supervisors in Applied Behavior Analysis (ABA). Using a single-case multiple-baseline across supervisors and supervisees research design, this study assessed the use of a modified behavioral skills protocol (BST) in the training of master’s and doctoral-level Board Certified Behavior Analysts (BCBAs/BCBA-Ds) in the supervision of DTT. Additionally, this study evaluated supervisee’s performances before and during the implementation of the modified BST protocol. Included in the new supervisor training curriculum outline (2.0) released by the BACB in 2019, BST is an evidence-based protocol whose efficacy has been primarily established in the area of staff training. Results revealed that the use of the new modified BST protocol for the training of BCBAs on the supervision of DTT had a positive effect on supervisors’ and supervisees’ performances, when compared to baseline values.
## Table of Contents

Chapter 1: Introduction .................................................................................................................. 1  
  Statement of the Problem ........................................................................................................... 1  
  The Topic .................................................................................................................................... 2  
  The Research Problem ................................................................................................................ 3  
  Background and Justification ....................................................................................................... 5  
  Deficiencies in the Evidence Base ............................................................................................... 8  
  Audience ...................................................................................................................................... 9  
  Setting of the Study ..................................................................................................................... 9  
  Materials ..................................................................................................................................... 11  
  Researcher’s Role ......................................................................................................................... 12  
  Purpose of the Study .................................................................................................................... 12  
  Definitions and Terms .................................................................................................................. 14  

Chapter 2: Literature Review ......................................................................................................... 21  
  Multisystem Therapy Supervision .............................................................................................. 22  
  Supervision of School Psychologists ......................................................................................... 24  
  Developmental Approach to Supervision ..................................................................................... 25  
  Competency-based Supervision Tool .......................................................................................... 27  
  Development of APA Guidelines for Supervision ....................................................................... 28  
  Competency-based Supervision .................................................................................................... 28  
  Attachment Informed Approach to Supervision .......................................................................... 31  
  Couple and Family Therapy Supervision ...................................................................................... 32  
  Clinical Neuropsychology ........................................................................................................... 33  
  Competency-based Training in Medicine .................................................................................... 34  
  Internationalization of Competency-based Supervision ............................................................. 40  
  Evidenced-based Approach to Supervision .................................................................................... 41  
  Competency Benchmarks for Professional Psychologist Supervisors ......................................... 43  
  Supervision in Play Therapy ........................................................................................................ 43  
  Early Behavioral Research on Supervisory Behavior .................................................................... 44  
  Using behavioral Methods to Improve Supervisor Behavior ....................................................... 45  
  Changing Supervisees’ Performances Through Behavioral Training ......................................... 46  
  Recent Research on Supervision for Behavior Analysts ............................................................... 47  
  BACB Requirements for Supervisors ........................................................................................... 52  
  Fidelity and Treatment Adherence ............................................................................................... 53  
  BST: An Evidence-based Method .................................................................................................. 58  
  Research Questions ...................................................................................................................... 63  
  Hypotheses ................................................................................................................................... 64
Chapter 1: Introduction

Statement of the Problem

Throughout the history of the human-services related disciplines, supervision has been considered the primary method of training and monitoring of clinicians. Traditionally, the attainment of skills competency in clinical supervision has been an essential part of clinician training, mainly in the health care-related professions (Falender & Shafranske, 2012; Lehman, Goldman, Dixon, & Churchill, 2004, as cited in Leffler, Jackson, West, McCarthy & Atkins, 2012). In similar fashion, the phenomenon of clinical competence in supervision has extended to other human services-related fields such as psychology, education, mental health counseling, medicine, and Applied Behavior Analysis (ABA; Behavior Analyst Certification Board, 2012; Falender & Shafranske, 2012; Milne & Reiser, 2012).

Bernard and Goodyear (2004) as cited in Rings, Genuchi, Hall, Angelo, & Erickson Cornish (2009) defined supervision as:

“A method of evaluative intervention provided by a senior member of a particular profession, a supervisor, to one or more junior members within the same profession. A supervisor works to enhance the professional functioning of his or her supervisees and ensures that they provide quality professional services to their clientele. In kind, supervisors ultimately act as gatekeepers for those entering their respective professions” (p. 140).

Expressing the definition in medical terms, clinical competence is “the habitual and judicious use of communication, knowledge, technical skills, clinical reasoning, emotions, values, and reflection in daily practice for the benefit of the individual and community being
COMPETENCY-BASED SUPERVISION IN ABA

served” (Epstein & Hundert, 2002, p. 226). Furthermore, “clinical competence depends on habits of mind, including attentiveness, critical curiosity, self-awareness, and presence” (Kaslow, Falender, & Grus, 2012, p. 48).

The topic. In the field of Applied Behavior Analysis (ABA), competence in supervision is considered an important practice due to its principal role in the attainment of certification, as well as its role in the service-delivery tier system that has historically been exercised by behavior analysts (BACB, 2012). This service-delivery tier system may consist of a Board Certified Behavior Analyst (BCBA) or Board Certified Behavior Analyst – Doctoral (BCBA-D), who may (or may not) be supported by a Board Certified Assistant Behavior Analyst (BCaBA), overseeing a treatment team of Registered Behavior Technicians (RBTs) and/or therapists (BACB, 2014). This model relies on the use of Assistant Behavior Analysts and Behavior Technicians as the primary mechanism for achieving behavioral improvements (BACB, 2014). Therefore, it is the responsibility of BCBAs and/or BCBA-Ds to efficiently train and supervise Assistant Behavior Analysts and Behavior Technicians (BACB, 2014). For this and other reasons, the Behavior Analyst Certification Board (BACB) defined supervision as an activity intended to “improve and maintain the behavior-analytic, professional, and ethical repertoires of the supervisee and facilitate the delivery of high-quality services to his/her clients” (BACB, 2012, p. 3). However, since its inception in 1998, the BACB had not addressed the topic of competency-based supervision until the establishment of the supervision task force in 2012 (BACB, 2012).

Based on a broad review of the literature, a possible reason why the BACB may not have addressed the topic of supervision until recent years may be the heavy focus that ABA professionals have placed on clinical practice. Historically, behavior analysts appear to have concentrated their research efforts on changing the behaviors of their clients, by means of either
COMPETENCY-BASED SUPERVISION IN ABA

skill acquisition, or behavior reduction procedures. Additionally, therapists’ skills development in community settings has often followed the “train and hope” approach, and clinical supervision has typically focused on administrative requirements and providing practitioners with social support (Bickman, 1999 as cited in Henggeler, Schoenwald, Liao, Letourneau, & Edwards, 2002). Nevertheless, the lack of research on supervision practices in community settings is most likely due to the significant challenges in conducting such studies (Henggeler, et al., 2002). To illustrate, a search of 24 databases containing the words supervision and applied behavior analysis revealed only 12 articles focused on supervision, none of which belonged to an ABA peer-reviewed journal. This suggests that the literature on supervision in ABA is currently very limited. The heavy emphasis on the clinical aspect of the ABA field may have taken away focus from the development of quality control mechanisms, as the field as a whole, did not focus on the behaviors of the professionals providing clinical services until recent years.

The Research Problem. Over the past decade and a half, the field of psychology has witnessed a shift towards accountability in the provision of clinical supervision (Watkins, 2012). In 2005, Stoltenberg affirmed, “there has been a recent increase in interest in defining and describing the competencies for professional practice in psychology” (p. 857). Likewise, an analysis of the supervision literature provided support for a continuing trend toward the use of competency and evidence-based methods of supervision (Falender & Shafranske, 2012).

An appraisal of the literature in the mental health-related fields yielded at least six qualitative philosophical research studies on supervision (Welsh, Stanley, & Wilmoth, 2003; Barret & Barber, 2005; Stoltenberg, 2005; Falender & Shafranske, 2007; Celano, Smith, & Kaslow, 2010; Milne & Reiser, 2012; and Leffler, Jackson, West, McCarty, & Atkins, 2012), one supervision tool (Madson, Campbell, Barrett, Brondino, & Melchert, 2005), one report on
COMPETENCY-BASED SUPERVISION IN ABA

competence in professional psychology with guidelines for the ethical provision of clinical supervision (Hatcher et al., 2013), one article focusing on explaining the theoretical underpinnings and key features of competency-based models (Gonsalvez & Calvert, 2014), one competency-based supervision model coupled with a psychotherapy-based approach (Stein, Clair, Soenksen-Bassett, Martin, & Clarke, 2015), one model of supervision training for social work field instructors (Bennett, 2008), one study providing a framework and guidance for the development of competency standards for training of neuropsychological supervisors (Stucky, Bush, & Donders, 2010), one study that developed an international competency checklist (Ng, Choudhuri, Noonan, & Ceballos, 2012), as well as literature review articles (Hudspeth, 2015). Additionally, the review yielded six articles on competency-based training in medicine. Three studies focused on the development of competency training programs/questionnaires (Yuen, Barrington, Headford, McNulty, & Smith, 1998; Pryor, et al., 2006; and Rubulotta, Gullo, Iapichino, Pezzi, Bion, Barret, & the CoBaTrICE Italian Collaborative, 2009), two others were descriptive articles (Emmett, & Green-McKenzie, 2001; and Louie, Coverdale, & Weiss Roberts, 2004), and one encompassed a literature review (Basnet, Clapham, Shakya, & McCall, 2004).

On the behavioral side, only three studies empirically evaluated supervisors’ behaviors (Rosenbaum, 1975; Ivancic, Reid, Iwata, Faw, & Page, 1981; and Methot, Williams, Cummings, & Bradshaw, 1996), two studies focused on empirically evaluating supervisee’s performances (Dillon, Kent, & Malott, 1980; and Hundert & Hopkins, 1992), and only one research article addressed recommended practices for supervisors when engaging in the supervision of students of behavior analysis (Sellers, Valentino & LeBlanc, 2016). The aforementioned article was largely based on recent criteria developed by the BACB, which requires BCBAs to master a number of specific competencies before engaging in the practice of supervision. However, as
COMPETENCY-BASED SUPERVISION IN ABA

previously mentioned, these criteria can be easily met by successfully by completing an 8-hour training in supervision based on a BACB-developed supervision curriculum and passing a 90-minute training on the BACB website regarding guidelines for supervised experience. At the time this manuscript was completed, accomplishment of the latter requirement (i.e., passing a 90-minute training) was no longer a condition for supervision. Overall, the existing research literature examined supported the need for empirically studying the effectiveness of supervision. However, the findings lacked several important attributes. They include: (a) live demonstrations of specific level of supervisor competence, (b) live demonstrations of supervisory competence at a level commensurate with best practices’ standards, (c) empirical methodology of assessing and monitoring supervisor performance, (d) reliance on observation data collection and analysis as a means of assessing supervisee performance, (e) any empirical validation, (f) standardized methods, and (d) controlled research design. Conversely, the proposed study sought to fill the current gap in the literature by examining live demonstrations of supervisor and supervisee competency, via direct observations and data collection. Additionally, the study proposed the implementation of an empirical method for assessing and training supervisor performance, within a highly-controlled research design. Literature on this empirical method selected for the examination of supervisor behavior (i.e., BST), will be discussed at the end of chapter two.

**Background and justification.** The field of ABA has recently begun to define a set of skills and knowledge base necessary for the effective supervision of professionals who deliver behavior-analytic services, as well as for prospective certificants (BACB, 2012). However, a definition and subsequent acquisition of supervisory skills was slow to arrive, as the earliest description of attainment of some type of supervisory competency was not addressed by the BACB until 2012 (BACB, 2012). In the interim, supervisors’ qualifications remained the same
as non-supervisors until new BACB requirements came into effect in January of 2015. Previous supervision requirements included the possession of a Board-Certified Behavior Analyst (BCBA) or Board-Certified Behavior Analyst – Doctoral (BCBA-D) certification, as well as maintaining “good standing” professional status (BACB, 2012). The 2015 requirements included two additional requisites: successful completion of an eight-hour supervision training workshop and passing a 90-minute BACB-issued training module on supervision standards (BACB, 2012). At the time this manuscript was completed, passing the 90-minute training was no longer a requirement for supervisors. Instead, the BACB on its October 2017 newsletter announced changes to supervisor qualifications that will take effect on January 2022 and stated that newly certified BCBAs will not be allowed to supervise trainees within their first-year post-certification (BACB, 2017). Junior BCBAs may therefore supervise trainees “only if they receive monthly consultation throughout that year from a BCBA with at least 5 years of post-certification experience in supervising trainees” (p. 3). That same year, the BACB revised the requirements for supervisors of trainees to include asides from the 8-hour training and appropriate credentials, the procurement of three hours of continuing education credits in the area of supervision during each two-year certification cycle (BACB, 2019). Furthermore, in February of 2019 the BACB also published a new Supervisor Training Curriculum Outline (2.0) which will take effect on November 1, 2019 (BACB, 2019). While considered significant steps in the right direction, these supervisor requirements presently allow any junior certificant in possession of the appropriate credentials and completion of the new training requirements to provide clinical supervision services and receive compensation for it, but do not require the certificant to demonstrate any specific level of supervision competence (BACB, 2012).
COMPETENCY-BASED SUPERVISION IN ABA

To date, neither the literature on supervision in ABA, current graduate school curriculum content, nor the BACB have offered clinicians a systematic method for training BCBAs (or prospective BCBAs), as most of the content has thus far focused on changing the behaviors of supervisees, while little attention has been placed on the development of competent BCBA supervisors. The absence of a systematic method to train professional behavior analysts in acquiring specified clinical supervision competencies and subsequently asses their level of supervision competency may ultimately affect the welfare of consumers of behavior analytic services, as inexperienced supervisors may lack accountability for their skills and readiness to take on the mentorship of a trainee (Falender & Shafranske, 2012). “If these professionals make their competency expectations and successful assessment approaches more transparent, they can be more easily shared and will contribute to the new culture” (Roberts et al., 2005, p. 359).

Behavioral Skills Training (BST) appears to be an encouraging method for training supervision competencies. The research literature on BST suggests this protocol as a promising method for training supervision competency in ABA, due to its was demonstrated effectiveness in other behavioral skills training areas. Specifically, BST was included in the supervisor-training curriculum issued by the BACB in 2012, as well as the new Supervisor Training Curriculum Outline (2.0) (BACB, 2019) as an evidence-based protocol whose efficacy was first established in the area of staff training (Miltenberger & Fuqua, 1985; Miltenberger & Veltum, 1988; Bakken, Miltenberger & Schauss, 1993). In addition, BST has been successfully utilized in the teaching of critical skills to persons with intellectual disabilities (Lee & Tang, 1998; Miltenberger, 2000; Egemo-Helm, Miltenberger, Knudson, Finstrom, Jostad, & Johnson, 2007; Burke, Andersen, Bowen, Howard, & Allen, 2010), as well as the teaching of personal and gun safety, plus abduction prevention skills to children (Miltenberger & Thiesse-Duffy, 1988;
COMPETENCY-BASED SUPERVISION IN ABA


Deficiencies in the Evidence Base. Although most disciplines including psychology, ABA and counseling are making the shift towards competency-based supervision, the lack of empirically validated protocols that describe what effective supervision training looks like continues to prevail (Falender & Shafranske, 2012). Additionally, substantial controversy still exists surrounding the concept of core competencies, with supporters highlighting a better definition of performance criteria as a key requirement (Hatcher et al., 2013; Summerall et al., 2000, as cited in Falender et al., 2004), “greater articulation between instruction, goals, and assessment, ensuring breadth of training in an era of specialization, and provision of core identity for psychologists” (Benjamin, 2001; Kaslow, 2002, as cited in Falender et al, 2004 p. 782).

While performance criteria have been developed for both supervisors and supervisees, they only address the problem from the perspective of what competencies to master. Attempting to tackle the problem from the perspective of how to attain competency in supervision is yet to be examined in the ABA literature.

As previously mentioned, the field of ABA has recently begun to address the topic of competency in supervision. Nevertheless, though related topics such as staff monitoring and training have been reasonably researched, no study to date has addressed the topic of competency-based supervision in the field. The lack of a body of research in this area presents an opportunity to develop a scientifically validated teaching model for the attainment of clinical supervision competency. Comparable to other disciplines within psychology, a teaching model
COMPETENCY-BASED SUPERVISION IN ABA

for competency in supervision should include clearly articulating instructions or goals, demonstrating desired behavior, providing opportunities for role play and practice, effective feedback, and assessing the supervisee’s behavior (BACB, 2012), among other components.

**Audience.** The primary beneficiaries of this study were BCBA supervisors of DTT sessions, for they were provided with a detailed description of the ABA supervisory process as it pertained to DTT, a clear definition of relevant ABA supervisory activities, as well as a means of identifying critical aspects in ABA supervision. In addition, the study allowed BCBA supervisors to identify and engage in effective supervisory practices, along with pinpoint challenging problems encountered in ABA supervision. Supervisee participants benefited from a description of problem-solving strategies during the ABA supervision process that included clear instructions, role play, and corrective plus positive feedback. With respect to their performance during DTT sessions, supervisees benefited from the acquisition of more complex DTT skills, which allowed them to solve specific previously addressed issues, in the moment (i.e., during the DTT session with the child). Finally, supervisees’ awareness of their own performance increased with the introduction of video reviews. Overall, both supervisors and supervisees benefited from the identification of consistent feedback received from each participant.

Though not addressed during this study, it is hypothesized that the above benefits also transferred to child participants. The study generated no unfavorable effects for participants.

**Setting of the study**

The study took place at four local ABA clinics and one ABA school. On two of the settings, two different rooms were utilized in the study. On the other three settings, the same room was utilized for all the sessions. At two of the clinics, the supervisee, BCBA and Uber supervisor used the first room, where individual supervision and training sessions occurred. The
second room, the therapy room, was used for the DTT sessions, and accommodated the supervisee and client. During baseline, the BCBA also was present in the therapy room. Supervision sessions between the Uber supervisor and the BCBA did not take place during baseline, as they were part of the treatment phase. Both phases (baseline and treatment) took place in the same rooms. The therapy rooms’ dimensions typically consisted of an eight-feet by ten-feet enclosure with a small glass window or one-way mirror, equipped with a kid-sized table, two to three kid-sized chairs, and other relevant materials (e.g., toys, books, games, DVDs, and other potential reinforcers, as well as therapy materials such as puzzles, cards, blocks, and data collection sheets, among others).

Although DTT sessions in the field of ABA may also take place in other settings such as home, school, aftercare, and the community, a clinic setting was selected for this study in order to allow for the most controlled environment possible, thereby limiting the effect of potential confounding variables on the behaviors of children, trainers and BCBA supervisors. For example, clinic settings do not include the presence of other family members, the indiscriminate availability of media and other potential unmonitored reinforcers, as well as other unexpected distractors/events. Minimization of distractors prevented the children from engaging in competing behaviors, thereby facilitating supervisees’ ability to demonstrate newly shaped DTT skills. The clinics selected for this study were previously set up and equipped for DTT sessions with cameras and other necessary materials and served as a hub where all participants could meet separately or as a group. This setup facilitated video recordings, reduced the potential for client and trainer reactivity, minimized distractions and changes to the child’s routine, as well as saved valuable time for all parties involved. Overall, the clinic setting provided a standardized setting,
which along with the supervision training materials and intervention approach were designed to facilitate reproducibility in future studies.

**Materials.** The first room was typically equipped with one of more adult-size desks, two or more chairs, a computer that included a monitor and speakers, as well as paper, a pen, and relevant DTT session materials. A digital video camera containing a built-in microphone and memory card was used to record each supervision session. The therapy room was equipped with one or more child-size tables, two or more chairs, a small drawer cabinet containing session materials such as language cards, toys, timers, individual program data collection sheets, language-builder objects, pens, papers, and hand sanitizer among other items, as well as a camera, and the client’s program book. The therapy room also contained potential reinforcers in the form of food, drinks, larger toys, iPads, smart phones, tape players, a trampoline, puzzles, and a two-way window. Potential reinforcers are highly preferred stimuli that are presented contingent upon desired behavior. During the study, these stimuli were always under the control of the supervisees and presented only after desired behaviors were exhibited by the children during DTT sessions. The above referenced digital video camera was also used to record each DTT session. All videotaped session data were stored in a file in the researcher’s password protected computer, at her home.

Intervention materials included (aside from the digital video camera) two main data collection forms: The *Evaluation of Supervisor Performance* (see Appendix C for a copy of this document) and the *Evaluation of Supervisee Performance* form (see Appendix D for a copy of this document). These two data collection forms were completed from the recordings. In addition, one form, the supervisor job aid (see Appendix E for a copy of this document), was introduced during treatment to help the BCBA with the prioritization of feedback. The BCBA
COMPETENCY-BASED SUPERVISION IN ABA

had the option of utilizing this form when providing feedback to the supervisee during supervision sessions. The last form, the *DTT Performance Feedback Sheet* (see Appendix E for a copy of this document) was utilized to measure BCBAs’ fluency with the implementation of DTT as part of their inclusion criteria. Additionally, the BCBA supervisor included relevant therapy room materials when necessary.

**Researcher’s role**

The researcher held no affiliation with the ABA companies in which the study was conducted. Early on in the study, the researcher served as video collector, but this role was soon fulfilled by a research assistant for the remainder of the study. The researcher collected coded data from the primary and reliability observers/data collectors, graphed it and analyzed it weekly. Based on the data provided to the researcher, she implemented the independent variable and made treatment decisions.

**Purpose of the Study**

This study aimed to fill the gap in the ABA supervision literature by empirically examining a systematic method of teaching BCBA supervisors how to engage in the ethical and effective practice of supervision. Due to the wide scope of potential applications of supervisory behavior, this study focused on supervision within Discrete Trials Teaching (DTT), by using a modified version of a promising evidence-based method: BST. A secondary aim of the study was the evaluation of supervisees’ performances. Therefore, the proposed study attempted to provide an evidence-based, systematic method for training and assessing ABA supervisors’ performance on the effective and ethical provision of competency-based clinical supervision. Specifically, the study evaluated the effectiveness of a modified behavioral skills training-based supervision-
COMPETENCY-BASED SUPERVISION IN ABA

training program with ABA supervisors in developing competency and readiness to supervise trainers engaged in DTT.

It was proposed that the use of a modified BST protocol for the training of current certificants would not only have a greater positive effect on supervisor competency, but also on supervisee’s performances, when compared to existing behavioral practices, during baseline. Additionally, this protocol served as the first of its kind competency-based assessment tool for the evaluation of supervisor competency, thereby fulfilling a current gap in the behavioral literature. The use of the modified BST protocol for the assessment and training of supervision behaviors was the main focus of this research.

Requiring supervisor training to advance beyond a basic level of performance allowed BCBA’s to demonstrate supervisory competence at a level commensurate with best practices’ standards. “Competency-based supervision supports an emphasis on formative assessment, ongoing high-quality feedback, and outcomes leading to early identification of competency shortcomings and proactive planning to ensure remediation or constructive action” (Falender & Shafranske, 2012, p. 134). Additionally, it provided supplementary safeguards for ABA consumers “protection of the public from incompetent practice has been both a professional and regulatory matter that has driven the competency movement as well” (Falender & Shafranske, 2012, p. 131).

For the purposes of this study, the researcher defined three levels of supervisor competency. The first or basic level of supervisor performance referred to supervisory skills involving: (a) observing the trainer, (b) identifying and documenting relevant trainer errors during the implementation of DTT, (c) providing corrective feedback, (d) coaching and/or modeling appropriate trainer responses, and (e) providing positive reinforcement. Basic level of
supervisory performance was assessed during the baseline phase of the study. This is the minimal level of supervisor performance currently accepted post certification. The second or intermediate level of supervisor competency comprised the above-mentioned basic level skills plus: (a) prioritizing corrective feedback, (b) prompting note taking, (c) providing opportunities for role play/practice, (d) prompting supervisee to engage in the relevant skills prior to the next DTT session, and (e) scheduling a follow-up appointment. Intermediate level supervisor performance was assessed during the treatment phase of the study. When achieved, this level of supervisor performance was superior to current standards, as BCBAs were evaluated while demonstrating relevant skills. Additionally, the performances of supervisees receiving supervision from intermediate level supervisors were also measured and assessed during live clinical sessions.

The third or advanced level of supervisor performance referred to expected BCBA participants’ performances beyond the modified BST/treatment phase (not assessed in this study). It includes the knowledge, skills and abilities described above under supervisor intermediate-level of performance, plus the demonstration of appropriate social (soft) skills. Appropriate soft skills may include a more in-depth operationalization of corrective feedback, as well as appropriate communication skills, social skills, problem solving skills, ethical applications, among others (Michnick Golinkoff, & Hirsh-Pasek, 2016).

Definitions and Terms

The following definitions were used to evaluate the performances of BCBA supervisor participants. The performance levels described below refer to a combination of knowledge, skills and abilities displayed by supervisors of ABA services.
COMPETENCY-BASED SUPERVISION IN ABA

Supervisor basic-level of performance: Refers to baseline performances displayed by BCBA participants. It includes prior knowledge of the implementation of DTT, as well as the following skills and abilities: a) observed trainer, b) identified and documented relevant trainer errors, c) provided corrective feedback, d) coached and/or modeled appropriate trainer responses, and e) provided positive reinforcement. The aforementioned five skills and abilities are further defined below. For this study, basic-level of supervisory performance was marked by scores of 50 percent or less on the Evaluation of Supervisor Performance form (see Appendix C for a copy of this document) during the baseline condition.

Supervisor intermediate-level of performance: Refers to expected BCBA participants’ performances during the modified BST phase. It includes the knowledge, skills and abilities described above under supervisor basic-level of performance, plus the following acquired skills and abilities: f) prioritized corrective feedback, g) prompted note taking, h) provided opportunities for role play/practice, i) prompted trainer re: relevant skills, and j) scheduled follow-up appointment. The abovementioned five skills and abilities were further defined below. For this study, intermediate-level of supervisory performance was denoted by scores of 80 percent or more on the Evaluation of Supervisor Performance form (see Appendix C for a copy of this document) during the Modified BST condition.

Supervisor advanced-level of performance: Refers to expected BCBA participants’ performances beyond the modified BST phase (not assessed in this study). It includes the knowledge, skills and abilities described above under supervisor intermediate-level of performance, plus the demonstration of appropriate social (soft) skills. Appropriate soft skills were not defined on this study and will be part of the researcher’s recommendations for future investigations.
The following terms were used for evaluating the performance of the supervisor. The corresponding data collection sheet encompassing the dependent measures below was named *Evaluation of Supervisor Performance*.

**Discrete trials teaching (DTT):** First researched by Lovaas (1987), any given discrete trial consists of three basic components: teacher’s instruction, pupil’s target response, and teacher-delivered consequence. Additionally, a discrete trial may also include a prompt between the instruction and response. For example, a teacher may instruct her pupil to “touch his nose”, the pupil touches his nose, and receives praise (e.g., “good job touching your nose”) from the teacher.

**Observed trainer:** The supervisor observed the trainer for a period of at least 30 minutes during the trainer’s DTT session. This may be scored when the supervisor records and talks about having watched a videotape of the previous DTT session observation and has taken notes regarding the trainer’s behavior during the session. Additionally, this may be scored when supervisor is observed to be present during the trainer’s DTT session.

**Identified and documented relevant trainer’s errors:** The supervisor documents on paper (or other means in lieu of paper) the trainer’s errors demonstrated during the previously observed DTT session. During treatment, the supervisor should document the trainer’s errors on the *DTT Performance Feedback Sheet*.

**Prioritized corrective feedback:** After the supervisor has documented all the trainer’s errors, he or she selected a few errors to address first with the supervisee. The supervisor should express something along the lines of: “although these (insert number of errors) errors were identified, we will be concentrating on X Y and Z errors first”. During treatment, the supervisor should be using the *DTT Performance Feedback Sheet* and concentrating on category 1.
Prompted note taking: The supervisor should prompt the trainer to take notes during the supervision session, if the trainer is not prepared to do so already. At this point, the supervisor may offer pen and paper, or prompt the trainer to take notes via other means in lieu of paper.

Provided corrective feedback: During baseline, the supervisor provides corrective feedback regarding the supervisee’s performance. During treatment, the supervisor provides corrective feedback regarding the supervisee’s performance specifically, on the skills previously prioritized. During this period (treatment), the supervisor refers back to the videotape and points out to the trainer the instances in which he/she engaged in incorrect behavior.

Coached and/or modeled appropriate trainer responses: The supervisor should tell and/or demonstrate to the trainer the correct way in which to engage in the previously identified behavior/errors. This should be done live, during the supervision session, and repeated should the trainer have any questions/concerns. For example, the supervisor may tell the trainer which are the appropriate mands to target during a particular session, and/or may demonstrate the appropriate way to contrive a mand if the supervisee continues to ask, “what do you want”?

Provided opportunities for role play/practice: The supervisor should allow the supervisee to practice engaging in the correct way to perform the skills previously identified as errors. This entails an exchange where the supervisor demonstrates the skill(s) (previous step), and then the supervisee imitates the supervisor by performing the skill(s) in the manner in which the supervisor previously demonstrated. This should be done for every error-skill that was identified/addressed during supervision. The child should not be present during role play.

Provided positive reinforcement: The supervisor should provide positive reinforcement immediately after the trainer engages in the correct response during role play in the form of praise, or acknowledgement of the trainer’s correct performance for at least 60% of the
opportunities. Additionally, the supervisor may also provide positive reinforcement for previously identified trainer’s correct responses (something the trainer did well that was not identified as an error).

**Prompted trainer re: relevant skills:** If applicable, the supervisor reminds the trainer to engage in the relevant DTT skills that they discussed earlier in the session. This is typically done at the end of the supervisor’s feedback session/the onset of the trainer’s DTT session. It may be done in the form of a question where the supervisor asks the trainer whether he/she remembers the specific skills that were coached/role played during the last supervision session (e.g., do you remember the two skills you are focusing on today?).

**Scheduled follow-up appointment:** The supervisor should indicate to the supervisee when the next supervision meeting will take place. This includes date and time of supervision session.

**Prompt:** Supplementary antecedent stimuli that are used to occasion accurate responses in the presence of discriminative stimuli, that ultimately control the behavior (Cooper, Heron & Heward, 2007).

The following definitions were used for evaluating the performance of the supervisee. The corresponding data collection sheet encompassing the dependent measures below was named *Evaluation of Supervisee Performance* form.

**Conducts a preference assessment:** The trainer presents different stimuli to the client (i.e., one, two, or several at a time), and assesses client’s response to stimuli to determine partiality.

**Uses different ways to call attention:** The trainer uses a different way to gain/recapture the client’s interest each time (e.g., look at me, calls his/her name, entices client with a preferred item, among others).
Phrases $S^D$ as an instruction/statement: When applicable, the trainer presents the discriminative stimulus as a directive and or announcement every time, with the exception of one wrong (e.g., “ask can you touch vs. touch”) presentation during the five-minutes sample observation.

Differentially reinforces correct responding: The trainer consistently reinforces the client’s independent and socially-appropriate responses, with a higher intensity, quantity, or quality of a reinforcer. The trainer refrains from reinforcing challenging behaviors.

Contrives manding opportunities: The trainer contrives at least three manding opportunities for the client during the 5-minutes sample observation (e.g., concealing an object, providing incomplete set of objects, pausing a video, among others). Manding is the action of asking.

Interspersing skills: The trainer randomly presents a different operant each trial (e.g., tact, intraverbal, and echoic presented as: “what is this”? “You play with a_”, and “say_” (referring to a ball), so that the same operant is not presented twice in a row.

While shaping, DOES NOT reinforce incorrect responses. Only reinforces correct approximations.

Accurately collects data: The trainer gathers data in a timely manner, every time, as indicated on the first trainer video.

Praise: The trainer provides at least three different approving statements during the sample observation (e.g., a combination of: “great job”, “phenomenal”, and “beautiful”).

Incorporates previous supervision feedback: During baseline DTT sessions, the trainer demonstrates mastery of skills that were addressed at the previous supervision meeting. During treatment DTT sessions, the trainer demonstrates mastery of skills that were just role played at the previous supervision meeting.
Discriminative stimulus: A stimulus in the presence of which some type of responses was reinforced (Cooper, Heron, & Heward, 2007).

Mands/Manding: Verbal responses followed by a characteristic consequence where the verbal responses are evoked by conditions of deprivation or aversive stimulation (Skinner, 1957). Manding is the action of asking.

Tacts: Verbal operant responses that are evoked by discriminative stimuli and followed by reinforcement (Skinner, 1957). Tacting is synonymous with labeling.

Intraverbal: Verbal operant that involves responses evoked by verbal discriminative stimuli that do not have point-to-point correspondence with the previous verbal stimulus (Skinner, 1957). For example, and intraverbal response would be evoked when a person says “one, two, and …”, and a second person immediately says “three!”.

Echoics: Verbal operant involving responses that are evoked by verbal discriminative stimuli that have point-to-point correspondence and formal similarity with the response (Skinner, 1957). For example, an echoic response would be evoked with a person says, “chair” and a second person immediately repeats the same word “chair”.

Verbal operant: The unit of analysis of verbal behavior, which illustrates the functional relationship between different nonverbal variables and their corresponding types of responses. An operant is a type of behavior that is distinct from a particular response instance (Skinner, 1957).
Sigmund Freud initially spearheaded supervision in the field of psychology with the introduction of psychoanalysis (Watkins, 2011). “Supervision took embryonic form in 1902—when Freud opened up his home for regularly scheduled, Wednesday-night discussion meetings” (Freud, 1914, p. 403; as cited in Watkins, 2011 p. 403). Although there is still some debate as to whether Freud’s activities at the time can legitimately be called supervision, most psychologists tend to accept the idea that psychoanalytic supervision had its informal beginnings then (Watkins, 2011). Originally, supervision incorporated clinical techniques and approaches rooted in psychoanalysis (Falender & Shafranske, 2004). The aforementioned approach continues to guide current supervisory practice. Consequently, supervision approaches have been developed for each of the major theoretical foundations (Falender & Shafranske, 2004), including developmental approaches (Barret & Barber, 2005; Stoltenberg, 2005), attachment-informed approach to supervision (Bennett, 2008), evidence-based approach to supervision (Leffler, Jackson, West, McCarty, & Atkins, 2012), and most importantly, competency-based supervision (Welsh, Stanley, & Wilmoth, 2003; Madson, Campbell, Barrett, Brondino, & Melchert, 2005; Falender, & Shafranske, 2007; Celano, Smith, & Kaslow, 2010; Stucky, Bush, & Donders, 2010; Ng, Choudhuri, Noonan, & Ceballos, 2012). The above-mentioned researchers have approached supervision in a variety of ways, consistent with their school of thought. The following is a review of the existing literature on supervision, presented chronologically by supervision approach/model. Subsequently, a review of the medical literature on supervision will be presented, followed by an examination of the behavioral literature on supervision, as well as literature on fidelity and treatment adherence, ending with a review of BST research.
Multisystem Therapy Supervision

Schoenwald, Brown, and Henggeler (2000) provided a summary of the theoretical and empirical foundations of Multisystemic Therapy (MST), and of evidence supporting the effectiveness of the model. The manuscript focused on therapist implementation of the model, supervisory practices, as well as the use of consultation and program practices. With regards to supervisory practices, the authors expressed that supervision sessions are the primary setting in which supervisors attain evidence of clinicians’ implementation of the behavioral skills required to effectively implement MST. In addition, they listed three interrelated purposes of supervision: “(a) development of case-specific recommendations to speed progress toward outcomes for each client family, (b) monitoring of therapist adherence to MST treatment principles in all cases, and (c) advancement of clinicians’ developmental trajectories with respect to each aspect of the ongoing MST assessment and intervention process” (p. 119). Relatedly, Schoenwald et al. (2000) expressed that supervisors must promote clinicians’ adherence to the MST protocol. For new supervisors, an MST consultant provides ongoing training and consultation aimed at enabling him/her to effectively engage in this practice. Generally, supervision in MST is done in a group format. During sessions, the consultant first discusses five assumptions about MST supervision, followed by case discussion, role play, and feedback aimed at developing interventions for a family in need. Individual supervision is only provided if a crisis develops between scheduled group supervision sessions, a clinician develops a specific clinical competency, and for the purposes of assessing and remediating clinician’s consistently poor adherence to the MST model. Schoenwald et al. (2000) successfully argued for the importance of training qualified supervisors who can effectively monitor clinicians’ and clients’ progress. This article initiated the conversation that led to empirical research in MST including supervision practices.
Schoenwald, Henggeler, Brondino, and Rowland (2000) discussed challenges associated with specifying a complex and individualized treatment model and measuring fidelity, using MST as an example. In addition, they examined the relationships between therapist adherence to MST and outcome variables, as well as the relationship between clinical supervision and therapist adherence. In terms of the latter, the authors shared the results of a study they conducted aimed at fortifying adherence to MST protocols. During the study, MST therapists’ adherence to protocol was measured by experts who listened to session audiotapes and completed an adherence questionnaire. The questionnaire required that each expert answer questions regarding the extent to which audiotapes reflected therapists’ adherence to MST principles on a scale of one to seven. Supervisors used this information to provide structured feedback to the therapists during weekly individual meetings. When necessary, the frequency of supervisory meetings increased if therapists demonstrated repeated adherence problems, and audiotapes were conjointly reviewed by both parties. In addition, supervisors developed specific recommendations for promoting changes in therapists’ behaviors during sessions and continued to monitor therapists’ progress via the adherence questionnaires. Results of the study revealed that therapists’ adherence to MST protocols improved from pre- to post-implementation of the adherence monitoring procedure, and all therapists were able to achieve scores of seven within three to four months of initiation.

Henggeler, Schoenwald, Liao, Letourneau, and Edwards (2002) validated a measure of clinical supervision practices, a measure of therapist adherence, and examined the association between supervisory practices and therapist adherence to an evidence-based treatment model (i.e., multisystemic therapy [MST]) in real-world clinical settings. Results of the study revealed that supervisor expertise in MST and evidence-supported treatments was associated with
therapist fidelity to the MST model, based on caregiver reports. Conversely, supervisor focus on the MST analytic process, treatment principles, and developing therapists’ competencies were generally associated with low therapist fidelity to the MST model. Therefore, it may be safe to assume that high supervisory fidelity might have been validly associated with low therapist adherence. Overall, though the study employed a seemingly robust methodology, both the supervisor and therapist adherence measures contained a number of subjective items. Thus, true empirical measures were largely lacking.

**Supervision of School Psychologists**

Utilizing a qualitative case study research design, the focus of Welsh, Stanley, and Wilmoth’s (2003) study, was to discuss a competency-based model of training for the supervision of school psychology candidates “completing a two-semester field-based training at the pre-internship level” (p. 178). Participants attended a 72-hour Specialist in School Psychology degree program. This empirical problem-solver model was focused on the development of single-case research designs and their application to different mental health practices in which school psychologists engage. Specific practice competencies defined in behavioral terms, were generated and included “additional global practice elements of accurate work completion, dissemination of information, and professional conduct” (186). Supervision and assessment of candidate progress toward acquiring professional practice competencies were assessed using skill present or not present indicators. Although a promising arena due to its potential for increasing the quality of supervision in school psychology, this study left much to be desired. For one, it provided no data on the implementation of the model, which questions its basis and credibility. Furthermore, though it provided a sample list of specific competencies to be evaluated during the supervision of school psychology students, it lacked a systematic process
of implementing these competencies. Identification of what is considered among experts’
competent behavior alone is not enough, the process should also be carefully considered.
Moreover, the identification of 145 competencies may be challenged by a lack of resources,
time, and skill on the part of the supervisors.

**Developmental Approach to Supervision**

Barret and Barber (2005) presented a qualitative, philosophical research study in which they discussed a developmental approach to the supervision of therapists in training. In the paper, the authors first provided a research account of the problems with supervision. Some of these problems included negative feelings towards supervisors on the part of trainees due to a lack of knowledge of trainee’s emotional development. The writers then discussed several approaches to supervision, starting with those that follow a particular theoretical orientation, to those that include cross-theoretical approaches and others that focus solely on the training of technical skills. Next, the researchers discussed a variety of reasons why supervisors may not be attending to the needs of supervisees. These reasons they argued might have to do with a lack of assessment of the trainee’s cognitive and emotional needs during the span of the supervisory relationship. Subsequently, the researchers discussed a few models of therapist professional development, including the *Integrated Developmental Model (IDM)* by Stoltenberg, McNeill and Delworth (1998) (as cited in Barret & Barber, 2005), whose most noticeable limitation is the lack of a link between personal and professional development of the trainee. With a more in-depth look at a model of cognitive and emotional development, the authors offered a discussion of the *Kegan’s Model* (Kegan, 1994; Kegan, Lahey, & Souvaine, 1998; as cited in Barret & Barber, 2005). This mode addresses developmental stages as well as theoretical underpinnings of developmental transitions. Lastly, the authors proposed guidelines on the application of
COMPETENCY-BASED SUPERVISION IN ABA

developmental approaches to supervision as it applied to students of psychology. Some of the strengths of this proposed model of supervision include a systematic methodology of advancing through stages in the trainee’s clinical skill development, and a focus on the supervisee’s developmental needs, which may decrease potential negative interactions between supervisor and trainee. However, some limitations of the developmental model were also noted. Namely, this model fails to address the effects of supervision on its most important member: the client. Additionally, the model does not propose a systematic approach to teaching the supervisor how to carry out the model. Lastly, this study lacked an empirical methodology for assessing and monitoring supervisor performance.

Stoltenberg (2005) also published a qualitative, philosophical research study discussing the use of developmental approaches to enhance professional competence for counseling and psychology students. In this position article, the author presents a review of the literature in supervision to argue for the use of developmental models of supervision for the enhancement of trainees’ professional competence. To this end, the author discussed the Integrated Developmental Model (IDM), which addresses how supervisees change over time, the supervision environments, and interventions. IDM is composed of four stages ranging from high degree of structure toward less structured, nondirective supervision. In this way, the model accounts for both supervisee development and the supervision process. When considering markers of development, the author discussed ways in which to monitor progress in trainee development. Specifically, these methods reflected qualitative changes between the levels of supervisee development. Supervision environments were also part of the model’s discussion. The author described five main environments of clinical supervision: facilitative, prescriptive, conceptual, confrontive, and catalytic, and provided an example of their use within specific
COMPETENCY-BASED SUPERVISION IN ABA

domains of practice (i.e., individual therapy versus couple’s therapy). Similar to any other model, the present one also revealed some strengths and limitations. As noted by the authors, some of the strengths include the model’s consistency with general conceptualizations of human development, and the progressive achievement of autonomy that may generalize to other domains of professional practice. In contrast, some limitations included the understanding that it would require an experienced clinical supervisor to carry out the model with precision and accuracy. Additionally, the model requires a great deal of training, and it has not been adequately tested. Furthermore, this study did not mention reliance on data collection and analysis as a means of assessing supervisor or supervisee performance.

Competency-based Supervision Tool

Madson, Campbell, Barrett, Brondino, and Melchert (2005), developed an alternative measure for evaluating competency in supervision, which assesses information similar to that measured by the Motivational Interviewing Skill Code (MISC), which is used for rating therapist-client interactions in audio or videotaped Motivational Interviewing (MI) sessions in order to evaluate therapist’s adherence to the MI approach, as well as quality of implementation of the MI therapy model targeted for use in either a clinical or research setting. This measure was intended for use in the training and supervision of therapists implementing motivational interviewing. Four therapists participated in the study, and were trained by a team of researchers with experience in the implementation of MI. The training was based on the treatment manual used in the study with individual supervision provided on a regular basis. Results of the study suggested that the data obtained with the Motivational Interviewing Supervision and Training Scale (MISTS) were reliable and reasonably valid. Satisfactory inter-rater reliability was established between raters who independently used the MISTS to analyze audiotaped therapy
COMPETENCY-BASED SUPERVISION IN ABA

sessions. “This suggests that the MISTS have the potential to provide reliable data when used in similar ways to the way it was used in the study” (p. 307). Despite being the only study in this review that provided actual data and statistical analysis of the data, the use of a mixed methods research design on the reliability of the instruments lacked a true measure of experimental control.

Development of APA Guidelines for Supervision

In 2006, the American Psychological Association (APA) published a report on competence in professional psychology, which included guidelines for the ethical provision of clinical supervision as a required competency (APA, 2006). Since then, interest in this area has ignited the development of a variety of models for supervision in clinical psychology and counseling, spearheaded by the works of Falender and Shafranske (2007).

Competency-based Supervision

Falender and Shafranske (2007) utilized a qualitative, philosophical research, coupled with historical research design to analyze viewpoints on competence as a concept, and therefore provide a definition of competency-based clinical supervision in the training of graduate-level psychology students, while paying particular attention to the nature of contextual, legal, ethical, and practice issues that may arise from the establishment of a standard of competency-based supervision practice. The researchers began by providing a definition of competence, followed by a discussion of the core competencies in psychology. Then, the authors discussed the idea of competence as an ethical standard, debating the concept of Metacompetence. Within the construct of competency-based clinical supervision, the writers discussed clinical competence, and the idea of a competency-based approach for the enhancement of clinical skills in supervision. The authors concluded with a discussion of challenges confronted within clinical
COMPETENCY-BASED SUPERVISION IN ABA

supervision, and 12 recommendations for competency-based supervision best practices. Some of the challenges listed by the authors regarding the use of a competency-based model of clinical supervision involved the identification of essential practices needed to perform specific protocols, how to determine whether these practices will be accurately and reliably assessed, preparation prior to conducting clinical supervision, supervisor self-assessment, supervisor ethical competence as well as areas and levels of incompetence, supervisor’s diversity and multicultural competence, and finally supervisor’s level of professional development (Falender & Shafranske, 2007). Strengths of the model proposed by the authors involved the establishment of clearly defined behavioral expectations and standards of performance as specific competencies, which are observable, and therefore may be evaluated (Falender & Shafranske 2007). However, a major limitation of the model was it was theoretical in nature, without empirical validation.

Gonsalvez and Calvert (2014) provided an explanation of the theoretical underpinnings and key features of competency-based models, explained the principles and implications of these models for supervision, and discussed the potential merits, limitations and challenges of the competency-based model. The authors concluded that within psychology, there appears to be quite the consensus favoring a three-dimensional model of supervision, which involves a combination of “foundational and functional competency domains across developmental stages” (p. 202). Therefore, as Gonsalvez and Calvert (2014) expressed, a clearer conceptualization of relevant dependent measures and potential mediating independent variables may facilitate the planning and execution of research aimed at comparing specific effects of supervisory methods and strategies, which may in turn increase the effectiveness of current (and future) supervisory practices and protocols. However, one key limitation of this publication, was that the authors
COMPETENCY-BASED SUPERVISION IN ABA

provided no empirical data to support the theoretical underpinnings and key features of competency-based models.

Stein, Clair, Soenksen-Bassett, Martin, and Clarke (2015), applied a competency-based supervision model coupled with a psychotherapy-based approach to the training and supervision of supervisees in the area of motivational interviewing (MI). To evaluate this psychotherapy-based approach, the authors selected five major aims: (a) measuring pre and post training MI skills, (b) measuring the relationship between post training skills and conduct of training, (c) determining the extent to which this approach based on MI had been applied to both training and ongoing supervision sessions, (d) determining the extent to which supervisors’ experiences might impact both training and supervision processes, and (e) determining the extent to which this approach to training and supervision would be applicable across different types of trainees. A total of 36 individuals participated in the study. During baseline, participants were required to complete the Helpful Responses Questionnaire (HRQ; Miller, Hedrick, & Orlofsky, 1991). Following a pretest, training sessions in MI were conducted, after which a training evaluation was administered, along with the HRQ posttest. After approximately five months of providing MI with coached supervision, supervisors conducted an evaluation of supervision. Results of the study revealed that HRQ pre and posttest scores were significantly different (aim one), with raters identifying significantly less roadblocks in the posttests when compared to the pretests. Additionally, HRQ posttest scores correlated with training evaluation scores (aim two), and correlations between training and supervision were found to be statistically significant, producing effect sizes in the large range (aim three). Regarding differences between supervisors on HRQ posttest and training evaluation scores (aim four), effect sizes were also found to be in the large range, which suggested differences on outcomes by supervisor. When considering the
COMPETENCY-BASED SUPERVISION IN ABA

last aim (number five), HRQ posttests and training evaluations showed that research staff had better outcomes than student trainees. However, though the aims of the study were fulfilled, it lacked measurement of specific behavior variables responsible for the results. Lastly, the study reported survey data. The accuracy of survey data may be questioned, as verbal reports do not always match observable behavior.

Attachment-informed Approach to Supervision

Bennett (2008) presented a study in which he used an attachment-informed model of supervision to train social workers. Specifically, the author presented an 8-month model of supervision training for social work field instructors of Master in Social Work (MSW) students. Meant to educate beginning social work field instructors about the basics of attachment theory and its usefulness for effective supervision in the MS field internship, the model emphasized the establishment and maintenance of a positive supervision working alliance, defined as relational bond. This pilot mixed-methods AB group design program consisted of three in-person daylong workshops, five online classes, as well as online conversations through Blackboard. The in-person workshops offered at the beginning, middle, and end of the academic year, as well as the online classes integrated Power Point presentations, journal articles, and questions. Participants consisted of one male and nine females volunteer novel supervisors of first year MSW social work students and other (selected) experienced supervisors. Each participant received 25 hours of continuing education credits in supervision. The training program consisted of eight modules offered in the following way: modules one, five and eight were offered in-person, and modules two through four, as well as six and seven were offered online. During the last workshop, an evaluation was provided to the participants encouraging them to describe their supervisory experience and training program throughout the academic year. Results of the study showed that
COMPETENCY-BASED SUPERVISION IN ABA

from the participants’ perspective, the program was a success, as supervisors grew in confidence and skills. Additionally, the subjects gained a deeper understanding of relational dynamics. Other strengths of the study included the opportunity for participants to earn continuing education credits, as well as the convenient access of the course material online, which facilitated a more individualized learning pace. However, the study provided very little quantitative and qualitative data, and there was a lack of standardized methods or controlled research design. Moreover, the group process utilized in the study is less attuned to individual supervisory needs, and the online classes revealed a lack of spontaneous participation.

**Couple and Family Therapy Supervision**

Utilizing a qualitative case study research design, Celano, Smith, and Kaslow (2010), provided a brief illustration of essential components and related functional and foundational competencies for intervention in the area of psychotherapy from a systemic standpoint (i.e., interactions among family members in their relational contexts), also known as couple and family therapy (CFT), and addressed these competencies within the supervisory process. First, the authors proposed eight essential components of CFT, including: (a) developing a systemic foundation; (b) forging a system of therapeutic alliance; (c) understanding family-of-origin issues; (d) reframing; managing negative interactions; (e) building cohesion, intimacy, communication; restructuring, parenting; and (f) understanding and applying evidence-based CFT models. Subsequently, the writers proposed a series of foundational and functional competencies informing CFT competency. These competencies encompassed the ability to demonstrate reflective practice, self-assessment, self-care, the application of ethical and legal standards to professional activities, individual and cultural diversity competency, competency of assessment, and consultation. Next, the authors discussed how the essential components of the
COMPETENCY-BASED SUPERVISION IN ABA

CFT competency would be addressed within a supervisory approach. Providing a case study illustration, the authors described how components of CFT could be successfully applied in the supervision of a graduate-level psychology trainee at a therapeutic impasse with a family. Two major strengths were identified in this study. These included a systemic approach to competency in CFT supervision by identifying essential skills required in the competent practice of this intervention, and the use of behavioral techniques such as modeling, which aided the trainee in the therapy process. Conversely, several limitations were found in the study including a lack of experimental control, and measurement systems, as well as a lack of clarity, which from a replication standpoint, threatens both the validity and reliability of the study.

**Clinical Neuropsychology**

Stucky, Bush, and Donders (2010), focused on the provision of a framework and guidance for the development of competency standards for training of neuropsychological supervisors, particularly at the residency level. Utilizing a qualitative, philosophical research design, Stucky et al. (2010) proposed a competency-based approach to supervision that “advocates for a science-informed, formalized, and objective process that clearly delineates the competencies required for good supervisory practice” (p. 741), and “was intended to capture core general features of neuropsychological supervision that could be applied in a variety of contexts” (p. 742). This model was an individually tailored, process-based, developmental approach where both foundational and functional supervisor competencies were emphasized. Additionally, basic supervisee competencies were identified, and a comparison was presented between supervision in neuropsychology and other forms of supervision in the field of clinical psychology. As part of the foundational competencies, the authors identified (and discussed in detail) six competency domains that make an effective neuropsychology supervisor. In addition, they identified and
explained another 6 functional supervisory competencies for effective supervision of neuropsychology trainees. Lastly, Stucky et al. (2010) identified and discussed a set of eight supervisee competencies for neuropsychology residents. The authors then ended the paper with a list of limitations regarding supervision in within this branch of psychology. These included: the fact that formal supervisory training standards for neuropsychology are in their beginning stages. Additionally, limitations with funding, small faculty size, limited access to formal supervision training programs, and a lack of residency exposure, may affect adequate training in the provision of competent supervision for neuropsychology students. Nevertheless, one strength of this approach to supervision is the individualization of the supervision process, taking into consideration each supervisor’s level of development. Additionally, a competency-based approach to supervision will always benefit from a decrease in subjective evaluation, thereby also decreasing potential discrepancies between supervisors and supervisees. Yet, one key limitation of this paper was the lack of empirical data to support the results of the study.

**Competency-based training in medicine**

Yuen, Barrington, Headford, McNulty, and Smith (1998) developed a competency-based palliative care program for the training and evaluation of doctors who practice in hospices, hospitals and home-care services. According to the authors, training in palliative care is needed for these doctors. In particular, training programs that articulate the appropriate knowledge, skills and attributes required by doctors wishing to work in the field. The authors further noted that training programs should be developed that ensure an initial competency level of these doctors as well as maintain their appropriate skill levels, thereby filling the gap between a more formalized assessment system linked to an appropriate training mechanism in medical palliative care education.
COMPETENCY-BASED SUPERVISION IN ABA

The program development was undertaken in Perth, Western Australia. The Silver Chain Hospice Care Service (SCHCS) worked with 29 doctors encompassing primarily family care physicians. Most had limited instruction in palliative care prior to beginning this work. SCHCS established a work group to develop a competency-based program for the training and evaluation of these doctors. Competency was measured at work under typical conditions, including all the attendant resources, opportunities, limitations, distractions, and pressures. The objective was to obtain a comprehensive and realistic picture of an individual’s abilities. Measuring competencies required both a detailed description of the work, and a set of standards for performing each part of it. Standards of performance were based on those outcomes recognized as having an enhancing or generally positive effect on the care of the patient and family and the functioning of the caregiving team. Assessment of doctors’ competency encompassed the effectiveness of both his/her medical interventions and his/her interactions with other team members.

Seven doctors began the development of the program. Among the six applications of competencies identified were monitoring doctor’s own competencies and identify areas that required further competency, as well as specifying the required quality of work. The final practice model developed included six categories. In each category, competency units were identified. The units were then further divided into a set of specific competence elements.

The competency program was given a trial as part of an educational package for 23 practitioners who participated in a palliative care education program. The practitioners were mentored for 14 weeks. Though feedback from mentors and participants was mixed, 53% of the participants found the competency document useful and relevant for evaluating their progress. This study illustrated the development and use of a competency-based training program for educating medical doctors in palliative care. Although the study lacked true empirical behavioral
COMPETENCY-BASED SUPERVISION IN ABA

measures, as well as a sound research design, it did provide evidentiary support for the use of competency-based training methodologies in medicine. The current study also proposes the use of a competency-based training.

Emmett, and Green-McKenzie (2001) described a competency-based training program that allowed physicians employed full-time in occupational and environmental medicine to satisfy the supervised practicum year of training required by the American Board of Preventive Medicine (ABPM). The program was designed for trainees with greater clinical experience than the 1 clinical year required by the ABPM. Results of the study revealed that the practicum-year training was highly successful, as evidenced by improvements in resident self-assessment of competency, resident satisfaction with the training, faculty evaluation of resident performance, and success rate in the ABPM examination. The training program was competency-based, and the objectives were designed to incorporate the list of program requirements for Residency Education in Preventive Medicine for Occupational Medicine3 and for the ABPM.4 The competencies are arranged in six subject groupings. The first grouping incorporates several General Competencies. At the beginning of the supervised practicum-year program, each resident completed a self-evaluation against each competency objective. Each competency was given a self-rating using a five-point Likert scale (1 lowest, 5 highest). The self-ratings were reviewed by the resident’s site supervisor and faculty preceptor to ensure that they were realistic. Based on the self-evaluation, an educational plan for the year was developed. Each resident had three levels of supervision for different aspects of the training program. These included: (a) a full-time faculty preceptor, who was either the program director or the associate program director, and had overall responsibility for the program and the progress of the residents; (b) one or more faculty supervisors for each competency module, who were faculty members at the
COMPETENCY-BASED SUPERVISION IN ABA

University of Pennsylvania; and (c) site supervisors, who were suitably qualified physicians whose credentials were reviewed as part of the residents’ application process. Both formative and summative evaluations were used, to evaluate outcomes and enhance quality. Formative evaluations such as resident reviews of each module and feedback from faculty and the Residency Advisory Committee, were used to ensure continuous improvement and relentless adaptation to training needs. The evaluations of program outcomes were achieved using three independent outcome measures: resident self-evaluations, faculty evaluations, and the success of graduates in the ABPM examination in Occupational Medicine. Resident self-evaluations were performed at the start and after the program. This study supports the use of competency-based training/supervision.

The “external” supervised practicum-training year described here successfully met its competency learning objectives, as measured by resident self-evaluations demonstrating competency enhancement, faculty evaluation of resident competence following training, and achievement in the ABPM certification examination. It has also met the perceived training needs of the residents. This innovative, competency-based, supervised practicum-year training program demonstrated that by using principles of adult learning, physicians making a mid-career shift to OEM can successfully be trained, as measured by resident self-evaluation, evaluation of competence by experienced faculty, and success in ABPM specialty examinations. Though this study involved a competency-based program, one major weakness concerned the lack of performance measurements rooted in direct observation.

Louie, Coverdale, and Roberts (2004) expressed their support of critical and creative thinking in the process of competency-based training in Psychiatry residency. To this end, the authors offered the following comments: (a) the change to competency-based training must
COMPETENCY-BASED SUPERVISION IN ABA

involve a broader cultural change, and training duration may be determined by the achievement of competencies as opposed to any other factor; (b) competency should be understood within a developmental framework - continuum of ability; and (c) psychiatry educators should guard against “teaching to the test”, and in doing so, forsake the most sophisticated skills of a physician. Competence measures should have to include observing therapists’ work.

Basnet, Clapham, Shakya, and McCall (2004) provided a literature review that presented the findings of the first 4 years (1999–2002) of implementation of post abortion care (PAC) services outside of major urban centers in Nepal, where a significant proportion of services are provided by nurses. Clinical competency assessments and service utilization data from three district hospital-based post abortion service sites supported by the Nepal Safer Motherhood Project were analyzed. The relationship between the findings of this assessment and two previous assessments, were discussed. This review found that nurses are at least as competent as physicians in providing post abortion care services. Once again, this study offers support for the use of competency-based assessments in medicine.

Pryor, et al. (2006) as part of a multidisciplinary team designed, developed, and offered a four-day, functional exercise, competency-based course at a national training center. The report provided a description of the development and initial evaluation of the course in changing participants' perceptions of their capabilities to respond to weapons of mass destruction (WMD) events. Course participants encompassed healthcare professionals, including physicians, nurses, emergency medical services administrators, hospital administrators, and public health officials. Each course included three modified tabletop and/or real-time functional exercises. A total of 441 participants attended one of the eight course offerings between March and August 2003. An intervention group only, pre-post design was used to evaluate change in perceived capabilities
related to administrative decision-making for WMD incidents. Paired evaluation data were available on 339 participants (81.9%). Self-ratings for each of 21 capability statements were compared before and after the course. A 19-item total scale score for each participant was calculated from the pre-course and post-course evaluations. Paired t-tests on pre-and post-course total scores were conducted separately for each course. Results of the study revealed that there was consistent improvement in self-rated capabilities after course completion for all 21 capability statements. Paired t-tests of pre- and post-course total scale scores indicated a significant increase in mean ratings for each course (all \( > 0.001 \)). Overall, the tabletop/real-time-exercise format was effective in increasing healthcare administrators' self-rated capabilities related to WMD disaster management and response. Integrating the competencies into training interventions designed for a specific target audience and deploying them into an interactive learning environment allowed the competency-based training objectives to be accomplished.

Like the previous studies, Pryor, et al. (2006) offered evidence to support the use of competency-based training.

Rubulotta, et al. (2009) developed a questionnaire to assess 21 elements of professional competence. The aim of the Competency-Based Training in Intensive Care medicine in Europe (CoBaTrICE) project is to create an internationally acceptable competency-based training program for specialists in intensive care medicine. The CoBaTrICE project performed a survey, in collaboration with the Picker institute, to identify desirable characteristics of Intensive Care Unit (ICU) specialists, as expressed by patients and their relatives. A questionnaire was developed to assess 21 elements of professional competence. Each element was assigned to one of four categories of a Likert scale: 1=essential; 2=very important; 3=not too important; 4=does not matter. The results were dichotomized into essential (score: 1) and not essential (scores: 2-4)
COMPETENCY-BASED SUPERVISION IN ABA

categories. Further, the elements were related to three key concepts: “medical skills and competencies”, “communication with patients”, and “communication with relatives”.

Questionnaire statements grouped by theme were also ranked for each item using a number: 1=highest rank; 21=lowest rank. Free text responses were also invited. Ten Italian ICUs were enrolled in the study. There were 249 questionnaires completed (18% total return rate). Priority in Italy was given to medical skills and competence. Overall, involvement of patients and relatives in decision-making processes were among the items considered least important. Italian families preferred a paternalist approach to the end of life decision-making process. Once more, this study provided support for the use of competency-based tools in training.

International Competency-based Supervision.

Utilizing both qualitative and quantitative research designs, Ng, Choudhuri, Noonan, and Ceballos (2012), focused on developing an international competency checklist based on literature and data from individual experts regarded as knowledgeable about and having experience with the development of instruments of higher education for international use, particularly in general and professional counseling arenas. The development of this checklist underwent a validation process, which resulted in 43 items meeting a predetermined 60% positive rate of support. The aforementioned items were distributed across six domains. Therefore, the checklist was the first to be verified by a distinct group of expert individuals in and outside of the U.S. Based on the findings, the authors recommend that U.S. counseling training programs consider using the 43 items that are supported by experts in the field as a set of primary indicators to guide them in the propagation of their work on international grounds. Additionally, they proposed to name the 43 indicators set the “Internationalization Competency Checklist” (p. 33). However, a review of several databases found no studies in which this checklist was empirically evaluated or validated.
Evidence-based Approach to Supervision

Milne and Reiser (2012) used a qualitative, philosophical research design coupled with historical research design to provide a rationale for the need for evidence-based clinical supervision of graduate-level students in psychology. First, the authors illustrated some of the problems that can arise in the absence of evidence-based practice (EBP) by alluding to historical arguments for the need of EBP, such as lack of formal training, licensing or regulation in the area of supervision, whether unscientific methods are currently being utilized and clinicians are practicing without fluent knowledge of the methods they are implementing, poor training, lack of an empirically validated assessment of supervisory competence, among others. Additionally, the researchers indicated a high degree of variability in clinical supervision competence amongst qualified professionals. Conversely, the authors also provided historical and professional arguments against the use of EBP in supervision. Mostly, these arguments consisted of subjectivity in professional judgment, perceiving guidelines as “cookbook” approaches, understanding EBP as a threat to the practitioner’s autonomy of action, a higher level of accountability, and a step toward greater objectivity and transparency. Actually, the authors viewed these last two as strengths of this model, and further noted that EBP raises the standard of care, has the potential to improve clinical outcomes, as well as to rectify the gap between research and practice. Additionally, they stressed the need for EBP to engage in practices based on the best evidence available on effective and ethical supervision. Clinicians, including supervisors, need a set of key tasks, guidelines, training manuals and evaluation instruments to aid in the ethical practice of clinical supervision. EBP also provides a methodological stance on practice, and protection from legal and other professional competence challenges. To illustrate the process of EBP in supervision, the authors provided a diagrammatic representation, as well as
three examples of EBP in clinical supervision (i.e., supervisor training manual plus consultancy, corrective feedback, and applying guidelines).

Leffler, Jackson, West, McCarty, and Atkins (2012) also utilized a qualitative, philosophical research design coupled with historical research design to describe several training modules used to enhance skills and techniques of EBP in psychology trainees from three graduate-level programs. The first model, from the University of Kansas Clinical Child Psychology Program (KUCCPP) is based on the Boulder model (Frank, 1986), which requires that a scientific approach be applied to all training efforts. The aforementioned KUCCPP model evaluated coursework and practicum expectations of graduate-level psychology students, as well as tools for training and documentation of competency. The second model, from the University of Illinois at Chicago (UIC) Department of Psychiatry: Clinical-Scientist Internship Training, focuses on ensuring that trainees of clinical psychology receive adequate training in the provision of EBP interventions, by incorporating clinical research experiences into the internship year. The third model involved the implementation of an intervention to address multi-family psycho-education psychotherapy (MF-PEP). MF-PEP is composed of a manual-guided 8-week group therapy for children from 8 to 12 years of age suffering with mood disorders, and their families. Some of the common strengths shared by all three models include an individualized training approach tailored to meet the needs of the trainee, the use of activities as learning strategies, and the provision of a means of measuring trainee’s progress with their learning outcomes across time and people. In contrast, some of the limitations shared by all three models reviewed within this study include lack of trainee interest, lack or qualified supervisors, considerable time to be committed to the accurate implementation of each model, and difficulty when training.
COMPETENCY-BASED SUPERVISION IN ABA

Competency Benchmarks for Professional Psychology Supervisors

In recent years, Hatcher et al., (2013) presented a description of the latest version of “The Competencies Benchmarks for Professional Psychology” (p. 84), which included a condensed version of the benchmarks aimed at facilitating the application of these competencies by supervisors of clinical psychology. This publication also included a discussion on the use of rating forms designed to assess the performance of supervisees on a Likert-type scale. More importantly, Hatcher et al., (2013) efficiently pointed out the need for empirically studying the effectiveness of supervision, in order to aid supervisees in attaining competency, as well as ensuring the provision of comprehensive and effective service delivery to clients.

Supervision in Play Therapy

Recently, Hudspeth (2015) provided a review of the literature of supervision within the context of play therapy. Two of the articles discussed employed qualitative research methodologies, while another one reported correlational results, and three additional ones relied on practice and application. The first qualitative article reviewed in this study described the use of Lego Seriously Play and its practical application within play therapy supervision. Similarly, the second qualitative study described critical incidents that may occur in the process of teaching play therapy, how these incidents are processed, as well as implications resulting from processing the incidents and recommendations for how a Critical Incidents Method may be used within play therapy training and education. In contrast, the third article reviewed examined the use of mindfulness and critical reflection as a means to appraise the supervisory relationship from the supervisor’s perspective. Further into the article, the author presented a literature review of play therapy supervision that discussed research related to supervision methods and models, and included aspects specific to play therapy and recommendations for future research. The
second-to-last study discussed in this article pertained to a correlational, quantitative design study in which the authors examined the relationship between trainees’ current supervision experiences and their satisfaction with supervision overall. Results of the study suggested that supervisors’ professional identity, and trainees’ years of experience may predict trainees’ levels of satisfaction with the supervision process. The last article discussed by Hudspeth (2015) presented a description of the supervision process from a developmental model perspective with four identifiable categories/stages for students, and recommendations regarding expressive arts activities that may aid in the improvement of communication, facilitating awareness, increasing empathy, and improving overall skills.

**Early Behavioral Research on Supervisory Behavior**

Behavior analysis can trace the beginnings of research on supervisory behavior back to the early 1970’s. Perhaps the first paper related to supervisory behavior was published by Sorcher (1971). In this paper, the author provided several arguments for the use of behavior modification in supervisor training. Among the reasons included, the author argued that psychotherapists had been successfully using behavior modification procedures to change behavior, and that “just as many of the problems treated by the therapists involve difficulties with interpersonal or social relationships, the problems of supervising or managing are often quite similar in this respect” (p. 401). Additionally, Sorcher (1971) indicated that behavior modification appeared to provide a bridge between clinical practice and industrial psychology, and that the objective of a behavior modification program would be to provide adaptive alternative responses to employee problems that supervisors may encounter in their daily work. The author further argued that supervisor’s maladaptive behaviors could successfully be treated
COMPETENCY-BASED SUPERVISION IN ABA

using principles of behavior modification, by means of observation, modeling, role playing, imitation, and practice.

A few years later, Rosenbaum (1975) proposed a behavioral modeling approach to changing supervisory behavior. The training emphasized modeling, behavior rehearsal, feedback and social reinforcement, as well as transferring training to the job. Participants were taught to focus on actual behaviors of the supervisees as opposed to theorizing.

From a review of the behavioral literature on supervision encompassing performance management journals, two main areas of research appear to emerge: studies that focus on affecting the behaviors of supervisors, and studies that focus on affecting the behaviors of supervisees.

Using Behavioral Methods to Improve Supervisor Behavior

Ivancic, Reid, Iwata, Faw, and Page (1981) evaluated a supervision program that incorporated language training into direct staffs’ care routines for institutionalized profoundly handicapped children. The principal aim was to increase and maintain the use of these training activities by staff. A second aim was to evaluate the generalization of the systematic supervision program across direct care routines. Utilizing a multiple baseline across staffs’ behavior research design, staffs were taught to vocalize more during bath routines, praise vocalizations by the children, imitate children’s sounds, and provide sound prompts. Procedures were comprised of a brief in-service meeting, which was followed by a series of supervisory prompts and feedback. Results indicated an increase in staffs’ performance levels in each category (i.e., antecedent vocalizations, descriptive praise, sound imitation, and sound prompts) during treatment, when compared to baseline.
COMPETENCY-BASED SUPERVISION IN ABA

Methot, Williams, Cummings, and Bradshaw (1996) investigated the effects of a supervisor training program designed to increase the extent to which supervisors would use contingent consequences and objective measures when monitoring and evaluating direct-care staffs’ performances in human service agencies. To achieve this goal, the authors measured generalization of skills from the training context to on-the-job interactions with staff. Additionally, data were collected to examine whether changes in supervisor’s performances produced changes in the direct-care staff’s use of contingent consequences for clients’ performances as well as changes in clients’ behaviors. Overall, desired behavior changes were observed for all participants at the supervisory/management level, for six out of seven direct care staff members, for nine out of 15 clients on target behavior measures, and for eight out of 12 clients on on-task measures.

Changing Supervisees’ Performances Through Behavioral Training

Dillon, Kent, and Malott (1980) described a supervisory system designed to help 15 master-level students maintain a steady rate of work toward the completion of their master’s thesis. The system was composed of written task specification, weekly sub goals/deadlines, monitoring, feedback, as well as an added point incentive system. The incentive system consisted of allowing the students to earn performance-based positive and negative points toward a letter of recommendation from the faculty advisor. Utilizing a BAB reversal design, the authors first implemented a letter of recommendation and feedback phase, then a no letter/no feedback phase, and finally a return to letter of recommendation and feedback. Results of the study revealed that the master’s students showed a high rate of task completion when the letter of recommendation and feedback were in place, an immediate decrease when the letter and feedback were removed, and a significant recovery when the letter and feedback were reintroduced.
Hundert and Hopkins (1992) utilized a multiple baseline across subjects’ design to examine the effects of training three supervisors of integrated preschools to encourage teachers to develop strategies that promoted peer interaction of all children (including those with disabilities). Supervisors were trained using a manual, modeling, and role-playing, which they implemented with classroom and resource teachers. Results of the study showed that all classroom teachers increased their behaviors toward children with disabilities in both training and generalization settings. At the same time, teachers’ focus toward children without disabilities decreased. In addition, children with developmental disabilities more than doubled their level of interactive play in the training setting and showed a similar increase in the generalization setting. In some ways similar to the proposed research, this study utilized a multiple-baseline design, and the independent variable consisted of a package that included modeling and role-playing.

The present study aimed at filling the gap on previous behavioral studies on supervision by proposing for the first time: (a) the implementation a training protocol that concentrates on changing the behaviors of BCBA supervisors so that they may in turn, affect the performances of ABA trainers; (b) the provision of a systematic means of achieving this change drawn from current BACB requirements, as well as past and current research on supervision and ABA; (c) the focus of supervision within one critical area of ABA: Discrete Trials Teaching; and (d) the modification and use of well-established behavioral protocol in the ABA literature: BST.

**Recent Research on Supervision for Behavior Analysts**

Sellers, Valentino, and LeBlanc (2016), addressed five specific fundamental recommended practices for the supervision of aspiring behavior analysts. They also provided strategies and resources for each recommended practice in order to aid the supervisor in structuring the supervision process. The proposed guidelines include: (a) establishing an
COMPETENCY-BASED SUPERVISION IN ABA

effective supervisor-supervisee relationship by issuing supervision contracts, establishing clear expectations including setting the expectation that feedback will be provided and should be accepted in a positive manner, and creating a committed and positive learning environment for the supervisee; (b) establishing a plan for supervision that is structured in content and includes competence evaluations by relying on the task list items, and creating competency-based evaluations where observations, assistance, and feedback are embedded; (c) evaluating the effects of supervision by tracking and rating the number of competencies completed by the supervisee, assessing the supervisee’s language during meetings, soliciting feedback directly from the supervisee, and creating structured surveys for the supervisee to complete; (d) incorporating ethics and professional development into the supervision process by exposing supervisees to an assortment of ethical dilemmas, and fostering discussion and problem solving, as well as encouraging and providing opportunities for professional development including participation in local ABA associations/chapters, attending conferences, and reviewing the ABA literature; and (e) ensuring the continuation of the professional relationship after certification by establishing parameters for the continued support of the new certificant, including the creation of professional networks, access to university resources, and maintaining a stream of active communication.

As part of the Journal of Behavior Analysis in Practice’s special section in supervision practices, Sellers, Alai-Rosales, and MacDonald (2016) discussed rationales and consequences related to supervision concerns, provided guidelines for professional development in supervisory areas within the BACB ethics code, encouraged critical discourse relative to supervision, as well as provided case examples that illustrated supervisory subcomponents of the “Behavior Analysts as Supervisors” section of the BACB Professional and Ethical Compliance Code for Behavior
COMPETENCY-BASED SUPERVISION IN ABA

*Analysts* (BACB, 2019). In addition, the authors provided a rationale for each component, as well as a discussion of possible undesirable consequences resulting from failure to follow the rule. The article begins with a discussion of ethics code item 5.01 regarding supervisors only supervising within their scope of practice. The discussion then moves on to item 5.02 related to supervisory volume. Item 5.03 addressed appropriate supervisory delegation of duties, and item 5.04 focused on designing effective supervision and training, that encompass evidence-based procedures and best practices. Communication responsibilities during the supervisor-supervisee relationship were addressed via items 5.05 and 5.06. They pertained to communication of supervision conditions as well as providing feedback to the supervisees, respectively. The last item addressed evaluated the effects of supervision, listing three evaluation levels: a) consumer performance, b) staff performance, and c) supervisor performance. Overall, this article explored some of the nuances inherent in each section of the supervision code and offered a rationale for the codes to promote critical analysis. Lastly, the article provided BCBA supervisors with directions and resources to maintain as well as enhance their supervisory repertoires.

Sellers, LeBlanc, and Valentino (2016) proposed a systematic approach to issues that may arise during the supervisory relationship. Furthermore, they offered guidance for supervisors to identify and address barriers to successful supervision, that may be related to potentially damaged supervisory relationships and persistent interpersonal and professional skills’ challenges of the supervisee. The authors also extended a general call to supervisors to “continually and thoughtfully reflect on their own history, repertoires, and behavior, such that they may continue professional growth as supervisors” (p. 309). The authors shared a table of resources for addressing issues during supervision. This table was divided into three categories: a) persistent supervision issue, b) potential indicators, and c) assessment and intervention ideas.
and resources. The authors discussed ways to address persistent professionalism and interpersonal problems, disorganization and poor time management, as well as poor interpersonal skills, and difficulty accepting and applying supervisor feedback.

Turner, Fischer, and Luiselli (2016), presented a practice model and offered considerations for supervising ABA students, that are consistent with the BACB’s Supervisor Training Curriculum Outline (BACB, 2012), the Professional and Ethical Compliance Code for Behavior Analysts (BACB, 2019), and existing literature from ABA as well as other fields. The practice model for behavior analytic supervision encompassed: a) setting the occasion for collaborative and ethical supervision; b) defining the relationship and performance expectations; c) evaluating performance expectations; d) confidentiality; e) remote supervision; f) termination of the supervisory relationship; g) baseline assessment of the supervisees’ skills; h) methods to teach and maintain skills; i) developing case conceptualization, problem-solving and decision-making repertoires; j) considerations in the delivery of performance feedback; k) evaluation of the process and outcomes of supervision; and l) continuing education in the area of supervision. Lastly, the authors discussed initiating an agenda in behavior analytic supervision.

Valentino, LeBlanc, and Sellers (2016), focused on the benefits of group supervision and recommended a structure for its implementation. Specifically, they concentrated on critical supervision characteristics such as the quality of the supervision experience and the supervisor’s arrangement of the experience components to maximize the efficacy of learning opportunities. In addition, the authors suggested ways behavior analysts can optimize the group learning experience. The article described the primary characteristics of a group supervision arrangement that are distinctive, as well as discussed how to plan and conduct group supervision. The authors provided resources, specific strategies, and an overall structure for group supervision
implementation. First, the authors discussed characteristics and benefits associated with learning that occurs in a group environment. Within this topic, the authors discussed the importance of peer feedback, social networking, having multiple listeners for the same event, observational learning and developing empathy, modeling and rehearsing productive group discussions, practicing public speaking, and developing professional repertoires. Regarding structuring the group supervision experience, the authors proposed the creation of a schedule and standardized format, the use of group supervision for the generalization of skills learned during individual supervision, incorporating public presentation opportunities, planning for specific behaviors to ensure productivity, and managing interpersonal dynamics. Lastly, the authors provide thoughts for future research in this area.

Relatedly, Hartley, Courtney, Rosswurm, and LaMarca (2016) presented an apprenticeship model of supervision that could be adopted by organizations to allow its BCBA to successfully supervise student-analysts. This model consisted of dividing the supervisees’ experiences into 750 hours of supervision of direct therapy implementation skills (of the 1500 required by the BACB) during their first year, followed by the completion of 750 hours of BCBA apprenticeship during the supervisees’ second year. In the latter year, student analysts become familiar with BCBA’s responsibilities such as client and technician oversight. According to the author’s description of this model’s utilization within an organization, “its supervisees become experts with the autism population by shadowing and purposefully learning the BCBA role in autism-specific programming by applying behavior analytic principles to skill acceleration and behavior deceleration objectives” (p. 333).

Similarly, Garza, McGee, Schenk, and Wiskirchen (2018) outlined a systematic approach to BCBA student-analyst supervision and provided tools for supervisors. Items included in the
COMPETENCY-BASED SUPERVISION IN ABA

approach consisted of a) establishing a supervisory relationship, b) assessing the supervisee’s skill-level, c) developing a responsibility matrix or “job model” (p. 64), d) assessing the supervisee’s skills relative his/her job position, e) setting appropriate goals for supervisees, f) utilizing BST for the training of supervisees, g) ongoing monitoring of supervisee’s performance, h) guidelines on how to appropriately end the supervisory relationship, i) evaluation of the supervisee’s performance before, during and after supervision, j) evaluation of supervisory effectiveness, and k) supervision professional development planning. The authors encouraged supervisors to use the materials discussed and share them with others.

**BACB Requirements for Supervisors**

The field of ABA has also made efforts to provide professional guidelines for the ethical provision of supervision first by publishing a supplementary newsletter containing a list of required topics (in the form of modules) to be addressed in the training of clinical supervisors (BACB, 2012), and afterwards by announcing new supervisor qualification requirements in 2017 along with the development and subsequent dissemination of a new edition of the *Supervisor Training Curriculum Outline (2.0)* in 2019 (BACB 2017; 2019). The original set of modules were to be completed by current and future professional behavior analysts by December 31st, 2014, in order to fulfill eligibility criteria to continue to, and in some cases begin to provide clinical supervision (BACB, 2012). The curriculum outlined six specific competencies which included the following: “I) the purpose of supervision, II) important features of supervision, III) behavioral skills training, IV) delivering performance feedback, V) evaluating the effects of supervision, and VI) ongoing professional development” (p. 1-6). The new *Supervisor Training Curriculum Outline (2.0)* will take effect in November of 2019. This module was divided into two main categories: a) supervision of ongoing services, and b) supervision of trainees. The first
COMPETENCY-BASED SUPERVISION IN ABA

category was further apportioned into six main items: “1) the supervisor should be able to state the purpose of supervision to the supervisee or trainee, 2) the supervisor should be able to describe the strategies and potential outcomes of ineffective supervision, 3) the supervisor should be able to prepare for the supervision relationship with the supervisee or trainee, 4) the supervisor should be able to establish a plan for structured supervision content and evaluation of competence for supervisees and trainees, 5) the supervisor should be able to create committed and positive relationships with supervisees or trainees, and 6) the supervisor should be able to use behavioral skills training to improve supervisee or trainee performance” (BACB, 2019 p. 3-8). The second category was also further apportioned into three main items: “1) the supervisor should be able to comply with relevant BACB fieldwork requirements when supervising a trainee pursuing BCBA or BCaBA certification, 2) the supervisor should be able to evaluate the effectiveness of supervision of the trainee, and 3) the supervisor should be able to incorporate ethics and professional development into supervision of trainees” (BACB, 2019 p. 8-12).

However, to date, there are no studies that address the attainment of supervision competency for clinicians in the field of ABA.

**Fidelity and Treatment Adherence**

In their article, Hogue, Liddle, and Rowe (1996) focused on new possibilities of treatment adherence research. First, the authors reviewed common methods of conducting adherence research and demonstrated that advanced-level procedures, which yield clinically rich information about therapist behavior in session, were congruent with the methods and goals that characterize traditional psychotherapy process research. Second, they provided both a theoretical framework and some practical recommendations for conducting advanced-level adherence research with family therapy models.
Building upon previous work, Hogue at al. (1998) used a treatment adherence and differentiation process measure to evaluate dynamic cognitive-behavioral therapy and multidimensional family therapy for adolescent substance abuse. Ninety full-length videos of treatment sessions were reviewed by non-participant raters. Results obtained from factor analysis of observational ratings demonstrated that high degrees of both treatment adherence and treatment differentiation were achieved by therapists conducting the two interventions. Therapists in each condition emphasized model-unique interventions, avoided model-proscribed interventions, and used a mixture of theoretically shared and facilitative interventions.

Additionally, the authors addressed challenges in conducting adherence research and implications of the results for advancing treatment development for adolescent drug users. Later, Hogue, Liddle, Singer, and Leckrone (2005) assessed fidelity in multidimensional family prevention (MDFP), in comparison to two other empirically-based treatments for adolescent drug abuse: multidimensional family therapy (MDFT), and cognitive-behavioral therapy (CBT). Videotapes of each treatment model were randomly selected and observationally rated along two dimensions of implementation: intervention parameters and intervention techniques. Results revealed that MDFP demonstrated adherence to its core intervention principles in comparison to MDFT, and in contrast to CBT. Overall, MDFP and MDFT produced almost identical profiles of intervention parameters, with each spending about half of every session focusing specifically on family issues.

Specifically, The Therapist Behavior Rating Scale—Competence (TBRS-C) was used to measure adherence and competence in individual cognitive–behavioral therapy (CBT) and multidimensional family therapy (MDFT) for adolescent substance abuse. Participants included 136 clinically referred adolescents and their families. During both the CBT and MDFT conditions, therapists’ adherence, competence, global competence, overall competence, skill, and responsiveness were assessed. Fidelity judges participating in one of two groups scored sessions’ videotapes and rated therapists’ performances. Results revealed initial reliability and validity findings for the TBRS-C, when applied to both CBT and MDFT. For both treatment models, the TBRS-C exhibited strong inter-rater reliability for adherence items, fair to poor reliability for individual competence items, but enough reliability for global competence ratings, and patterns of correlations among items that support construct validity. Variance in adherence, especially competence scores, was attributable more to client effects than to therapist effects. These results demonstrated that fidelity measures focusing on molar therapeutic goals and incorporating therapist responsiveness into the assessment of competence can demonstrate acceptable psychometric properties to complement a high level of clinical validity.

Subsequently, Hogue et al. (2008) examined the impact of treatment adherence and therapist competence on intervention outcomes of individual cognitive–behavioral therapy (CBT) and multidimensional family therapy (MDFT) for adolescents engaging in substance use and related behavior problems. Study participants included 136 adolescents (62 CBT, 74 MDFT). Participants were assessed at intake, discharge, and at 6-months after the intervention. Observational ratings of adherence and competence were collected on early and later phases of treatment by using a contextual measure of treatment fidelity. Adherence and competence effects were tested after controlling for therapeutic alliance. Results found that treatment adherence
predicted treatment outcome in manualized behavioral interventions. Treatment adherence was linked to improvement in multiple outcomes up to 6 months after discharge. Adherence promoted therapeutic change across the two different outpatient approaches. In CBT only, stronger adherence predicted greater declines in drug use. In CBT and MDFT, (a) stronger adherence predicted greater reductions in externalizing behaviors, and (b) intermediate levels of adherence predicted the largest declines in internalizing behaviors, with high and low adherence predicting smaller improvements. Contrary to hypotheses, therapist competence was not related to any outcome in either condition, or did it moderate the impact of adherence on outcome. The findings supported the contention that treatment adherence plays an important role in the success of empirically based behavioral interventions for adolescent mental health problems.

Lastly, Sheidow, Donohue, Hill, Henggeler, and Ford (2008) developed and conducted a preliminary evaluation of an adherence monitoring system for 27 community-based practitioners. The system assessed clinician fidelity to an evidence-based treatment for adolescents engaging in substance abuse. Session tapes were scored for adherence to a family-based approach to contingency management during baseline, post-workshop, and follow-up periods. Approximately half of the clinicians were randomized to receive intensive quality assurance following a family-based contingency management workshop as part of a larger study. Findings supported the clinical feasibility of the developed system.

Further studies on treatment fidelity were published in 2013 and 2015 spearheaded by the works of Hogue and colleagues. Namely, Hogue, Ozechowski, Robbins, and Barrett Waldron (2013) discussed instrumental shifts in operating practices and procedures that could fortify evidence-based practices (EBPs), and insulate against protocol drift and erosion, which may naturally occur in real-world settings. To this end, the authors described three innovations for
localizing EBP’s quality assurance: a) adaptation of observational fidelity methods for therapist self-report and supervisor observation of EBPs, b) process control benchmarking methods for continuous tracking of EBP fidelity strength and consistency, and c) development of intramural clinical expertise grounded in local management of EBP implementation and outcome data. According to the authors, these innovations exemplified a fundamentally empirical approach to sustaining quality EBP implementation in frontline settings.

Additionally, Hogue and Dauber (2013) described the use of a multi-method assessment of treatment fidelity to the family therapy (FT) approach demonstrated by front-line therapists in a community behavioral health clinic. Fifty adolescents with conduct and/or substance use problems were randomly assigned either to routine family therapy (RFT) or to a treatment-as-usual clinic not aligned with the FT approach. Observational analyses showed that RFT therapists consistently demonstrated adherence to the FT approach. In addition, analyses of therapist self-report measures found that compared to TAU, RFT demonstrated stronger adherence to FT and skill than did TAU therapists. However, as the authors noted, the study would have been enhanced by observationally measuring the competence (quality) of FT implementation, which has the potential to predict clinical outcomes over and above adherence (quantity).

Finally, Hogue, Dauber, Lichvar, Bobek, and Henderson (2015) assessed the reliability and accuracy of two groups of community therapists who reported on their use of family therapy (FT) and motivational interviewing/cognitive-behavioral therapy (MI/CBT) interventions during routine treatment of adolescents with conduct and substance use problems. This was achieved by comparing therapist ratings on the Inventory of Therapy Techniques for Adolescent Behavior Problems (ITT-ABP) to ratings made by non-participant observers. Forty-five participants were
COMPETENCY-BASED SUPERVISION IN ABA

randomized into two conditions: (a) Routine Family Therapy (RFT), or (b) Treatment As Usual (TAU). Therapists and trained observational raters provided FT and MI/CBT adherence ratings on 157 sessions (104 RFT, 53 TAU). Results revealed that overall, family therapists were fairly reliable with trained observers in reporting on the extent to which they implemented FT interventions, though unreliable in reporting on MI/CBT interventions. TAU therapists also displayed relatively stronger correspondence with observers for FT interventions than for MI/CBT, though their reliability was low for FT and non-existent for MI/CBT. Lastly, both groups of therapists overestimated the extent to which they implemented FT and MI/CBT interventions. These results offer support for the feasibility of using existing therapist-report methods to anchor quality assurance procedures for FT interventions in real-world settings. However, this was not the case for MI/ CBT.

BST: An Evidence-Based Method

Currently, BST holds a robust place in the behavioral literature. BST has been successfully utilized in the teaching of appropriate social behaviors, and social-vocational skills to individuals with intellectual disabilities and behavior challenges (Miltenberger, 2000; Burke, Andersen, Bowen, Howard, and Allen, 2010). Additionally, BST has been employed in the teaching of safety skills, including gun safety, as well as firearm injury and gun play prevention (Jostad, Miltenberger, Kelso, & Knudson, 2008; Kelso, Miltenberger, Waters, Egemo-Helm, & Bagne, 2007; Gross, Miltenberger, Knudson, Bosch, & Breitwieser, 2007). Moreover, BST has proven a valuable protocol to adhere to when teaching individuals with intellectual disabilities abduction and sexual abuse prevention skills (Johnson, et al., 2005; Egemo-Helm, Miltenberger, Knudson, Finstrom, & Jostad, et al., 2007; Lee & Tang, 1998). Lastly, BST has also been applied in the training of adults with intellectual disabilities to implement “mand” training and discrete
trials teaching (Lerman, Hawkins, Hillman, Shireman, and Nissen, 2013; Lerman, Hawkins, Hillman, Shireman, and Nissen, 2015). Indeed, the use of BST and its many applications for the teaching of new skills to individuals of varying ages and exceptionalities, has presented new and exciting avenues of research. As such, it appears to be a promising method for training supervisory competencies. Nonetheless, before embarking on novel research ventures, existing works on the efficacy of BST should be inspected. The following content will provide a more in-depth view of the current literature on BST, and its applications, and value.

**Using BST to Teach Appropriate Social Skills.** One study that illustrates the efficacy of the BST protocol was conducted by Miltenberger (2000). The author implemented discrimination training and BST procedures to teach appropriate social behaviors to a 13-year-old boy diagnosed with intellectual limitations and multiple behavior problems, who exhibited social skills deficits and social behavior problems. As expected, the participant showed improvements in his social behavior during the treatment sessions, as well as in the home and school environment, especially as his therapist, peers and staff maintained these newly acquired appropriate social behaviors across the aforementioned environments.

Another study that demonstrates the effectiveness of BST was conducted by Burke, Andersen, Bowen, Howard, and Allen (2010). In this study, the authors evaluated the efficacy of a vocational training program that included BST and a performance cue system (a patented iPhone application that was adapted to the study), to teach social-vocational skills to six individuals with Autism Spectrum Disorder (ASD). The cue system in particular was programmed to present a total of 63 textual cues, at a frequency of one cue per task. Cues were displayed on an iPod screen mounted at eye level inside a mascot. The application was programmed so that a research assistant could touch a text cue on the iPhone, which would then
display that same cue on the iPod. In experiment one, the participants were exposed to a training program that included BST components, and if necessary, for mastery, the implementation of the cue system. In the second experiment, participants were exposed to the cue system alone, and later the BST training if necessary, to achieve skill mastery. Results of the study revealed that all subjects acquired the required social-vocational skills when the cue system and BST were implemented.

**Using BST to Teach Manding and DTT.** A third example of the efficacy of BST when applied to individuals with intellectual disabilities was conducted by Lerman, Hawkins, Hillman, Shireman, and Nissen (2013). The authors found that BST was effective when used to train five adults with Asperger syndrome or pervasive developmental disorder not otherwise specified (PDD-NOS) on the implementation of mand training via incidental teaching, and on the implementation of discrete trial training (DTT). The study further demonstrated that the DTT skills generalized across untrained targets and children.

**Using BST to Teach Safety Skills.** Additionally, BST has been effectively used to teach safety skills, including gun safety to young children and individuals with intellectual disabilities. Jostad, Miltenberger, Kelso, & Knudson (2008) examined the use of a BST model in the training of six peers ages six and seven, as tutors to aid in decreasing the time and resources needed to teach gun safety skills to six younger peers ages four and five. Results of the study revealed that following BST, all peer trainers demonstrated the gun safety skills correctly, as well as maintenance of the skills, in naturalistic assessments. Furthermore, all six younger peers acquired the safety skills through training conducted by an older peer trainer.

In another study, Kelso, Miltenberger, Waters, Egemo-Helm, & Bagne (2007) evaluated the effectiveness of implementing BST and a program called Eddie Eagle, in the teaching of
firearm injury prevention skills to 30 children ages eight and nine. During the study, participants were randomly assigned to one of three conditions: the Eddie Eagle group, the BST group, or the control group. The researchers found that the children in both the Eddie Eagle and BST groups performed significantly better than the control group, in the identification of correct safety behaviors. Overall, both BST and the Eddie Eagle program were equally effective training procedures when paired with in-situ training for teaching firearm injury prevention skills to the 28 children who eventually completed the study. Although children needed various numbers of training sessions before acquiring the target skills, all the children eventually performed to criterion with in-situ training.

Furthermore, Gross, Miltenberger, Knudson, Bosch, & Breitwieser (2007), evaluated the use of BST coupled with an in-situ training program for parents to teach their own children safety skills to prevent gun play, in an effort to increase the efficiency of this training program. Results of the study showed that the training was effective for three of the four children. Additionally, results also revealed that the training program was implemented with fidelity by three of the four parents. As an added measure of social validity, the parents expressed their support for the treatment, stating that it was easy to implement, and that they were satisfied with the program.

**Using BST to Teach Prevention Skills.** BST has also been successfully utilized to teach abduction and abuse prevention skills to individuals with developmental disabilities. Johnson et al. (2005) examined the effectiveness of BST with in situ training for teaching abduction-prevention skills to 14 preschool children. Likewise, they evaluated the long-term maintenance of abduction prevention skills following training, to determine whether BST with in situ training was more effective than previously evaluated implementations of this protocol. As hypothesized,
all children acquired the required abduction prevention skills following BST with in situ training. Furthermore, all the children maintained the safety skills after one month.

Egemo-Helm, Miltenberger, Knudson, Finstrom, and Jostad, et al. (2007) evaluated a BST program in combination with in situ training to teach sexual abuse prevention skills to five women with mild to moderate mental retardation. Following BST with in situ training, all five women performed to criterion levels, especially on role-play assessments. Additionally, three of the five participants demonstrated generalization of sexual abuse prevention skills in the natural environment following one to two in situ training sessions. Lastly, three of the four participants maintained the skills at a one-month evaluation, and two of the participants maintained the skills after three months.

Lee and Tang (1998) evaluated the effectiveness of the BST program in the prevention of sexual abuse for 72 female Chinese adolescents diagnosed with mild mental retardation. Their aims also included exploring the presence of any negative effects resulting from the implementation of the BST program, and examining whether a sexual abuse prevention program designed for use with Western populations could feasibly be extended to Chinese female adolescents diagnosed with mental retardation. Participants were randomly assigned to either the control or the treatment groups and were presented with the Attention Control Program (Wurtele et al., 1992; as cited in Lee & Tang, 1998) or the BST program, respectively. Following a MANOVA (Tabachnick & Fidell, 2013; as cited in Lee & Tang, 1998) statistical analysis, results of the study revealed that Chinese female adolescents diagnosed with mild mental retardation did in fact improve their knowledge about sexual abuse and self-protection skills after participating in a sexual abuse prevention program. Specifically, the authors found that participants in the treatment group significantly increased their recognition of appropriate-touch requests when
compared to participants in the control group, though gains were not maintained after 2 months. No negative effects were documented. Overall, the authors concluded that with modifications, the BST sexual abuse prevention program implemented could be used with Chinese adolescents who have been diagnosed with mild mental retardation.

Tarasenko, Miltenberger, Brower-Breitwieser, and Bosch (2010) assessed the effectiveness of a peer-delivered BST model to teach typically developing children abduction prevention skills, using a multiple-baseline design across subjects. Results of the study demonstrated that students successfully utilized the BST model to teach their younger peers how to avoid a potential abduction.

Prior to this, Knudson et al. (2009) implemented a BST protocol to teach individuals with severe and profound mental retardation fire safety skills. Specifically, the aim of the study was to decrease the subjects’ house-exit latency upon hearing the fire alarm. Although only one subject was able to exit the house independently within the first three minutes after hearing the fire alarm, the training did decrease the number of prompts provided by the staff to exit the home.

The literature review of BST presented above offered support for the potential effectiveness of a modified BST approach for the development of a systematic protocol to use in supervision based on its proven effectiveness for various clinical uses.

**Research Questions**

1. Are standard (basic-level of competency) BCBA supervisory practices effective in improving supervisee behavior during DTT (see *Definitions and Terms* for a description of standard/basic supervisory level of competency)? Basic level supervisory competency was measured by the *Evaluation of Supervisor Performance* form (see Appendix C for a copy of this document) during pre-treatment (baseline) until stability was reached.
COMPETENCY-BASED SUPERVISION IN ABA

2. After the modified BST protocol is introduced, will BCBA supervisors demonstrate intermediate levels of competency (see Definitions and Terms for a description of intermediate and advanced supervisory levels of competency) during supervision sessions? Intermediate and advanced levels of competency were also measured by the Evaluation of Supervisor Performance form (see Appendix C for a copy of this document).

3. Does the introduction of the modified BST protocol affect the performance of supervisees receiving supervision from BCBAs during the implementation of DTT? Supervisees’ performances were measured by the Evaluation of Supervisee Performance form (see Appendix D for a copy of this document).

Hypotheses

First, it was expected that, after the introduction of the modified BST protocol, supervisors’ performances would improve from basic to intermediate levels. These improvements would indicate that supervisors would have demonstrated proficiency past merely completing a training, to validating their supervisory skill performance in vivo via observation, measurement, and evaluation. Scores of 80 percent or more on the Evaluation of Supervisor Performance form (see Appendix C for a copy of this document) indicated as well as validated improvements in supervisory performance from basic to intermediate levels.

Second, it was expected that supervisees’ performances would also improve as a result of improved supervisory performance. Collateral data collected on supervisees’ performances indicating scores of 70 percent or higher on the Evaluation of Supervisee Performance form (see Appendix D for a copy of this document) indicated as well as validated improved performance in the application of DTT skills by supervisees.
Chapter 3: Methodology

Introduction

It is important for any new intervention to be compatible with the context in which it will ultimately be used (Wolf, 1978; Schwartz & Baer, 1991). The intervention evaluated in this study was designed in response to a call from the BACB to employ effective supervision training systems for newly certified behavioral practitioners (BACB, 2012). To optimize interpretation of the results of this study by those who are most likely to use it (after the intervention turned out to be effective), the research methods used to conduct this study paralleled those commonly used in the field of ABA. Therefore, though the study will have potential implications for other disciplines such as clinical psychology, counseling psychology, mental health counseling, marriage and family counseling, as well as education among others, it focused on a special area within behavioral science. In so doing, the study offered new technology, and the potential to impact current environmental events in the area of ABA supervision and beyond.

The purpose of the proposed study was to evaluate the effectiveness of a behavioral skills training-based supervision-training program with ABA supervisors. The dependent variable was the observed and measured change in supervisor behavior with their supervisees. In addition, collateral data was collected on supervisees’ skills to determine if the introduction of the modified BST protocol affected their performances. This study was conducted following the single-case multiple baseline methodology outlined by Bailey and Burch (2018).

Participants

Participants included BCBA supervisors. Three BCBAs and two BCBA-Ds were recruited for this study. BCBA participants were selected if they had at least one year of experience in the implementation of DTT. Proficiency was demonstrated by achieving scores of 85% or better on
the *DTT Data Collection Sheet* (See Appendix G). BCBA supervisors had to also score 40% or less on the *Evaluation of Supervisor Performance* checklist (See Appendix C). The latter criterion was included as a means to target BCBAs who need enhanced supervision competencies. Because the researcher was attempting to make a distinction between training and supervision, the supervisors selected for this study had already met the BACB supervisor qualifications which included passing an eight-hour, competency-based training on supervising individuals working toward their certification in ABA, and passing an online competency-based training about the BACB’s experience standards found at the BACB website (Behavior Analyst Certification Board, 2015) before the onset of the supervisory relationship in accordance with the *BACB Professional and Ethical Compliance Code for Behavior Analysts* (BACB, 2016). A total of five BCBA supervisors participated in the study while accounting for attrition and disqualification. Exclusion criteria for BCBA participants included engaging in unethical behavior and/or having had complaints filed against them with the BACB or other regulatory company (e.g., insurance company), three or more unjustified absences during the study, verbally expressing refusal to implement the modified BST protocol (i.e., the independent variable), and being simultaneously employed by another ABA company different from the one in which the study was conducted. No BCBA participant was disqualified or withdrew from the study.

Each BCBA participant also worked with one or more therapists/supervisees. Supervisees were working toward a Bachelor’s or Master’s degree. They may also have earned either degree. In addition, supervisees had worked in the ABA field for no more than six months. Supervisee proficiency with DTT was demonstrated by scores of 80% or higher on the *DTT Data Collection Sheet* (See Appendix G) form. They also reached scores 40% or less on the *Evaluation of*
Supervisee Performance checklist (See Appendix D for a copy of this document). At least eight supervisees were recruited, to account for attrition or disqualification. Two of the supervisees recruited did not meet inclusion criteria (i.e., scored higher than 40%, on the Evaluation of Supervisee Performance checklist) and a third one participated in protocol generalization. A total of six supervisees participated in the study. Supervisees would have been excluded from the study had they met one or more of the following exclusion criteria: a) engaged in unethical behavior and/or had complaints filed against them with the BACB or other regulatory company (e.g., insurance company), b) accumulated three or more unjustified absences during the study, c) verbally expressed refusal to implement the feedback provided by the supervisor during BST supervision sessions, and d) also maintained employment with another ABA company different from the one in which the study was conducted. No supervisee participant was disqualified or withdrew from the study.

Each supervisee also worked with one or more children using DTT. Children were between the ages of two and 10. They had all been diagnosed with an intellectual disability, or Autism Spectrum Disorder (ASD) in addition to an intellectual disability. Furthermore, children participated in DTT for at least six hours a week. During DTT, children were taught a number of skills including requesting, visual performance, vocal and motor imitation, receptive language, expressive language, self-help skills, among others. Children’s data determined if they did not engage in self-injury, aggression, property destruction, elopement, or excessive crying more than three times per week. At least eight children needed approval to participate in the DTT sessions with the supervisee, to account for attrition or disqualification. Two of the children recruited were disqualified from the study, and three withdrew during the treatment phase due to reasons unrelated to the implementation of the independent variable. A total of six children participated
in the study. Meeting one or both of the following criteria excluded the children from the study. These included three or more unjustifiable absences throughout the duration of the study; and increases in frequency, duration, and/or intensity of challenging behaviors during the study.

**Instruments and Measures**

Two instruments were developed by the researcher. The first instrument, *Evaluation of Supervisor Performance* form (see Appendix C for a copy of this document), was used to evaluate the performance of the BCBA supervisors. In so doing, this form was used to answer the first and second research questions. Research question number one attempted to assess whether basic-level of competency supervisory practices are effective in improving supervisee behavior during DTT. Research question number two attempted to assess whether BCBA supervisors demonstrated intermediate levels of competency during supervision sessions after the introduction of the modified BST protocol. The second instrument, *Evaluation of Supervisee Performance* form (see Appendix D for a copy of this document), was used in the collection of collateral data on supervisees’ performances. As such, this form was used to answer the third research question. Research question number three attempted to assess whether the introduction of the modified BST protocol affected the performance of supervisees receiving supervision from BCBAs during the implementation of DTT.

Both instruments mentioned above were used during the pre-intervention (baseline) and intervention sessions and each yielded a percentage score that later was analyzed and presented in graphical display. Prior to instruments development, a review of 24 databases on “ProQuest” of both the behavioral and non-behavioral literature on supervision was conducted. This appraisal yielded no appropriate reliable or valid instruments that aided the researcher on the implementation of the study’s independent variable. Therefore, the researcher relied on
COMPETENCY-BASED SUPERVISION IN ABA

information included in relevant behavioral studies (i.e., Hanley, Jin, Vanselow, & Hanratty, 2014; Jin, Hanley, & Beaulieu, 2013; Perrin et al., 2011; Gillett & LeBlanc, 2007; and Wolf, 1978) for the development of the instruments.

**Data Forms Development.** For the development of *Evaluation of Supervisor Performance* form (see Appendix C for a copy of this document), the researcher selected specific elements of the *Supervisor Training Curriculum Outline* (Behavior Analyst Certification Board, 2012), entitled “Important Features of Supervision” (Behavior Analyst Certification Board, 2012 p. 2) and “Behavioral Skills Training” (Behavior Analyst Certification Board, 2012 p. 3), respectively. A more in-depth description of the development of this instrument is included in Appendix A.

The second instrument created for this study was the *Evaluation of Supervisee Performance* form (see Appendix D for a copy of this document). For the development of this instrument, the researcher initially relied on two research studies on discrete trials teaching (DTT). The first study by Catania, Almeida, Liu-Constant, and Reed (2009) utilized video modeling to train staff on the implementation of this fundamental element of behavior change in ABA. As part of the video utilized to demonstrate the DTT skills, a voiceover script was incorporated, which provided an explanation of each of the 10-modeled teaching steps. Additionally, a list of the skills was also supplied in the form of a table on the publication, from which the researcher selected some of the initial skills on the *Evaluation of Supervisee Performance* form (see Appendix D for a copy of this document).

The second study in which the researcher relied for the creation of the *Evaluation of Supervisee Performance* form (see Appendix D for a copy of this document) was conducted by Lerman, Hawkins, Hoffman, and Caccavale (2013). In their study, the authors used BST to pilot
COMPETENCY-BASED SUPERVISION IN ABA

a vocational program in which they taught adults with high-functioning Autism (HFA) how to conduct DTT. A more in-depth description of the development of this instrument is included in Appendix A.

Reliability of Developed Instruments. According to Kimberlin and Winterstein (2008), a research instrument should account for reliability errors. In this case, reliability errors found within the instruments were addressed during the pilot study, by refining each operational definition belonging to each of the skills listed under each data collection sheet. The researcher and three research assistants worked on refining the definitions, and separately recoded all session videos (i.e., assessed interrater reliability). Reliability data obtained during the pilot study on the use of the two instruments revealed overall scores of 80 to 100 percent for both data collection sheets (for more information regarding reliability results, see inter-observer agreement on table 2). Likewise, subsequent instruments’ retesting using repeated measures during the pilot also strengthened their reliability (Kimberlin & Winterstein, 2008). Though the instruments may likely need further refinement, results of the pilot study provided useful information regarding their internal consistency, as different observers scored the same behaviors/skills using the same instrument (Kimberlin & Winterstein, 2008). Also, instruments were used to measure behaviors once per week, which contributed to their stability (Kimberlin & Winterstein, 2008). Lastly, responsiveness to change on the part of the subject whose behavior is being measured was also assessed during the pilot study. Data revealed that the introduction of the modified BST protocol resulted in 40 percent positive behavioral changes. Accordingly, both subjects’ behaviors were responsive to change.

Validity of Developed Instruments. The proposed study served as validation of the instruments developed. Nevertheless, the description below provided insight into steps already
COMPETENCY-BASED SUPERVISION IN ABA

taken by the researcher to aid in the validation of the *Evaluation of Supervisor Performance* form (see Appendix C for a copy of this document), as well as the *Evaluation of Supervisee Performance* form (see Appendix D for a copy of this document). First, the researcher identified other research articles that studied the key variable: supervision (Korb, 2012). However, none of the literature reviewed provided an appropriate instrument that could potentially be used in this study. Second, after collecting and analyzing extensive video data over the course of one and a half years (i.e., from June 2013 through January of 2015), the researcher aided by a team of approximately ten behavior analysts developed construct definitions for each of the items under each of the forms mentioned above (Korb, 2012). These definitions may be found under the section of the same name on page 95. Additionally, the researcher relied on the *Supervisor Training Curriculum Outline* (Behavior Analyst Certification Board, 2012), as well as the articles referenced above (i.e., Catania, et al., 2009; and Lerman, et al., 2013) for the development of construct definitions. During this period, construct definitions were also refined into operational definitions (Korb, 2012). The researcher once again relied on the aid and knowledge-base of approximately ten BCBA's, some of whom are considered senior researchers and practitioners. The resulting operational definitions were structured into observational instruments, namely checklists encompassing 10 operational definitions each, which were chosen to be pilot-tested during the pilot study (Korb, 2012). Regarding construct validity, results of the pilot study revealed that the *Evaluation of Supervisor Performance* form (see Appendix C for a copy of this document) measured a single construct: supervision, and the *Evaluation of Supervisee Performance* form (see Appendix D for a copy of this document) also measured a single construct: supervisee performance. This type of item selection also contributed to the instruments’ stability (Kimberlin & Winterstein, 2008). Furthermore, though the internal
COMPETENCY-BASED SUPERVISION IN ABA

consistency of these instruments was not measured (other than the pilot, this is the first time the instruments will be used), a survey conducted by the researcher and completed by several BCBA supervisors from a local ABA company prior to the initiation of the pilot study, validated the items and definitions chosen for inclusion on the instruments. Concerning theoretical underpinnings, results of the pilot study also revealed that the instruments behaved according to theoretical propositions about the constructs (Korb, 2012). In this case, both instruments registered improvements on supervisors and supervisees’ performances. The instruments’ content validity was also examined by the researcher (Korb, 2012). After surveying other BCBA in the field, as well as relying on her own clinical experience to ensure that the instruments adequately covered the entire content area that they should cover, the Evaluation of Supervisor Performance form (see Appendix C for a copy of this document) and the Evaluation of Supervisee Performance form (see Appendix D for a copy of this document) were finally approved and pilot tested. Moreover, the researcher conducted informal surveys at the end of the pilot, which revealed that both users and takers thought the instruments were worth using. Overall, the information gathered by the researcher also provided substantiation towards the instruments’ face validity. They in fact appear to measure what they were developed to measure (Korb, 2012).

**Supplementary Instruments.** The researcher developed two additional instruments. The first additional instrument (i.e., the DTT Data Collection Sheet form; see Appendix G for a copy of this document) was used to evaluate the performances of BCBA supervisors and supervisees prior to inclusion in the study. The second additional instrument (i.e., the DTT Performance Feedback Sheet form; see Appendix E for a copy of this document) was used by BCBA
supervisors as a job aid during the intervention, in order to help them structure the feedback that was provided to supervisees during the supervision sessions in the treatment phase of the study.

For the development of the *DTT Data Collection Sheet* form (see Appendix G for a copy of this document), the researcher once again relied on the two studies listed above (i.e., Catania, et al., 2009; and Lerman, et al., 2013). Please see Appendix B for a mode in-depth description of the development of this instrument.

For the development of the *DTT Performance Feedback Sheet* form (see Appendix E for a copy of this document), the researcher incorporated all of the elements encompassed within the *Evaluation of Supervisee Performance* form (see Appendix D for a copy of this document), plus nine more (i.e., safety related, ethical, school/setting policy, acquisition plan, repeated errors, appropriate affect, generalization of skills, maintenance of skills, and problem solving), and organized them in the order in which the BCBA supervisor addressed them with the supervisee during the treatment phase of the study. This hierarchy divided the form into four categories. The first category encompassed issues that the BCBA prioritized first when providing feedback to the supervisee, and included safety related issues, ethical issues, school/setting policy issues, acquisition plan issues, and repeated errors that the supervising BCBA observed the supervisee execute during a given DTT session. These first five issues were deemed of most importance by the researcher when prioritizing and preparing the feedback to be provided to the supervisee. The second and third categories of the *DTT Performance Feedback Sheet* form (see Appendix E for a copy of this document), encompassed the ten elements listed within the *Evaluation of Supervisee Performance* form (see Appendix D for a copy of this document), and the fourth category – to be addressed last, encompassed appropriate affect, generalization of skills, maintenance of skills, and problem solving skills. These last four skills were deemed as “soft skills” by the researcher,
and therefore less critical (though important nevertheless) when prioritizing and preparing the feedback provided to the supervisee. Once again, the researcher surveyed several BCBA supervisors in the field regarding higher-level skills within DTT that may make the difference between initially trained skills, and skills that develop after ongoing supervision. *Appropriate affect, generalization of skills, maintenance of skills, and problem-solving* skills, though deemed higher-level skills, are not typically prioritized in clinical practice as much as the skills listed in categories one through three.

**Reliability of Supplementary Instruments.** The supplementary instruments’ reliability was addressed much in the same way as the previously mentioned instruments. Documents were analyzed by a team of behavior analysts, and reliability errors found within were also addressed during the pilot study, by refining each operational definition belonging to each of the skills listed under each data collection sheet Kimberlin and Winterstein (2008). The researcher and three research assistants worked on refining the definitions. Because data on the use of supplementary instruments were not relevant to the results of the pilot study (i.e., the *DTT Data Collection Sheet* form was used to select appropriate BCBA participants and supervisees for inclusion in the study, and the *DTT Performance Feedback Sheet* form was used as a job aid when providing feedback to the supervisee), reliability data on these documents were not collected. However, the collection of reliability data on these forms was included as part of the proposed study. Also, results of the pilot study provided useful information regarding their internal consistency, as different observers agreed on the same behavioral definitions within each instrument (Kimberlin & Winterstein, 2008). Additionally, one instrument was used to provide feedback to the supervisee once per week, which contributed to its stability (Kimberlin & Winterstein, 2008). Lastly, responsiveness to change on the part of the subject whose behavior
was being measured was also assessed during the pilot study. Data revealed that the introduction of the *DTT Performance Feedback Sheet* form as part of the modified BST protocol resulted in 40 percent positive behavioral changes for the supervisee. Accordingly, the supervisee’s behaviors were responsive to change.

**Validity of Supplementary Instruments.** Like the previously described documents, the proposed study served as validation of the supplementary instruments developed. Not unlike the description provided above, the researcher first identified other research articles that studied the key variable: DTT (Korb, 2012). However, none of the literature reviewed provided an appropriate instrument that could potentially be used in this study. After collecting and analyzing extensive video data over the course of one and a half years (i.e., from June 2013 through January of 2015), the researcher aided by a team of approximately ten behavior analysts developed construct definitions for each of the items under each of the forms mentioned above (Korb, 2012). These definitions may be found under the section of the same name on page 95. Additionally, the researcher relied on the articles referenced above (i.e., Catania, et al., 2009; and Lerman, et al., 2013) for the development of construct definitions. During this period, construct definitions were also refined into operational definitions (Korb, 2012). The researcher once again relied on the aid and knowledge-base of approximately ten BCBAs, some of whom are considered senior researchers and practitioners. The resulting operational definitions were structured into observational instruments, namely checklists encompassing seven operational definitions under the *DTT Data Collection Sheet* form, and four major categories encompassing four to five items under the *DTT Performance Feedback Sheet* form, that were chosen to be pilot-tested during the pilot study (Korb, 2012). Regarding construct validity, results of the pilot study revealed that the *DTT Data Collection Sheet* form (see Appendix G for a copy of this
document) measures a single construct: performance during DTT, and the *DTT Performance Feedback Sheet* form (see Appendix E for a copy of this document) also measures the same construct: supervisee performance during DTT. In addition, multiple items on each data collection sheet were used to measure different behaviors/skills as part of the same construct: supervisee performance during DTT. This type of item selection also contributed to the instruments’ stability (Kimberlin & Winterstein, 2008). Furthermore, though the internal consistency of these instruments was not measured (other than the pilot, this is the first time the instruments were used), the researcher surveyed several BCBA supervisors in the field regarding appropriate number of skills within DTT that supervisors demonstrated fluency on before engaging in the supervision of this procedure, as well as types of feedback that should be prioritized during the feedback-provision part of supervision. All surveyed supervisors validated the items and definitions chosen for inclusion on the instruments. Concerning theoretical underpinnings, results of the pilot study also revealed that the instruments behaved according to theoretical propositions about the constructs (Korb, 2012). In this case, both instruments aided in improving the supervisee’s performance. The instruments’ content validity was also considered by the researcher (Korb, 2012). After surveying other BCBAs in the field, as well as relying on her own clinical experience to ensure that the instruments adequately covered the entire content area that they should cover, the *DTT Data Collection Sheet* form (see Appendix G for a copy of this document) and the *DTT Performance Feedback Sheet* form (see Appendix E for a copy of this document) were finally approved and pilot tested. Moreover, the researcher conducted informal surveys at the end of the pilot, which revealed that both users and takers thought the *DTT Performance Feedback Sheet* form was worth using. Overall, the information gathered by
the researcher also provided substantiation towards the instruments’ face validity. They in fact appear to measure what they were developed to measure (Korb, 2012).

**Questionnaire Development.** Lastly, the researcher developed one questionnaire document to be used as a social validity measure. The *Supervision Social Validity Questionnaire* (see Appendix H for a copy of this document) was composed of five questions. For the development of this questionnaire, the researcher relied on the studies discussed below under the section labeled *Social Validity*. The first two questions formulated for this instrument served as subjective evaluative tools for the improvement of supervisory skills and continued use of this protocol. The following two questions addressed whether supervisors would recommend the intervention and thought that all the skills were important to achieving and maintaining excellent supervisee performance. On the last question, supervisors were asked to list any skills they did not deem important for a supervisor to demonstrate.

Once again, a review of both the behavioral and non-behavioral literature on supervision yielded no appropriate reliable or valid instruments that could aid the researcher on the compilation of social validity data specific for this type of study. Therefore, the researcher relied on information included in relevant behavioral studies for the development of the instruments.

**Procedures**

**Methodology and Research Design.** A quantitative, single-case, multiple-baseline across supervisors’ research design (Copper, Heron & Heward, 2007; Bailey & Burch, 2018) was used to evaluate the effectiveness of a BST-based supervision-training program with ABA supervisors.

**Data collection procedures.** Baseline supervision sessions attempted to answer research question number one: were standard (basic-level of competency) BCBA supervisory practices
COMPETENCY-BASED SUPERVISION IN ABA

effective in improving supervisee behavior during DTT (see Definitions and Terms for a description of standard/basic supervisory level of competency)?

Sessions took place as they typically would in clinical practice, where the supervisee was engaged in teaching various skills to a child, in the presence of the BCBA. Teaching sessions consisted of a supervisee and a child sitting at the table facing each other either directly across or diagonally or sitting beside each other. They lasted anywhere from half an hour to three hours with other transitional activities and breaks in between (for sessions where the children were seated at a table for more than 30 minutes) and generally encompassed the supervisee preparing materials and electronic data collection devices, as well as in some cases conducting mini preference assessments and probe trials for the first part of each session to determine which targets had been mastered, which targets the client may need to work on, as well as the child’s readiness to acquire new target responses. Next, the supervisee engaged in the teaching component of DTT for the remainder of the session, interspersing maintenance and target skills trials; plus, data collection activities either throughout or during the last few minutes of the session. The BCBA supervisor observed the supervisee and provided feedback throughout the DTT session. An average of 20 minutes was filmed for each DTT session where the supervisor provided feedback to the supervisee in the manner described above. Video data were coded focusing on the supervisor’s performance using the Evaluation of Supervisor Performance form (see Appendix C for a copy of this document). Baseline data were collected once a week for approximately three weeks, or until the graphic display of the data showed stability. Afterward, the treatment phase was implemented. Stability was determined by visual inspection of the data. If the researcher could begin to predict where the next data point would be, the data were considered stable (Bailey & Burch, 2018). This also meant that the data could be trending downward (i.e., stable downward trend), prior to the
implementation of the independent variable (Bailey & Burch, 2018). Similarly, a steady state, when it refers to data, is a pattern of responding that illustrates relatively slight variations in its measured dimensional quantities over some time period (Johnston & Pennypacker, 2009). Stability of the data then, refers to the fluctuation in the participant’s performance over time. Generally, “the greater the variability in the data, the more difficult it is to draw conclusion about the effects of the intervention” (Kazdin, 2011, p. 126).

**Uber training sessions.** These sessions encompassed the actual implementation of the independent variable (i.e., the Modified BST Protocol) and took place between the researcher and the BCBA participants, on the same day as the regularly scheduled DTT sessions, approximately one and a half hour prior, in a room designated for the purposes of this training/meeting. During the Uber sessions, the researcher (AKA the “Uber Supervisor”) first described to BCBA participants how the supervision sessions would occur (i.e., how BCBAs and trainers should be seated, what materials should be used, where the computers should be placed, and where the camera will be placed). Subsequently, BCBA participants were taught how to use the modified BST protocol checklist, as well as the *DTT Performance Feedback Sheet* form (see Appendix E). The researcher followed the protocol outlined on the *Evaluation of Supervisor Performance* checklist (see Appendix C) while training the BCBAs. She was prepared with documented feedback to provide to the supervisor, an uploaded 10-40-minutes video of the last supervision session between the BCBA supervisor and the trainer, a pen, paper, and relevant forms. They were also seated side-by-side facing a computer. First, the researcher went over the *DTT Performance Feedback Sheet* form (see Appendix E) with the BCBA. Next, the researcher prompted the BCBA to take notes, watched the relevant footage with the BCBA, provided coaching and/or modeling, opportunities for role-play/practice, and positive reinforcement where
COMPETENCY-BASED SUPERVISION IN ABA

Appropriate. Role plays were repeated until the BCBA had successfully demonstrated the new skills. The last two steps consisted of the researcher reminding the BCBA to engage in the new skills within the upcoming supervision session with the supervisee and scheduling the next Uber Training session (follow-up appointment) if appropriate. At this point, the Uber Training session ended. These training sessions lasted for approximately 20 minutes each and were filmed, for purposes of treatment integrity. Focusing on the behavior of the researcher, treatment integrity data were collected on each session. Results revealed that two to five Uber Training sessions were be conducted per BCBA participant.

**Booster Sessions.** If BCBA supervisors’ performances declined below 100%, following protocol maintenance, booster sessions were conducted in the same manner as the Uber Training Sessions described above. No booster sessions were conducted during the Pilot Study. Booster sessions were intended to help BCBA supervisors’ performances maintain the highest scores in intermediate levels of competency. A total of four booster sessions were carried out throughout the study, one for each of four out of the five BCBA participants.

**Modified BST sessions.** These sessions attempted to answer research question number two: after the modified BST protocol was introduced, did BCBA supervisors demonstrate intermediate levels of competency (see *Definitions and Terms* for a description of intermediate supervisory levels of competency)? During the intervention, Modified BST supervision sessions occurred an hour before each scheduled DTT session in the same room as the Uber training sessions, where the BCBA and supervisee were also seated next to each other, facing a computer. In preparation for this meeting, the BCBA supervisor has already uploaded a five-minute video sample of the previous DTT session that took place with the supervisee and the client, gathered relevant materials needed for role play, as well as paper and a pen. BCBA participants used the
COMPETENCY-BASED SUPERVISION IN ABA

*DTT Performance Feedback Sheet* form (see Appendix E) as an aid to help prioritize the feedback they were providing to the supervisees. First, the BCBA addressed with the supervisee the first (highest priority) skill noted on the *DTT Performance Feedback Sheet* form (see Appendix E). The BCBA provided corrective feedback on that skill, prompted the supervisee to take notes, watched the relevant footage with the supervisee, provided coaching and/or modeling, opportunities for role-play/practice, and positive reinforcement where appropriate. Role plays were repeated until the supervisee had successfully demonstrated the corrected skill. Next, the BCBA addressed the second skill noted on the *DTT Performance Feedback Sheet* form (see Appendix E), following the steps outlined above. Once again, role plays were repeated until the supervisee had successfully demonstrated the second priority skill. The last two steps consisted of the BCBA reminding the supervisee to engage in the corrected skills within the upcoming DTT session and scheduling the next supervision session (follow-up appointment). At this point, the Modified BST session ended. Supervisors were evaluated on their adherence to the *Evaluation of Supervisor Performance* checklist (see Appendix C) until a greater positive effect on supervisor’s performance was achieved for three consecutive sessions or more, at which point the intervention ceased. A greater positive effect on supervisors’ performance was defined by a satisfactory score of 80 percent or higher on the *Evaluation of Supervisor Performance* (see Appendix C for a copy of this document) checklist when compared to basic performance (see p. 94 for a definition of basic supervisory performance level) values obtained on the same instrument during the pre-treatment phase. This percentage was selected after careful consideration of the data analyzed during a pilot study (see *Pilot Study* below for more information) conducted by the researcher, as well as the current passing score for the BCBA examination (BACB, 2015). Specifically, results of the pilot study revealed supervisor’s scores
at 50 percent during the baseline phase. In accordance with common-knowledge current academic guidelines that require a score of at least 80 percent or better (equivalent to a grade of B) in order to successfully complete graduate work, as well as the passing grade on the BCBA examination of 76 percent or better (BACB, 2015), it was determined that supervisors had to achieve a score of at least 80 percent in order to attain mastery, thereby matching both the academic standard and BACB criteria for the mastery of supervisory skills. As such, when supervisors achieved this increase in performance values during the intervention phase, they were considered as having attained moderate to advanced levels of competency (see p. 94 for a definition of moderate and advance supervisory performance levels) on these particular sets of supervisory skills. Throughout the study, supervisors demonstrated intermediate levels of performance for an average of eight sessions, ranging between four and 13 sessions.

Similar to the Uber sessions, the BST supervision sessions also had a duration of approximately 20 minutes. The decision regarding the duration of the films for both the training and supervision sessions was primarily influenced by the current BACB requirement of an hour of supervision per supervisory period, where a supervisory period is defined as one or two weeks depending on the experience category (e.g., practicum or independent fieldwork; BACB, 2015). Because each hour of supervision also includes observation, the current study proposed as part of the Modified BST Protocol, that BCBA supervisors engaged in at least 30 minutes of observation, documentation and evaluation of trainer’s performance; 10 minutes of feedback preparation including prioritizing corrective feedback, and 20 minutes meeting with the supervisee, during which time the supervisor provided appropriate positive and corrective feedback, as well as opportunities for role play and skill practice. Therefore, the researcher proposed to name this protocol The 30-10-20 Method of Competency-Based Supervision.
COMPETENCY-BASED SUPERVISION IN ABA

Since the proposed study assessed the use of a modified BST protocol for the training and evaluation of BCBA participants’ supervisory skills, the selection and subsequent modification of the BST protocol made by the researcher were primarily based on *The Supervisor Training Curriculum Outline* (Behavior Analyst Certification Board, 2012) published for the purposes of training current and new BCBAs on evidence-based supervisory practices. This document includes eight areas in the form of a curriculum outline encompassing a task list of the “knowledge and skills required for the effective supervision of those who deliver behavior-analytic services and those who are pursuing BACB certification” (Behavior Analyst Certification Board, 2012 p. 1). Specific modifications to the BST protocol included observing the supervisee, prioritizing corrective feedback, prompting note-taking, prompting supervisee to engage in relevant skills, and scheduling a follow-up appointment. The proposed study concentrated on use of the modified BST protocol for the supervision of DTT, which is a very specific area of ABA.

**Collateral Data.** Collateral behavioral data were collected on the performances of the supervisees/therapists during the study in an attempt to answer research question number three: did the introduction of the modified BST protocol affect the performances of supervisees receiving supervision from BCBAs during the implementation of DTT? Supervisees’ performances were evaluated using the checklist developed and tested by the researcher. The *Evaluation of Supervisee Performance* checklist (see Appendix D for a copy of this document) is composed of 10 skills. Supervisees’ performances were determined by dividing the number of steps performed correctly by the total number of steps, multiplied by 100, which yielded a percent correct performance result. Supervisees were evaluated on their adherence to the *Evaluation of Supervisee Performance* checklist (see Appendix D for a copy of this document)
until a greater positive effect on their performance was demonstrated throughout three consecutive DTT sessions, at which point data collection ceased. During this phase, supervisees demonstrated higher levels of performance when compared to baseline levels for an average of eight sessions, ranging between four and 13 sessions. Supervisees’ performances indicated a greater positive effect when their scores reached 70 percent or higher on the *Evaluation of Supervisee Performance* (see Appendix D for a copy of this document) form when compared to pre-intervention values reached on the same instrument. Comparable to the justification provided above for a greater positive effect on supervisor performance, this percentage was selected after careful consideration of the data analyzed during the pilot study, as well as the current passing grade for the BCaBA examination. Results of the pilot study revealed supervisee’s scores at 20 percent or less during the baseline phase. Current common-knowledge academic guidelines require a score of at least 70 percent or better in order to successfully complete undergraduate work, as well as a score of 76 percent or better in order to pass the BCaBA certification exam (BACB, 2015). Taking these guidelines into account, it was determined that scores of at least 70 percent would match similar standard criteria for mastery of trainer skills on the implementation of DTT. Therefore, supervisees who achieved this increase in performance values during the intervention phase, were considered as having attained a higher competency on these particular sets of DTT skills.

**Dependent Variable.** The dependent variable included the BCBA supervisors’ implementation of the modified BST protocol. It was measured as the number of skills performed correctly, and reported as a calculation of percent correct performance, by dividing the number of skills performed into the total number of skills required, and multiplying by 100.
Ensuring Data Quality and Observer reliability. Inter-observer agreement (IOA; Bailey & Burch, 2018; Cooper, Heron & Heward, 2007; Teddlie & Tashakkori, 2009) checks were conducted on 100 percent of the all the sessions (i.e., baseline and Modified BST sessions) for the BCBA supervisors and supervisee participants. Two (or more) different research associates previously trained on how to use the Evaluation of Supervisor Performance (see Appendix C for a copy of this document), and the Evaluation of Supervisee Performance (see Appendix D for a copy of this document) checklists, separately evaluated their scoring of the same videotapes until they reached 80 percent or better item-by-item agreement. Then, item-by-item IOA was calculated by dividing the items in which both research associates agreed, over the total number of agreements plus disagreements, and multiplying the result by 100. If two or more research associates reached at least 80 percent agreement or better on 33 percent of the sessions, the data were considered reliable. As previously mentioned, IOA (Bailey & Burch, 2018; Cooper, Heron & Heward, 2007; Teddlie & Tashakkori, 2009) was computed on collateral data collected on trainers’ performances in the same manner as the BCBA supervisors’ data.

Selecting one third of the data to conduct inter-observer reliability checks is a common practice in ABA research. Currently, peer reviewed journals such as The Journal of Applied Behavior Analysis, and The Journal of the Experimental Analysis of Behavior endorse this measure of data quality.

Procedural Fidelity. The Uber Supervisor Training Protocol Fidelity Checklist (see Appendix F for a copy of this document), containing a detailed task analysis of the behaviors in which the researcher had to engage during training, was used in order to evaluate and certify that the intervention was implemented according to plan. Additionally, the checklist served as a means to ensure standardization of the intervention across all BCBA participants. Every training
video was analyzed for treatment integrity adherence. Research assistants unrelated to the training sessions scored all the intervention videos. The checklist was developed and tested by the researcher skills.

**Social Validity.** Social validity data were collected after the conclusion of the study. Social validity questionnaires have been utilized in the ABA literature to assess whether participants perceived the treatment outcomes to be meaningful (Hanley, Jin, Vanselow, & Hanratty, 2014). Wolf (1978) further highlighted that relevant judges need to be employed in assessing the social importance of a study to ensure that it has value to society. With this in mind, there were several reasons why supervisors were provided with the opportunity to weigh in on the social validity of this study. First, the use of a modified BST protocol directly affects the behaviors of supervisors. As such, supervisors were provided with the opportunity to voice their opinions regarding the helpfulness of the intervention. The proposed treatment protocol affected the performances of supervisors and supervisees. Therefore, as stakeholders, participants were provided with opportunities to assess: a) the social significance of the research aims, b) the social appropriateness of the procedures, and c) the social importance of the effects (Wolf, 1978). One of the social validity checkpoints that ABA has is the requirement to meet the dimension of being “applied”, such that the behavior targeted is socially significant (Baer, Wolf & Risley, 1968). This means that the behavior targeted for change has importance to the participants and society. Perhaps one of the only ways to know this is to, “…Ask them.” (Finney, 1991, p. 245).

After the conclusion of the study, BCBA participants were asked to complete a five-question *Supervision Social Validity Questionnaire* (see Appendix H for a copy of this document). The questions were aimed at evaluating whether supervisors: a) thought the
COMPETENCY-BASED SUPERVISION IN ABA

intervention was helpful in improving their supervisory practice, b) would continue to use the protocol in their day-to-day supervision, c) would recommend the intervention to other BCBA supervisors, d) thought that all the skills were important to achieving and maintaining excellent therapist performance, and e) list any skills that did not seem important for a supervisor to demonstrate. Except for the last question, supervisors had two answer choices (on every question) from which to select: yes, and no. Following completion of the questionnaires, the researcher descriptively analyzed each questionnaire and developed a summary of the results.

Internal and External Validity Threats. In single-case experimental studies, there is almost always a high level of internal validity because subjects serve as their own control (Huitt, Hummel, & Kaeck, 1999). Nevertheless, based on a review of the behavioral literature, the behaviors examined were considered to be suitable representations of supervisory skills (Kazdin, 2011). Additionally, basic-levels of supervisory skills represented performance under normal conditions of supervisory practice (Kazdin, 2011).

Observational bias was reduced via measures of reliability (Kazdin, 2011). IOA checks were conducted on over 33% of the sessions (i.e., baseline and modified BST sessions) for the BCBA supervisors and was also computed on collateral data collected on supervisees’ performances (Bailey & Burch, 2018; Cooper, Heron & Heward, 2007; Teddlie & Tashakkori, 2009). Lastly, observational measures reflecting change aimed at objectively representing meaningful effects on the behaviors of supervisors and supervisees (Kazdin, 2011).

Conversely, the external validity of these type of single-case studies is considered to be extremely low (Huitt, Hummel, & Kaeck, 1999). To this end, these studies seek to acquire external validity through replication and generalization across different populations, settings, and
COMPETENCY-BASED SUPERVISION IN ABA

situations, among others (Huitt, Hummel, & Kaeck, 1999). To address this concern, the researcher conducted the study at five separate clinic-based settings.

**Recruitment Procedures.** Prior to proposal and Institutional Review Board (IRB) approval the researcher selected two local ABA companies that employed a minimum of 10 BCBAs each, on a full-time basis (BCBAs who work at least 30 billable hours per week), and provided each company official with a written description of the research, as well as information regarding inclusion criteria for both types of participants (see Appendix J for a copy of this document). In addition, the researcher requested a signed *Site Approval Letter* from each company official (see Appendix J for a copy of this document). After IRB approval, a formal meeting was requested via email, with each company’s official designee, in order to begin recruitment of participants. On the day of the meeting, the researcher requested that a recruitment email comprising a written description of the study and the researcher’s contact information was sent to all BCBAs, supervisees, and parents within the company (see Appendix K for a copy of this document). The researcher also requested that all potential participants contact her so that she may provide further clarification of the research aims, as well as develop a database comprising scheduling, availability, and contact information for each potential participant. This database was kept in the researcher’s computer, which is password protected, within a password protected file, in the researcher’s home. From the participants’ contact, the researcher identified BCBAs, supervisees, and children who met inclusion criteria for the study. In the case of BCBAs, the researcher conducted in-person and phone interviews, as well as verified their credentials and years of experience via the BACB public database. Additionally, the researcher conducted in-person interviews with supervisee participants, as well as parents, in order to obtain relevant inclusion information for the study. This information included supervisees’ tenure in the
COMPETENCY-BASED SUPERVISION IN ABA

field of ABA, as well a specific behavior reduction goals the children were attempting to achieve. Though children were not active study participants and no data were collected on their performance, it was important that those who were working with the supervisees had demonstrated the ability to acquire new skills via DTT for an average of at least 10 hours a week, every week, for a period of four months of longer. Furthermore, children working with supervisees did not engage in any of the following behaviors more than twice a week: self-injurious behavior, aggression, property destruction, elopement, or excessive crying (lasting for a period of 10 minutes or more) during each DTT session. Ensuring that the former criteria were met helped facilitate supervisees’ ability to demonstrate newly shaped DTT skills.

Once potential supervisor subjects were identified, the researcher contacted each BCBA via email, attached a BCBA consent form (see Appendix L for a copy of this document), and requested a videotape of the supervisor engaging in the implementation of DTT as well as information regarding the number of years she had practiced as supervisor. After the BCBAs met inclusion criteria for the study and returned the signed consent form to the researcher, she scheduled a second meeting with each supervisor to further discuss their role within the study, address any questions/concerns, and reach an agreement regarding dates and times to begin baseline data collection.

A similar procedure was utilized when recruiting supervisees. The researcher contacted each previously identified potential participant in person, handed presented them with the supervisee consent form (see Appendix M for a copy of this document), and requested a videotape of the supervisee engaging in the implementation of DTT. After the supervisees had met inclusion criteria for participation in the study, the researcher scheduled a second meeting
with each to further discuss their role within the study, address any questions/concerns and ultimately come to an agreement regarding dates and times to begin collecting baseline data.

After prospective children participants were identified, the researcher with the help of administration contacted each parent, and provided a consent form for disclosure of client information that also included a section for child assent (see Appendix N for a copy of this document), and requested access to the child’s data for the last two months. Though the collection of data to appraise child performance was not a principal aim of this study, consent forms were solicited for the selection of child participants who were videotaped along with supervisees during each DTT session. To this end, the researcher needed to review children’s records to ensure that the selected few meet behavioral participation criteria. After the children had met inclusion criteria for participation in the study, the researcher reached out to the supervisees and BCBAs who were part of each child’s treatment team and had expressed interest in participating in the study. Once all participants had signed consent forms and start dates were agreed upon, the study commenced.

**Timeline.** Based on the results of the pilot study, it was estimated that the study would be completed in a period of approximately four months. The researcher was estimating a period of two to three weeks for the recruitment of participants, approximately 12 weeks for both baseline data collection and the implementation of the independent variable, and approximately one week for the completion of the social validity questionnaires. Due to site cancellations related mainly to children and supervisees’ illnesses, the study ran for approximately eight months.

**Ethical Considerations and Review.** After this proposal was approved by the researcher’s doctoral committee, the researcher submitted an application to the university’s IRB committee for review and approval. Following IRB approval, the researcher begun recruitment
COMPETENCY-BASED SUPERVISION IN ABA

efforts. From this point forward, the researcher followed strict HIPPA guidelines. Baseline and treatment data collection were kept confidential in the researcher’s computer, which was password protected, within a password protected file, in the researcher’s home until the completion of the study. Unidentified data were shared only with research assistants and Co-PIs. Collateral data were also kept confidential and de-identified. Three years after the conclusion of the study, the researcher will destroy all documents containing identifiable information. All study-related files will be deleted from the researcher’s computer including videotapes, clients’ program data, data collection sheets, and questionnaires.

For children nonparticipants, the participation consent indicated that involvement in the research study was completely voluntary and had no bearing on the continuation of services currently being provided to the client by the company. Parents and children were also advised that they may withdraw from the study at any time, without any adverse consequences. Additionally, a similar message was conveyed to both trainers and BCBAs. The latter ones were informed in writing on the participation consent form that their involvement was completely voluntary and had no bearing on their continued employment. They were also advised that they could withdraw from the study at any time, without experiencing adverse consequences. Furthermore, they also were informed that the data obtained from the study was in no way utilized as job-related evaluative data for incentives or promotions.

Data analysis procedures. Single-case research designs have a distinguished history in psychological research, demonstrating their enduring value (McKay, 2008). They rely primarily on experimentation and visual inspection of data, and are characterized by increasing levels of scientific rigor, as well as by the strength of the inferences that can be made with their use. Contrary to group design, all that is needed is controlled variation of the independent variable
and measurement of the effects of this variation. This allows the experimenter to examine causal relations between intervention and outcome (McKay, 2008).

The most essential methodological requirement for single-case experimental designs is repeated assessments of the dependent variable across all phases of the study (McKay, 2008). This measurement should use valid and reliable measures, and should occur early, frequently, and consistently. Assessing early is important to establishing a baseline, or pre-intervention level of responding. Assessing frequently is important to capture change as it was occurring during the intervention. Finally, using consistent assessment methods across assessment opportunities is necessary to rule out the possibility that behavioral changes were the result of changes in assessment procedures (McKay, 2008). Another essential methodological requirement for single-case experimental designs is the use of a specified experimental manipulation (McKay, 2008).

In their review article Byiersa, Reichlea, and Symons (2012) provided further support for the use of this design. The authors first discussed the role of single-case experimental designs in evidence-based practice, followed by a discussion on analyzing effects in these experimental designs including the desirable qualities of baseline data, and visual data inspection. Subsequently, the authors discussed the areas of quality and replications for external and internal validity before describing the different types of single-case experimental designs. Among the designs reviewed were the AB, withdrawal, multiple baseline, multiple probe, changing criterion, as well as multiple and alternating treatment designs. As part of their review, the authors examined advantages and disadvantages of each design. On the last section of the article, Byiersa et. al. (2012), discussed how results are evaluated in single-case experimental designs. Specifically, they examined the issue of when, if ever, the data generated from single-case experimental designs should be statistically analyzed. To approach this issue, Byiersa et. al.
COMPETENCY-BASED SUPERVISION IN ABA

(2012) broke it into four parts: effect detection, magnitude of effect, quality of inference, and practice and data-based decisions. They concluded that: single-case experimental designs “can address the effects that intervention strategies and environmental variables have on performance at the individual level” (p. 21). Furthermore, Byiersa et. al. (2012) stated that this design allows for flexibility in a study to adjust the independent variable when it does not lead to the desired effect, while also not compromising the integrity of the experimental design.

**Research using single-case, multiple baseline designs.** The use of single-case research design has been well established in the ABA literature (Cooper, Heron & Heward, 2007; Bailey & Burch, 2018), and there is some support for the use of single-case research in training and supervision research. White (1993), argued that single-case designs, particularly multiple-baseline designs hold potential as evaluation designs for the provision of responsible supervision. Additionally, Meany-Walen, Davis-Gage, and Lindo (2016) used a wellness-focused supervision model within a single-case design to assess the model's degree of treatment effect at improving mental health practicum students' wellness over the course of the semester.

Examples from the behavioral literature include the following two studies. Iwata, Wong, Riordan, Dorsey, and Lau (1982) utilized a multiple baseline design across behaviors (Baer, Wolf, & Risley (1968) as cited in Iwata et al. 1982) to evaluate the effects of training (i.e., written materials, classroom instruction and practice, and quizzes) on therapists’ performances during clinical interviews. Results showed improvements in therapists' interviewing skills. Results were maintained four months after replication and expansion of the research. Additionally, Gallant, Thyer, and Bailey (1991) conducted two experiments as part of the same study, in which they employed a multiple baseline across trainees (first experiment), and a multiple baseline across clients design during the second experiment to show that the Bug in the
COMPETENCY-BASED SUPERVISION IN ABA

Ear (BITE) feedback during supervised therapy sessions was effective in producing precise and immediate improvements in therapists’ clinical skills. During the BITE supervision sessions, therapists wore a wireless earphone behind their ear. The supervisor observed the ongoing therapy sessions and provided feedback and reinforcement to the therapists.

Single-case research designs can demonstrate clear causal relations between intervention and behavior change with much more efficiency than large-sample designs (McKay, 2008). One of the main reasons for its continued support relates to the provision of a technologically sound approach for studying the effects of independent variables on the behavior of individuals without interference from other subjects. In addition, employing single-case design makes it easier to discern whether clinically significant improvements have occurred for individual participants in the study, while providing information about the extent to which the intervention was effective, and whether the level of effectiveness was clinically acceptable. Furthermore, it reveals the extent to which the researchers have adequately demonstrated experimental control (Copper, Heron & Heward, 2007; Fisher, Piazza, & Roane, 2011). Also, single-case designs offer more flexibility in the implementation and evaluation of treatment interventions than group designs. While group designs do not allow for a tailoring of the intervention to the individual, single-case research provides the opportunity to modify interventions as needed, which affords greater research and clinical options leading to more innovative treatment development (McKay, 2008). Most significantly, it allows the experimenters to make treatment decisions relatively quickly and individually (Bailey & Burch, 2018). Likewise, the use of single-case design allows for a quicker conceptualization of the research study, recruitment of subjects, and practical use of resources (Copper, Heron & Heward, 2007; Fisher, Piazza, & Roane, 2011). “Any adequately
trained researcher or clinician with a modest amount of resources and effort can use this design” (McKay, 2008, p. 347).

Ultimately, the researcher intended to evaluate patterns of behavioral changes as the subjects moved temporally through different experimental conditions (McKay, 2008). The use of repeated assessment and a well-specified intervention in this study provided useful information about the intervention as well as the timing and process of change over the course of treatment (McKay, 2008). Therefore, the effectiveness of the intervention was tested by evaluating the presence and degree of change in the dependent variable (BCBA supervisors’ performance) that occurred with the introduction of the modified BST phase. Provided that all extraneous variables remained constant and the only variable that changed was the introduction of the modified BST protocol, it could be concluded that any concurrent changed in supervisor performance was caused by this change (McKay, 2008).

Data collected from the present single-case, multiple baseline research design study provided a much greater level of detail than data commonly collected in group studies (McKay, 2008). One of the two main approaches to data analysis of single-case experimental designs was used: visual inspection of graphed data. Four criteria guided visual inspection of the data: a) change in mean, b) change in level, c) change in the rate of the behavior, and d) change in trend and latency to change (McKay, 2008). “A change in mean refers to the magnitude of change that occurs in the average behavioral frequency in one condition compared to another. A change in level refers to the immediate change in behavioral performance that occurs at the transition point between one condition and the next. A change in trend refers to a systematic variation in the slope of the data points from one condition to the next. Latency to change refers to the amount of time that passes between a change in condition and a change in behavioral performance”
(McKay, 2008, p. 346). Intervention effects were clearest when there was a stark difference in trends and no latency to change between conditions (McKay, 2008).

**Strengths and Weaknesses of the Multiple-baseline Design.** Multiple-baseline designs entail the successive application of some treatment condition across two or more already established baselines (Bailey & Burch, 2018; Kazdin, 2011). Selection of this design was based on the non-reversibility of the target behaviors (supervisor skills are non-reversible behaviors), the nature of the independent variable (i.e., the modified BST protocol), as well as the versatility of the multiple-baseline design itself. Lerman et al. (2015) illustrated the use of a multiple-baseline design to evaluate non-reversible behaviors. Through the implementation of this design, the authors successfully showed that BST was effective for teaching adult participants with Autism Spectrum Disorder the necessary behavioral procedures to teach children with Autism. Additionally, Hundert and Hopkins (1992) utilized a multiple baseline design to examine the effects of training supervisors of integrated preschools on a collaborative team approach. Like the study above, by utilizing a multiple baseline design the authors effectually demonstrated that all classroom teachers increased their behaviors toward children with disabilities in both training and generalization settings. Also, Gallant, Thyer, and Bailey (1991) employed two multiple baseline designs in two separate experiments as part of the same study, to examine the effectiveness of the Ear (BITE) feedback during supervised therapy sessions. Once again, the authors successfully demonstrated that through the implementation of the multiple baseline, feedback via the BITE provided produced precise and immediate improvements in therapists’ clinical skills. Likewise, Iwata, Wong, Riordan, Dorsey, and Lau (1982) utilized a multiple baseline design (Baer, Wolf, & Risley (1968) as cited in Iwata et al. 1982) to evaluate the effects of training on therapists’ performances during clinical interviews. Once more the study
COMPETENCY-BASED SUPERVISION IN ABA

successfully demonstrated through the use of a multiple baseline, that therapists’ skills improved after the training protocol was implemented (i.e., written materials, classroom instruction and practice, and quizzes).

When applied across participants (such as supervisors), the multiple-baseline employs each participant as his or her own control, ultimately demonstrating experimental control by showing that the treatment was equally effective across different individuals (Bailey & Burch, 2018). The multiple-baseline design’s power comes from demonstrating that changes occur when, and only when, the intervention is directed at the individuals in question (McKay, 2008). This was accomplished by replicating the intervention-behavior change relation in temporal sequence across different subjects (e.g., ABA supervisors). This temporal sequencing element is vital to rule out the likelihood that extraneous variables could account for the observed change in supervisor and supervisee behavior. Multiple-baselines are also flexible and demonstrate causality by replicating treatment effects across (in this case) different individuals (e.g., different supervisors and supervisees). Furthermore, they do not require the withdrawal of effective interventions, thereby avoiding some of the ethical concerns often raised against the use of reversal designs. One potential limitation is that a convincing demonstration of causality required that behavioral changes not occur across individual supervisors or supervisees until the intervention was applied (McKay, 2008). This design was deemed most appropriate, while affording the most experimental control for assessing the implementation of the independent variable (Bailey & Burch, 2002).

Therefore, taking into account previous uses of the multiple-baseline for evaluating training protocols, statistical assessments, and considering that supervision behaviors are learned responses that fit nicely into the category of non-reversible, the selection and use of a multiple-
baseline design for this study is supported. An illustration of this design is indicated below under Figure 3.1.

*Figure 3.1.* Sample multiple baseline design across three BCBA supervisors

Confounds. The researcher attempted to control for two types of confounds in the study. The first type concerned within-subject confounds, and the second related to confounds within the research design. Participants’ fatigue and practice are two of the most common carryover effects in single-case designs, as participation in one condition may affect performance in subsequent conditions (Hall, 1998). The proposed study aimed to reduce these effects by scheduling weekly observations sessions during the baseline and treatment conditions, thereby
reducing carryover effects. Relatedly, with the use of the multiple-baseline design (discussed below), participants’ individual experiences and self-monitoring were ruled out by replicating treatment effects across multiple supervisors and supervisees (Dallery, Cassidy, & Raiff, 2013). As replications were observed across participants, performances changed when, and only when, the modified BST protocol was introduced, allocating a high degree of confidence that performance changes were caused by the introduction of the independent variable (Dallery, Cassidy & Raiff, 2013). Therefore, experimental control was demonstrated by repeated changes in the dependent variables with each successive introduction of the modified BST protocol. Consequently, it was unlikely that a confound repeatedly coincided with the introduction of the independent variable (Dallery, Cassidy & Raiff, 2013), as the effects of confounding variables would have to have been perceived in performance changes in all the participants simultaneously (Dallery, Cassidy & Raiff, 2013). Other confounds may have included supervision intensity, supervisor qualifications, and caseload number (i.e., number of supervisees and children to which supervision will be provided; Dixon et al., 2016). These confounds were addressed by maintaining the aforementioned variables constant. To this end, supervision was only conducted once a week for all supervisee participants, supervisor’s qualifications remained the same across the board (i.e., all supervisors were BCBAs with well over one year of experience post certification in the implementation of DTT), and supervisory caseload remained the same (i.e., no more than two supervisees who were working with no more than two children).

**Research questions analysis.** To assess research questions number one (i.e., Were standard BCBA supervisory practices effective in improving supervisee behavior during DTT?) and number two (i.e., After the modified BST protocol was introduced, did BCBA supervisors demonstrate intermediate levels of competency?) data obtained from the *Evaluation of*
COMPETENCY-BASED SUPERVISION IN ABA

Supervisor Performance checklist (see Appendix C for a copy of this document) was analyzed by graphing the behavior of BCBA supervisor participants in a multiple-baseline fashion (i.e., stacked), so that supervision sessions (time component) were displayed along the X axis, and percentage of correct supervisory skills exhibited were displayed on the Y axis. Adding the number of skills displayed and dividing the result into the number of total skills required during each session, yielded the percentage of occurrence of correct behaviors. Each graph was visually and rigorously inspected for changes in trend, bandwidth (distance from the highest baseline point on the graph, to the highest point on the treatment condition), mean, level, percent correct, latency to change, variability, and stability of the data to determine intervention effects (Bailey and Burch, 2018). An illustration of this graph was indicated above under Figure 1.

Should a functional relation be demonstrated, data results would show immediate increasing changes in trend, mean, level, and percent correct, as well as decreasing latency to change values when the modified BST protocol is implemented, in relation to baseline performances. Standard (basic-level of competency) supervisory practices included the behaviors that BCBAs engaged in during the baseline (pre-treatment) phase of the study. Based on the results of the pilot study, these practices were comprised of: prior knowledge of the implementation of DTT, as well as the following skills and abilities: a) observed trainer, b) identified and documented relevant trainer errors, c) provided corrective feedback, d) coached and/or modeled appropriate trainer responses, and e) provided positive reinforcement.

Intermediate level of supervisory performance refers to expected BCBA participants’ performances during the modified BST (treatment) phase. These included the knowledge, skills and abilities described above under standard supervisory practices, plus the following newly acquired skills and abilities: f) prioritized corrective feedback, g) prompted note taking, h)
provided opportunities for role play/practice, i) prompted trainer re: relevant skills, and j) scheduled follow-up appointment. The Evaluation of Supervisor Performance checklist (see Appendix C for a copy of this document) was used to evaluate BCBAs during the baseline and treatment phases. Achievement of intermediate-level of competency was demonstrated by an increase of 40 percent or more on the above-mentioned checklist when compared to baseline values on the same instrument.

For the assessment of research question number three (i.e., Did the introduction of the modified BST protocol affect the performance of supervisees receiving supervision from BCBAs during the implementation of DTT?), collateral supervisee data was analyzed, in a similar fashion as supervisor data. Supervisees’ performances were also graphed in a multiple-baseline fashion (i.e., stacked), so that DTT sessions (time component) were displayed along the X axis, and percentage of correct supervisee skills exhibited were displayed on the Y axis. Adding the number of skills displayed and dividing the result into the number of total skills expected during each session, yielded percentage of correct behaviors observed. Each graph once again was visually and rigorously inspected for changes in trend, bandwidth, mean, level, percent correct, latency to change, variability, and stability of the data to determine intervention effects (Bailey and Burch, 2018).

Should the data establish a functional relation, data results would show immediate increasing changes in trend, mean, level, and percent correct, as well as decreasing latency to change values when the modified BST protocol is implemented, in relation to baseline performances. To evaluate the behaviors of supervisees before and after the implementation of the modified BST protocol, the researcher and used the Evaluation of Supervisee Performance checklist (see Appendix D for a copy of this document). The document includes the following set
COMPETENCY-BASED SUPERVISION IN ABA

of skills: a) conducting a preference assessment, b) using different ways to call the client’s attention, c) phrasing the discriminative stimulus (S^D) as an instruction, d) differentially reinforcing correct responding, e) contriving manding opportunities, f) interspersing skills, g) shaping, h) accurately collecting data, i) providing vocal praise, and j) incorporating previous feedback provided by the supervising BCBA (please see Definitions and Terms for a description of each skill).

After implementation of the independent variable (i.e., the modified BST protocol), a positive effect on supervisees’ performances was demonstrated by an increase of 50 percent or more on the Evaluation of Supervisee Performance (see Appendix D for a copy of this document) form when compared to baseline values on the same instrument.

The purpose of this research enterprise was to accurately arrive at truthful conclusions regarding relations among variables (McKay, 2008). The current training model evaluated whether BCBA s engaged in some (and ultimately all) aspects of the modified BST protocol during baseline, and after the training that was provided. Therefore, each BCBA’s performance was compared with his/her own performance before and after training. After the modified BST training was implemented, supervisors maintained standard supervisory practice skills (basic-level of competency) and acquired the second set of new supervisory skills into their repertoire. This signified that supervisors had achieved an intermediate-level of competency (please see Definitions and Terms for a description of basic and intermediate levels of supervisory competency). Additionally, the current study evaluated whether supervisees’ performances improved after their supervisors had completed the modified BST training. Like BCBA participants, each supervisee’s performance was compared with his/her own performance before and after the supervisor has undergone training. After the implementation of the independent
variable, supervisees’ performances improved. This signified the acquisition and maintenance of DTT skills on the part of each supervisee.

**Social Validity Analysis.** The *Supervision Social Validity Questionnaire* was analyzed descriptively. Data from the *Supervision Social Validity Questionnaire* (see Appendix H for a copy of this document) developed for this study yielded whether (a) supervisors thought that the intervention was helpful in improving their supervisory practice, (b) they would continue to use the protocol in their day-to-day supervision, (c) they would recommend the intervention to other BCBA supervisors, (d) all the skills were important to achieving and maintaining excellent therapist performance, and (e) they would like to list any skills that did not seem important for a supervisor to demonstrate.
Introduction

The study took place across five settings, at three different local ABA companies/facilities, over the course of eight months (i.e., July 2018 through March 2019). A total of 228 videos were filmed and coded throughout, which included five BCBA supervisors, six supervisees, and six children. This represented a total of six pairs of supervisors-supervisees, and six pairs of supervisees-children. One supervisor’s performance was also measured for generalization with a different supervisee and child, which constituted the sixth pair of supervisor-supervisee and supervisee-child participants. A total of four research assistants participated in the study, though one of them withdrew from the study early-on during the baseline condition. For the vast majority of the study, one research assistant collected the videos, and two other assistants coded the videos: the primary observer, and the reliability observer. Coded data were then forwarded to the researcher for graphing and analysis, which aided in determining the course of the study. Data were collected once per week.

Demographic Characteristics

As previously mentioned, five BCBA supervisors participated in the study. Two of the BCBA supervisors were doctoral-level BCBA-Ds who worked at a private ABA school for children, teenagers, and adults diagnosed with ASD and related intellectual disabilities, in leadership roles. Both BCBAs had been supervising teachers/therapists for over three and four years each, respectively, and had worked at their school for several years. Another pair of supervisors worked for a private ABA company that provided clinic-based drop-off ABA therapy services to toddlers and preschool-aged children diagnosed with ASD and other intellectual disabilities. Both supervisors had over two and four years of experience supervising
COMPETENCY-BASED SUPERVISION IN ABA

ABA therapists each, respectively, and had worked for the ABA company for a few years. The fifth supervisor was the owner of a local ABA company that also provided clinic-based drop-off ABA therapy services to toddlers and children (including preschool, school-aged, and teenagers) diagnosed with ASD and other intellectual disabilities. This last supervisor had over 4 years of experience supervising ABA therapists as well as bachelor’s level behavior analysts (BCaBAs). All BCBA participants were females in their late twenties to late thirties.

Each BCBA supervisor also worked with one or more supervisees. For the most part, one supervisor worked with only one supervisee, except one of the supervisors who also worked with a second supervisee for the purposes of skill generalization. The first two supervisees were teachers at the ABA school for children, teenagers and adults diagnosed with ASD and other intellectual disabilities referenced above. One of the supervisees/teachers worked in the preschooler classroom, and the other worked in the teenagers’ classroom. Both supervisees had been working at the school for at least six months and had expressed interest in furthering their career in the field of ABA. The second pair of supervisees along with the supervisee who participated in the protocol generalization worked for the previously mentioned private ABA company that provided clinic-based drop-off ABA therapy services to toddlers and preschool-aged children diagnosed with ASD and other intellectual disabilities. All three supervisees worked with preschool-aged children on a daily basis and had worked for the company for at least six months. Two of the supervisees had expressed interest in furthering her career in the field of ABA, while the other therapist was relatively new to the field and was unsure about her future within. The sixth supervisee worked for the local ABA company that also provided clinic-based drop-off ABA therapy services to toddlers and children (including preschool, school-aged, and teenagers) diagnosed with ASD and other intellectual disabilities. This supervisee had also...
COMPETENCY-BASED SUPERVISION IN ABA

worked with preschool-aged children for at least six months and was at the time enrolled in an ABA academic program. In addition, this supervisee was studying to take her BCaBA examination. All supervisees engaged in DTT on a daily basis, expressed interest in participating in the study, as well as learning new ways in which supervision could help improve their DTT skills. One male and five female supervisees participated, most of whom were in their late twenties, except for one who was in her fifties.

Each supervisee also worked with one child during the study. Children participants were preschool and elementary school-aged (i.e., between four and nine years old). All the children were male and attended the aforementioned programs every day for approximately 3-8 hours. Two of the children who participated in the study were fully vocal, one used an augmentative communication device to communicate his wants and needs, two others used an augmentative communication device as well as vocal approximations to communicate their wants and needs, and one learned to communicate using spoken language during the study. This last child would communicate by pointing and leading at the beginning of the study, and by the end (8 months later) was communicating using two-word sentences independently to both make requests and identify relevant items.

Data Analysis

Results will first be explained by providing a descriptive analysis of the data obtained (i.e., mean, trend, level, latency to change, bandwidth, variability, and stability) followed by a discussion of supervisory levels of competency, under the section of the same name. Data will be shown for all five supervisors first, after which the data for all supervisees will be presented, followed by one additional figure depicting all graphs included in the multiple baseline design.


Research question number one. For each BCBA participant, analysis of the baseline conditions described below will attempt to answer the first research question: Are standard (basic-level of competency) BCBA supervisory practices effective in improving supervisee behavior during DTT?

Figure 4.1 below shows all the supervisor participants’ baseline data depicted as two multiple-baselines. The percentage of correct responding for supervisors is shown on the Y axis. Number of sessions were plotted on the X axis. Data points denoted in black circles were obtained from the primary observer and represent percentage of skills performed. These data were used by the researcher to make decisions during the study. Data denoted in pins (●) were obtained from the reliability observer and represent inter-observer agreement (IOA). All sessions were scored for Inter-observer Agreement (IOA). The vertical dashed line represents a change in condition from baseline or pre-intervention to the treatment condition.

During the baseline condition, data indicated that all five ABA supervisor participants demonstrated basic-level of supervisory performance (M = SUP-01: 40%, M = SUP-02: 21%, M = SUP-03: 40%, M = SUP-04: 25%, M = SUP-05: 33%). In addition, IOA pertaining to supervisor performance remained over 80 percent during this condition (M = SUP-01: 79%, M = SUP-02: 83%, M = SUP-03: 91%, M = SUP-04: 89%, M = SUP-05: 84%), with the exception of SUP-01’s IOA. As previously mentioned, when two or more individuals reach at least 80 percent agreement or better on the sessions, the data should be considered reliable (Bailey & Burch, 2018; Cooper, Heron & Heward, 2007; Teddlie & Tashakkori, 2009). Therefore, baseline data for four out of the five BCBA supervisor participants were considered reliable. Regardless, mean IOA scores for SUP-01 were very close to the current requirement (i.e., 79%). Additionally, the mean IOA for all SUP participants during this condition was 85 percent.
Regarding data stability, for the most part, scores indicated somewhat stable values during the baseline condition of the study. Perhaps the most variable data belonged to SUP-04 whose scores fluctuated between 10 and 43 percent. Though presenting a similar data range, SUP-02’s performance was considered a stable (predictable) decline. Nevertheless, values never improved beyond this range for any of the supervisors. These scores may suggest that supervisors only demonstrated skills within their supervisory repertoire, leaving room for the introduction of new DTT supervision skills.

As a group, BCBA supervisor participants’ performances during the baseline condition showed some variability. Three out of the five supervisors displayed relatively stable performances, especially toward the last weeks of this condition, while the remaining two supervisors showed some performance variability. However, this variability remained within a specific range for all BCBAs, especially the latter two. The results guided the researcher to conclude that these particular set of participants were ready for the introduction of new skills related to the supervision of DTT.

An individual-level analysis of each BCBA supervisor participant’s performance during the baseline condition is provided below.

Scores displayed on the first graph of the left “leg” below for SUP-01 show that during the three baseline sessions conducted, SUP-01 engaged in four of the 10 skills included on the modified BST protocol. The types of skills demonstrated on each baseline session varied, except for observing the supervisee and providing corrective feedback. The mean score for baseline was 40 percent. Data were stable, at relatively low-to-medium levels, with zero variability, trend and bandwidth. Inter-observer agreement scores for this condition ranged between 78 and 80 percent, with a mean value of 79 percent.
The second graph of the left “leg” displayed below for SUP-02 showed values of 40 percent during the first two baseline sessions. Then, scores remained low fluctuating between 10 and 20 percent. A steady downward trend indicated that this supervisor’s performance consistently declined throughout this condition. Though the types of skills demonstrated on each baseline session generally varied, the only skill that this supervisor consistently demonstrated throughout this condition was coaching and/or modeling appropriate supervisee responses. The mean score for baseline was 21 percent. For this condition, data showed (as previously mentioned) a steady downward trend, from medium to low data levels, with relatively low variability and medium bandwidth (i.e., 30 data points). Inter-observer agreement data for this condition ranged between scores of 70 and 100 percent. The mean value was about 83 percent.

Figure 4.1. Percent correct and IOA scores for SUP-01 through SUP-05 during baseline
The third graph of the left “leg” in the multiple-baseline above depicts the percentage of correct responding for SUP-03 on the Y axis. The first baseline session for this participant took place with the same child as SUP-01/SEE-01, which is why it was not included. The researcher recruited a second child shortly after, who remained the official participant for this pair (SUP-03/SEE-03), throughout the study. Scores displayed below for SUP-03 indicated values of 40 percent during the first baseline session with the official new child participant. After, values remained at medium levels, first declining until finally stabilizing back at 40 percent, indicating that this supervisor’s performance eventually stabilized after showing a small downward trend. Though the types of skills demonstrated on each baseline session generally varied, this supervisor consistently demonstrated the following: observing the supervisee, providing feedback to the supervisee, coaching and/or modeling appropriate responses, and providing positive reinforcement. The mean score for baseline was 40 percent. For this condition, data showed medium levels, with relatively low variability and medium bandwidth (20 points). Inter-observer agreement scores ranged from 90 to 100 percent, with a mean value of 91 percent.

The second (right side) leg of the multiple baseline depicts data pertaining to the fourth supervisor participant. The top of Figure 4 above represents the percentage of correct responding for SUP-04 on the Y axis. Scores displayed below for SUP-04 indicated values of 33 percent during the first baseline session, followed by a small increase, and afterward a somewhat steady decline in performance, which culminated in a slight improvement. Nevertheless, baseline values remained between 10 and 43 percent throughout, indicating that this supervisor’s performance though showing some variability, mostly declined. The types of skills demonstrated on each baseline session varied very little; they mostly consisted of observing the supervisee, providing feedback to the supervisee, and coaching and/or modeling appropriate responses. The mean score
for baseline was 25 percent. For this condition, data remained at medium-to-low levels, with some variability and medium bandwidth (i.e., 33 data points). Inter-observer agreement during this condition ranged between 60 and 100 percent, with a mean value of roughly 89 percent.

The last supervisor participant graph on the second “leg” pertains to SUP-05. Similar to the previous participants, the percentage of correct responding for SUP-5 was represented on the Y axis. The first two baseline sessions collected for this supervisor included a child participant who was withdrawn from the study after the second session, due to meeting exclusion criteria. As a result, a second child was recruited, and baseline data collection was re-initiated. Scores displayed below depict data for SUP-05 with the second child participant and indicated values of 43 percent during the first baseline session, followed by a small fluctuation, and then very steady performance throughout the last five sessions. Therefore, for the most part, this supervisor’s performance though showing little variability at the beginning of the study, remained stable throughout baseline. The types of skills demonstrated on each baseline session did not vary except for one session; they consisted of observing the supervisee, providing feedback to the supervisee, and coaching and/or modeling appropriate responses. SUP-05 also scheduled a follow-up appointment on session 3. This was the only other skill that the supervisor demonstrated during baseline. The mean score for baseline was 33 percent. For this condition, data remained at medium levels, with very little variability and a small bandwidth (13 points). Inter-observer agreement values ranged from 60 to 100 percent, with a mean score of approximately 84 percent.

**Research question number two.** Similarly, for each supervisor participant, analysis of the treatment/modified BST conditions described below will attempt to answer the second
research question: After the modified BST protocol is introduced, will BCBA supervisors demonstrate intermediate levels of supervisory behavior during supervision sessions?

Figure 4.2 below shows all the supervisor participants’ baseline and treatment data depicted as two multiple-baselines. Once again, the percentage of correct responding is shown on the Y axis. Number of sessions were plotted on the X axis. Data points denoted in black circles were obtained from the primary observer and represent percentage of skills performed. Data denoted in pins (♀) represent inter-observer agreement (IOA). All sessions were scored for Inter-observer Agreement (IOA). The vertical dashed line represents a change in condition from baseline or pre-intervention to the treatment condition.

During the modified BST condition, data indicated that all five ABA supervisor participants demonstrated intermediate-level of supervisory performance (M = SUP-01: 97%, M = SUP-02: 98%, M = SUP-03: 91%, M = SUP-04: 96%, M = SUP-05: 95%). In addition, IOA pertaining to supervisor performance remained above 81 percent during this condition (M = SUP-01: 90%, M = SUP-02: 94%, M = SUP-03: 89%, M = SUP-04: 82%, M = SUP-05: 89%). Just as before, when two or more individuals reach at least 80 percent agreement or better on the sessions, the data should be considered reliable (Bailey & Burch, 2018; Cooper, Heron & Heward, 2007; Teddlie & Tashakkori, 2009). Therefore, all treatment data for the five BCBA supervisor participants were considered reliable with a mean value of 89 percent.

Regarding data stability, for the most part, scores indicated stable values during the treatment condition of the study. Perhaps the most variable data belonged to SUP-03 whose scores fluctuated between 78 and 100 percent. The other four supervisors showed mostly stable performance with some slight variability. Nevertheless, values never dropped beyond 78 percent
for any of the supervisors. These scores may suggest that for the great majority of the treatment sessions, supervisors’ performances remained stable at maximum criteria.

As a group, BCBA supervisor participants’ performances during the modified BST condition showed very little variability. Four out of the five supervisors displayed stable performances, while one supervisor showed some performance variability. However, this variability remained within a specific low range for this BCBA. The aforementioned results steered the experimenter to deduce that these particular set of participants had acquired and maintained new skills related to the supervision of DTT.

A more individualized analysis of each BCBA supervisor participant’s performance during the modified BST condition is provided below.

After the introduction of the modified BST protocol, SUP-01 immediately demonstrated eight skills out of nine possible skills included on the protocol. The BCBA continued to demonstrate all required possible skills, including throughout the last two maintenance sessions. For all BCBA participants, maintenance began the session following a score of 100 percent. In the first treatment session, SUP-01 failed to demonstrate the skill of providing positive reinforcement to the supervisee. This was the only skill that SUP-01 failed to demonstrate during the modified BST condition of this study. The mean score for this condition was 97 percent. Visual inspection of the data revealed an immediate change in data level from relatively low-to-medium, to high. There was a short latency to change (indicating rapid skill acquisition) and very little variability as the data remained stable for three sessions after a slight upward trend. The bandwidth was also relatively small (i.e., 11 data points). Performance on the last three sessions of this condition is indicative of behavior under stimulus control. Inter-observer agreement scores for this condition ranged between 80 and 100 percent, with a mean value of 90 percent.
Similarly, following the introduction of the modified BST protocol, SUP-02 immediately demonstrated eight out of nine possible skills (10 altogether) on the first session. This was the most dramatic effect seen during the study. Performance remained steady with the supervisor consistently demonstrating all possible required skills throughout the following six sessions, including five maintenance sessions, which was indicative of rapidly acquired skills, under stimulus control. On the eighth session, SUP-02’s scores decreased to 89 percent, at which time a booster session (session nine) was conducted by the researcher/Uber Supervisor, resulting once again in the supervisor engaging in all of the required possible skills throughout sessions nine through 13. Sessions 10 through 13 were maintenance. There were only two skills that SUP-02 failed to demonstrate during the modified BST condition of this study: providing positive reinforcement and prompting the supervisee regarding relevant skills, during sessions one and eight respectively. The mean score for this condition was 98 percent. Visual inspection of the data revealed an immediate change in data level from low, to high. There was a very short latency to change indicating rapid skill acquisition, and little variability. Supervisor performance remained stable at high levels for six sessions after a slight upward trend, including the first five and last four maintenance sessions, except for session eight. The bandwidth was also relatively narrow (i.e., 11 data points). Performance on this condition is indicative of behavior under stimulus control. Inter-observer agreement for this condition revealed scores that ranged from 80 to 100 percent, with a mean value of around 94 percent.

Likewise, after introducing the modified BST protocol, SUP-03 immediately demonstrated seven out of nine possible skills (10 altogether) on the first session. Supervisor performance steadily improved finally achieving maximum criteria for two sessions, at which point maintenance began. Though remaining at high levels, SUP-03’s performance slightly
COMPETENCY-BASED SUPERVISION IN ABA

declined during the fifth session, prompting a booster session conducted by the researcher on the sixth session, resulting again in an increase in performance during the last session. Shortly afterward, the researcher was informed that one of the participants was no longer able to participate in the study, due to health-related concerns, which were unrelated to the implementation of the modified BST protocol to SUP-03. Except for sessions four and five, SUP-03 failed to demonstrate the following skills during the modified BST condition of this study: observing the supervisee, prompting the supervisee regarding relevant skills, scheduling a follow-up appointment, and prompting note-taking. The mean score for this condition was 91 percent. Visual inspection of the data revealed an immediate change in data levels from medium, to high. There was a very short latency to change indicating rapid skill acquisition, and some variability. For the most part, supervisor performance showed a steady upward trend throughout. The bandwidth was also relatively small (i.e., 22 data points). Performance on this condition was indicative of behavior under skill acquisition. Inter-observer agreement scores showed a range in values between 40 and 100 percent, with a mean score of roughly 89 percent.
Like the first three BCBA supervisors, subsequent to the introduction of the modified BST protocol, SUP-04 immediately demonstrated nine out of ten possible skills on the first session. Supervisor’s performance achieved maximum criteria on the second treatment session, at which point maintenance began. Scores remained steady with the supervisor consistently demonstrating all possible required skills throughout, except for sessions six and nine. This was indicative of rapidly acquired skills, under stimulus control. On the sixth session, SUP-04’s
scores decreased to 89 percent, at which time a booster session (session seven) was conducted by
the researcher/Uber Supervisor, resulting once again in the supervisor engaging in all of the
required skills during the following session. Sessions three through six, as well as eight and nine
were maintenance. Though the last data point indicates a slight decline in performance, scores
nevertheless remained high throughout treatment. There were only two skills that the supervisor
failed to demonstrate during the modified BST condition of this study: prompting the supervisee
regarding relevant skills and scheduling the follow-up appointment. The mean score for this
condition was 96 percent. Visual inspection of the data revealed an immediate change in data
level from relatively low, to high. There was a very short latency to change indicating rapid skill
acquisition, and little variability. The bandwidth was also relatively small (i.e., 11 data points).
Performance on this condition is indicative of behavior under stimulus control. Inter-observer
agreement scores ranged between 50 and 100 percent, with a mean value of approximately 82
percent.

Not unlike the rest of the BCBA participants, following the introduction of the modified
BST protocol, SUP-05 immediately demonstrated eight out of ten possible skills on the first
session. Supervisor’s performance achieved maximum criteria on the second treatment session,
at which point maintenance began. Scores remained steady for the next two sessions with the
supervisor consistently demonstrating all possible required skills. This was indicative of rapidly
acquired skills, under stimulus control. On the fifth session, SUP-05’s scores decreased to 90
percent. Just prior to session six, the researcher reminded the BCBA participant of the one skill
undemonstrated on the previous session, resulting once again in the supervisor engaging in all of
the required skills during the following two sessions. On session eight, supervisor’s performance
again slightly declined again to 89 percent, at which time a booster session was conducted,
resulting in maximum required performance on the last treatment session. Though this supervisor’s performance slightly declined twice during treatment, scores still remained high throughout. There were only two skills that the supervisor failed to demonstrate during the modified BST condition of this study: prompting the supervisee regarding relevant skills and scheduling the follow-up appointment. The mean score for this condition was 95 percent. Visual inspection of the data revealed an immediate change in level from basic, to intermediate. Figure 4.3 below provides an illustration of the skills included in the aforementioned basic-level of supervisory performance, as well as the new skills comprised within the modified BST protocol. Skills next to a blue BST box depict behaviors classified as basic-level of supervisory performance. Skills next to the word “New” in red depict intermediate-level of supervisory performance. In addition, there was a very short latency to change indicating rapid skill acquisition, and some variability. The bandwidth was fairly medium (i.e., 20 data points). Performance on this condition is mostly indicative of behavior under stimulus control. Inter-observer agreement scores ranged from 70 to 100 percent, with a mean value of roughly 89 percent.
Research question number three. Finally, for each supervisee participant, analysis of both baseline and modified BST conditions will attempt to answer the third research question: Does the introduction of the modified BST protocol affect the performance of supervisees receiving supervision from BCBAs during the implementation of DTT?

Figure 4.4 below shows collateral data collected on the supervisee participants’ performances during baseline depicted as two multiple-baselines. The percentage of correct responding for supervisees is shown on the Y axis. Number of sessions were plotted on the X axis. Similar to the data described above for supervisors, data points denoted in black circles were obtained from the primary observer and represent percentage of skills performed. As before, these data were used by the researcher to make decisions during the study. Data denoted in pins (🪖) were obtained from the reliability observer and represent inter-observer agreement.
COMPETENCY-BASED SUPERVISION IN ABA

(IOA). All sessions were scored for Inter-observer Agreement (IOA). The vertical dashed line represents a change in condition from baseline or pre-intervention to the treatment condition.

During the baseline condition, data indicated that all five supervisee participants demonstrated basic-levels of performance (M = SEE-01: 59%, M = SEE-02: 34%, M = SEE-03: 51%, M = SEE-04: 52%, M = SEE-05: 30%). In addition, IOA pertaining to supervisees’ performances was somewhat on the lower range (M = SEE-01: 77%, M = SEE-02: 69%, M = SEE-03: 83%, M = SEE-04: 66%, M = SEE-05: 64%), for only one of the supervisee’s data reached a score of over 80 percent. As previously mentioned, when two or more individuals reach at least 80 percent agreement or better on the sessions, the data should be considered reliable (Bailey & Burch, 2018; Cooper, Heron & Heward, 2007; Teddlie & Tashakkori, 2009). Therefore, baseline data for four out of the five supervisee participants were not considered reliable. As it pertained to these participants, the observers for the most part scored the occurrence and non-occurrence of similar skills (about 60-70 percent of the time), while also scoring a significant number of different skills as either performed or not-performed (30-40 percent of the time). Regardless, mean IOA scores for these supervisees were close to the current requirement at 72 percent.

Regarding data stability, for the most part, scores indicated somewhat stable values during the baseline condition of the study for two supervisees: SEE-01 and SEE-03. SEE-02’s data showed some variability culminating with a steady decline in performance values. Perhaps the most variable data belonged to SEE-04 whose scores fluctuated the most. Though also a little variable, SEE-05’s data ultimately showed a predictable performance deterioration pattern. Nevertheless, all supervisees’ scores indicated room for improvement.
As a group, supervisee participants’ performances during the baseline condition revealed more variability when compared to supervisors’ performances. Most of the supervisees’ data (three out of five) showed some level of variability, though predictable data patterns emerged toward the end of the baseline condition. Once stabilized (i.e., behavioral patterns were predictable), supervisees’ performances appeared to also indicate that scores did not generally surpass a certain percentage. The results guided the researcher to conclude that these particular set of participants might benefit from the introduction of a new supervision protocol specific to DTT.

An individualized analysis of each supervisee participant’s performance during the baseline condition is provided next.

Data presented for SEE-01 revealed a score of 67 percent for the first baseline session, and subsequent scores of 55 percent for the second and third baseline sessions, which indicated that supervisee performance was slightly declining. SEE-01’s performance fluctuated between five and six possible skills included on the Evaluation of Supervisee Performance. Though the types of skills demonstrated on each baseline session varied, the supervisee consistently engaged in the following skills during baseline: phrasing S^D as instruction, differentially reinforcing correct responding, accurately collecting data, and providing different types of praise. The mean score for baseline was 59 percent. The data were mostly stable, at relatively medium levels, with a small variability, a slight downward trend, and a small bandwidth (12 data points). Inter-observer agreement scores for this condition ranged from 70 to 80 percent, with a mean value of approximately 77 percent.

Data from SEE-02 showed scores that fluctuated between of 55 and zero percent, signifying that this supervisee’s performance though showing an increase over two sessions,
COMPETENCY-BASED SUPERVISION IN ABA

consistently declined over the course of the baseline condition. The mean score for baseline was 34 percent. During this condition, data values remained at medium-to-low levels, with fairly low variability. The bandwidth for this condition was wide (i.e., 55 data points). Inter-observer agreement scores for ranged from 50 to 80 percent, with a mean value of roughly 69 percent.

Data shown for SEE-03 on the baseline condition showed scores that fluctuated between 40 and 55 percent, indicating that this supervisee’s performance slightly improved, until it reached a stable level of performance. The mean score for baseline was 51 percent. During this condition, data values remained at medium levels, with very low variability and short bandwidth. The data also showed a small upward trend consistent with slight performance improvement. Inter-observer agreement values for this condition ranged from 70 to 100 percent with a mean of about 83 percent.
The data for SEE-04 showed high levels of variability for the first five sessions, indicating unpredictable performance. However, scores stabilized at medium levels between 55 and 67 percent during the last four baseline sessions. Overall, baseline condition scores fluctuated between 22 and 78 percent, indicating that this supervisee’s performance was again highly variable, until it reached a somewhat stable level of performance. The mean score for baseline was 52 percent. During this condition, data values fluctuated between low, medium, to
somewhat high values, which also revealed a wide bandwidth (56 points). Results from the inter-
observer agreement revealed a range in scores from 50 to 80 percent, with a mean value of about
66 percent.

Data obtained for SEE-05 showed an improvement in performance and upward trend
during the first three baseline sessions. Performance then slowly declined over the next three
sessions, after which it mostly fluctuated between 22 and 40 percent. SEE-05’s performance
never went over 44 percent during this condition. It appeared that this supervisee’s data revealed
a pattern whereby performance fluctuated between improvements up to a certain level, followed
by a decline, improvement, and then a decline in performance again. The mean score for baseline
was 30 percent. During this condition, data values fluctuated between low and medium values,
which also revealed a somewhat medium bandwidth (32 points). Inter-observer agreement scores
showed values ranging from 60 to 89 percent, with a mean of about 64 percent.

Figure 4.5 below shows collateral data collected on the supervisee participants’
performances during the modified BST condition depicted as two multiple-baselines. The
percentage of correct responding for supervisees is shown on the Y axis. Number of sessions
were plotted on the X axis. Like the data described above, data points denoted in black circles
were obtained from the primary observer and represent percentage of skills performed. Likewise,
these data were used by the researcher to make decisions during the study. Data denoted in pins (📍)
were obtained from the reliability observer and represent inter-observer agreement (IOA). All
sessions were scored for Inter-observer Agreement (IOA). The vertical dashed line represents a
change in condition from baseline or pre-intervention to the treatment condition.

During the modified BST condition, data indicated that all five supervisee participants
demonstrated performance improvements (M = SEE-01: 77%, M = SEE-02: 81%, M = SEE-03:
COMPETENCY-BASED SUPERVISION IN ABA

97%, $M = \text{SEE-04: 93\%, } M = \text{SEE-05: 89\%}$) relative to baseline, with performance falling within the 77 to 97 percent rage. In addition, IOA pertaining to supervisees’ performances remained high at values over 83 percent ($M = \text{SEE-01: 87\%, } M = \text{SEE-02: 84\%, } M = \text{SEE-03: 91\%, } M = \text{SEE-04: 90\%, } M = \text{SEE-05: 87\%}$). As stated before, when two or more individuals reach at least 80 percent agreement or better on the sessions, the data should be considered reliable (Bailey & Burch, 2018; Cooper, Heron & Heward, 2007; Teddlie & Tashakkori, 2009). Therefore, baseline data for all five supervisee participants were considered reliable, as the mean IOA score for supervisees’ performances was 88 percent.

Regarding data stability, for the most part, scores indicated somewhat stable values during the modified BST condition of the study for three supervisees: SEE-01, SEE-03, and SEE-04. SEE-02’s data probably showed the most variability culminating with a steady increase in performance values. This supervisee had the most fluctuation in scores. Though also somewhat variable, SEE-05’s data ultimately showed a predictable pattern of performance improvement culminating in stable performance. Nevertheless, all supervisees’ scores markedly improved during the treatment phase.

As a group, supervisee participants’ performances during the modified BST condition revealed more variability when compared to supervisors’ performances under this same condition. A few of the supervisees’ data (two out of five) displayed some level of variability, though predictable data patterns emerged toward the end of the treatment condition. Once stabilized (i.e., behavioral patterns were predictable), supervisees’ performances appeared to also maintain for two or more consecutive sessions. Results indicated that these particular set of supervisee participants benefited from the introduction of a new supervision protocol that facilitated the acquisition and maintenance of new DTT skills.
A more specified analysis of each supervisee participant’s performance during the modified BST condition was provided below.

After the introduction of the modified BST protocol, SEE-01’s performance immediately increased to 70 percent during the first session, and then remained between values of 70 and 90 percent for the last three sessions, including the last two maintenance sessions. The mean score for this condition was 77 percent. Visual inspection of the data revealed an immediate change in data level from medium to medium-high. There was a relatively short latency to change and low variability as the data, though revealing an upward trend, remained within 70 and 90 percent values. The bandwidth was relatively medium (i.e., 20 data points). Performance during this condition was indicative of behavior under skill acquisition. This meant that the individuals, though performing at higher levels than during baseline, still failed to demonstrate a few skills during the first few treatment sessions, until their performances stabilized at maximum levels throughout the last one to four treatment sessions. Inter-observer agreement for the treatment condition revealed a range in scores from 80 to 90 percent, with a mean value of 87 percent.

Following the introduction of the modified BST protocol, SEE-02’s performance immediately reached a score of 50 percent during the first session, indicating rapid skill acquisition. Then, scores continued to increase, except for sessions four and seven, and remained steady at 100 percent for the last four treatment sessions. The mean score for this condition was 81 percent. Visual inspection of the data revealed an immediate change in data level from medium to high. There was a very short latency to change indicating rapid skill acquisition, and a moderate-to-high level of variability, for the data consistently trended upward, until reaching maximum criteria and stability over the last four maintenance sessions. The bandwidth for this condition’s data was large (i.e., 50 data points). Performance on this condition is indicative of
behavior under skill acquisition. Inter-observer agreement scores ranged from 50 to 100, with a mean value of 84 percent.

After the modified BST protocol was introduced, SEE-03’s performance immediately attained a score of 89 percent during the first session. Then, scores reached maximum criteria, and remained stable at 100 percent, except on session five. The mean score for this condition was 97 percent. Visual inspection of the data revealed an immediate change in level from medium to high, indicating rapid skill acquisition and a very short latency to change, as well as very low variability. The bandwidth for this condition’s data was relatively narrow (i.e., 11 data points). Performance on this condition is indicative of behavior under stimulus control. Inter-observer agreement scores ranged from 70 to 100 percent, with a mean value of 91 percent.
Figure 4.5. Percent correct and IOA scores for SEE-01 through SEE-05 during the baseline and modified BST phases

Once the modified BST protocol was introduced, SEE-04’s performance achieved a score of 70 percent during the first session. Then, scores reached maximum criteria for two sessions, and somewhat declined until attaining maximum criteria again on session six and remaining stable over the last four treatment sessions. The mean score for this condition was 93 percent. Visual inspection of the data did not reveal a change in level until the second treatment session,
from medium to high, indicating some latency to change, as well as some variability. The bandwidth for this condition’s data was relatively medium (i.e., 30 data points). Performance on this condition is indicative of behavior under both skill acquisition first, and then stimulus control. Inter-observer agreement showed values that ranged from 70 to 100 percent, with a mean score of 90 percent.

When the modified BST protocol was introduced, SEE-05’s performance immediately achieved a score of 80 percent during the first session. Subsequently, there was a slight decline in performance, followed by improvement over two sessions, and a slight decline again, similar to the pattern of responding during baseline. Then, performance consistently improved until reaching maximum criterion and remaining stable for the last three treatment sessions. Though this supervisee’s performance slightly declined twice during treatment, scores still remained high throughout between 70 and 100 percent. The mean score for this condition was 89 percent.

Visual inspection of the data revealed an immediate change in level from relatively medium to high values indicating rapid skill acquisition, as well as some variability. The bandwidth for this condition’s data was relatively medium (i.e., 30 data points). Performance on this condition is indicative of behavior under both skill acquisition first, and then stimulus control. Inter-observer agreement revealed scores that ranged from 60 to 100 percent, with a mean value of approximately 87 percent.

Figure 4.6 below depicts the performance of all supervisors and supervisees in a multiple baseline fashion, except the generalization pairs. The multiple baseline “leg” on the left depicts the behaviors of the first six participants described above (SUP-01, SUP-02, and SUP-03; SEE-01, SEE-02, and SEE03). The multiple baseline “leg” on the right shows the performance of the remaining four participants also described above (SUP-04, SUP-05, SEE-04, and SEE-05). As
displayed, following the first baseline, subsequent ones ran for a longer period until stabilization, after which time the modified BST condition was implemented. Consequently, all baselines varied from one set to the next. Similarly, all treatment conditions showed increases in performances following the implementation of the independent variable at the supervisor level, thus demonstrating experimental control.

Figure 4.6. Percent correct and IOA scores for SUP and SEE’s during the baseline and modified BST phases
The tables below illustrate mean percent correct performance and inter-observer agreement scores for all participants across both baseline and modified BST conditions. As shown, participants’ performance scores during baseline reached a mean value of 32 percent for supervisors, and 45 percent for supervisees. Following the introduction of the BST protocol, participants’ scores reached a mean value of 95 percent for supervisors and 87 percent for supervisees. Inter-observer agreement data presented below for all participants revealed that during baseline, observers reached a mean score of 85 percent agreement for supervisors’ behaviors, and a mean score of 72 percent agreement for supervisees’ performances. During the modified BST protocol condition, inter-observer agreement data revealed an overall mean score of 89 percent for supervisors, as well as a total mean score of 88 percent for supervisees’ data. In general, the primary observer’s data showed low performance values on the baseline condition (i.e., below 50 percent), and high-performance values (i.e., above 80 percent) on the treatment condition, for all participants. In addition, though data from the reliability observer showed values below 80 percent during the baseline condition, these improved during the treatment condition.

Table 4.1

Mean percent correct performance for all participants

<table>
<thead>
<tr>
<th>Participant</th>
<th>Baseline mean Percent correct</th>
<th>Modified BST mean Percent correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUP-01</td>
<td>40</td>
<td>97</td>
</tr>
<tr>
<td>SEE-01</td>
<td>59</td>
<td>77</td>
</tr>
<tr>
<td>SUP-02</td>
<td>21</td>
<td>98</td>
</tr>
<tr>
<td>SEE-02</td>
<td>34</td>
<td>81</td>
</tr>
<tr>
<td>SUP-03</td>
<td>40</td>
<td>91</td>
</tr>
<tr>
<td>SEE-03</td>
<td>51</td>
<td>97</td>
</tr>
<tr>
<td>SUP-04</td>
<td>25</td>
<td>96</td>
</tr>
<tr>
<td>SEE-04</td>
<td>52</td>
<td>93</td>
</tr>
<tr>
<td>SUP-05</td>
<td>33</td>
<td>95</td>
</tr>
<tr>
<td>SEE-05</td>
<td>30</td>
<td>89</td>
</tr>
<tr>
<td>Mean SUP</td>
<td>32</td>
<td>95</td>
</tr>
<tr>
<td>Mean SEE</td>
<td>45</td>
<td>87</td>
</tr>
</tbody>
</table>
COMPETENCY-BASED SUPERVISION IN ABA

Table 4.2
Inter-observer agreement scores for all participants

<table>
<thead>
<tr>
<th>Participant</th>
<th>Baseline mean percent IOA</th>
<th>Modified BST mean percent IOA</th>
<th>Mean percent IOA</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUP-01</td>
<td>79</td>
<td>90</td>
<td>85</td>
</tr>
<tr>
<td>SEE-01</td>
<td>77</td>
<td>87</td>
<td>82</td>
</tr>
<tr>
<td>SUP-02</td>
<td>83</td>
<td>94</td>
<td>89</td>
</tr>
<tr>
<td>SEE-02</td>
<td>69</td>
<td>84</td>
<td>77</td>
</tr>
<tr>
<td>SUP-03</td>
<td>91</td>
<td>89</td>
<td>90</td>
</tr>
<tr>
<td>SEE-03</td>
<td>83</td>
<td>91</td>
<td>87</td>
</tr>
<tr>
<td>SUP-04</td>
<td>89</td>
<td>82</td>
<td>86</td>
</tr>
<tr>
<td>SEE-04</td>
<td>66</td>
<td>90</td>
<td>78</td>
</tr>
<tr>
<td>SUP-05</td>
<td>84</td>
<td>89</td>
<td>87</td>
</tr>
<tr>
<td>SEE-05</td>
<td>64</td>
<td>87</td>
<td>76</td>
</tr>
<tr>
<td>Mean IOA SUP</td>
<td>85</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td>Mean IOA SEE</td>
<td>72</td>
<td>88</td>
<td></td>
</tr>
</tbody>
</table>

**Protocol Generalization.** Data were also collected on the performance of SUP-04 with a second supervisee and child. The top of Figure 10 below represents the percentage of correct responding for SUP-04G on the Y axis. Scores displayed below for SUP-04G indicated values of 43 percent during the first baseline session, followed by a steady decline in performance/downward trend, and subsequently a slight improvement. SUP-04G’s performance never reached 40 percent again during baseline. Therefore, for the most part, this supervisor’s performance declined during baseline, with this supervisee. The types of skills demonstrated on each baseline session did not vary except for one session; they consisted of observing the supervisee, providing feedback to the supervisee, and coaching and/or modeling appropriate responses. SUP-04G failed to also demonstrate providing corrective feedback on session four. The mean score for baseline was 31 percent. For this condition, data remained at medium-to-low levels, with very little variability (mostly a downward trend) and a somewhat small bandwidth (23 points).
Figure 4.7. SUP-04G and SEE-06’s percent correct generalization scores during baseline and modified BST phases.

After the introduction of the modified BST protocol earlier that day before meeting with SEE-04, SUP-04G immediately demonstrated all possible skills on the first session with SEE-06. Supervisor’s performance remained at maximum criteria with the supervisor consistently demonstrating all possible required skills. This was indicative of rapidly generalized skills, under stimulus control. After the third treatment session, the child participant withdrew from the study, for reasons outside of the researcher’s control (i.e., withdrew from ABA company – family moved to another state), at which time data collection ended for this pair. The mean score for this condition was 100 percent. Visual inspection of the data revealed an immediate change in data.
level from relatively medium, to high. There was a very short latency to change with no variability, and zero bandwidth. Performance on this condition is indicative of generalized behavior under stimulus control.

The bottom graph of Figure 4.7 above represents correct responding for SEE-06 on the Y axis. Data displayed for SEE-06 showed an improvement in performance and upward trend during the first four baseline sessions, followed by a slight decline. The mean score for baseline was 32 percent. During this condition, data values fluctuated between low and medium values, which also revealed a large bandwidth (55 points).

After the introduction of the modified BST protocol, SEE-06’s performance reached a score of 60 percent during the first session. Then, performance declined during a session in which the child fell asleep. This was followed by a vast improvement and score of 80 percent on the third and last session. The mean score for this condition was 58 percent. Visual inspection of the data revealed very variable performance, from which little was learned. Levels fluctuated from medium, to medium-low, to high, indicating unpredictable performance. The bandwidth for this condition’s data was large (i.e., 47 data points). Because of participant withdrawal, the researcher was unable to resume data collection with this supervisee.

**Treatment Integrity.** Data were collected on the researcher’s implementation of the modified BST protocol. Figure 4.8 below depicts correct treatment implementation for the Uber Supervisor/Researcher with each of the BCBA participants. As shown, the Uber Supervisor implemented the treatment with 100 percent integrity during two sessions with SUP-01, three sessions with SUP-02, five sessions with SUP-03, and three sessions with SUP-04 and SUP-05. Sessions not shown on the graph were maintenance, where the modified BST protocol was not implemented.
Figure 4.8. Treatment integrity data for the Uber Supervisor with each BCBA participant.

Social Validity Questionnaire. After the conclusion of the study, all five BCBA participants were asked to complete the Supervision Social Validity Questionnaire (see Appendix H for a copy of this document). The questionnaire consisted of four yes/no questions, and one short-essay question. Results were summarized on Table 4.3 below. Concerning question number one, all BCBA participants indicated that the intervention was helpful in improving their supervisory practice. In contrast, when answering the second question, two of the BCBA
COMPETENCY-BASED SUPERVISION IN ABA

participants indicated that they would not continue to use the protocol in their day-to-day supervision, and three of them expressed that they would. Feedback on this question included difficulty to incorporate supervision protocol into programming and schedule, sporadic use, and inability to film supervisee performance on a consistent basis. The third question pertained to whether participants would recommend the intervention to other BCBA supervisors. All BCBAs indicated that they would indeed make this recommendation to their peers. Similarly, the five BCBA supervisors agreed that all skills are important to achieving and maintaining excellent supervisee performance. On the last question – number five, participants were asked to list any skills that did not seem important for a supervisor to demonstrate. Feedback from two of the supervisors listed scheduling the next appointment, and only needing a score of 80 percent (as opposed to 100 percent) to demonstrate supervisor competency.

Table 3
“Yes” answers to questions in the Supervision Social Validity Questionnaire, and comments

<table>
<thead>
<tr>
<th></th>
<th>Question 1</th>
<th>Question 2</th>
<th>Question 3</th>
<th>Question 4</th>
<th>Question 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCBAs</td>
<td>5/5</td>
<td>3/5</td>
<td>5/5</td>
<td>5/5</td>
<td>Scheduling next appt</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Score of 80% for competency</td>
</tr>
</tbody>
</table>
Introduction

As previously mentioned, the intervention assessed was designed in response to a call from the BACB to develop effective supervision training systems for newly certified behavior analysis practitioners (BACB, 2012). To accomplish this purpose, the present study evaluated the use of a modified BST protocol on the performance of BCBA supervisors and their supervisees during the implementation of DTT. The study employed a quantitative, single-case, multiple-baseline across supervisors’ research design to answer three research questions. This design is widely used in the behavioral literature (Cooper, Heron & Heward, 2007; Bailey & Burch, 2018). In fact, White (1993), argued that multiple-baseline designs held potential as evaluation designs for the provision of supervision. Two hypotheses were proposed and later supported by the results of the study. First, it was predicted that, following the introduction of the modified BST protocol, supervisors’ performances would improve from basic to intermediate levels, thereby ascertaining their proficiency past merely completing a training, to validating their supervisory performance in vivo via observation, measurement, and evaluation. Second, it was expected that supervisees’ performances would also improve as a result of improved supervisory performance, thereby suggesting that supervisor behavior is an important variable to consider when evaluating supervisee behavior. Data obtained from the results of the study supported both of the proposed hypotheses.

All five BCBA participants successfully completed the training and implemented the protocol with their respective supervisees. Therefore, as mentioned before, the implementation of the modified BST protocol not only had a greater positive effect on supervisor competency, but also on supervisee’s performances during DTT, when compared to existing practices during
baseline. Moreover, this protocol may stand as one of the foundational empirical studies that assessed the use of a competency-based tool for the evaluation of supervisor competency, thereby filling a current gap in the behavioral and supervision literatures. Overall, behavioral changes occurred immediately for supervisors, and maintained throughout the course of the study. For supervisees, though changes were significant and very positive, for some participants it was slower to come and maintain than others. This suggests that the modified BST protocol was successful in improving supervisor performance in a more immediate manner, while also successfully improving supervisee performance though in a less experimentally reliable fashion. It is hypothesized that client participants also benefited from the implementation of the modified BST protocol at the supervisor’s level, though these data were not collected or analyzed.

**Summary of Findings**

Guided by three research questions, the researcher assessed changes in performance levels for all BCBA participants and supervisees. Starting with the first research question (i.e., Are standard [basic-level of competency] BCBA supervisory practices effective in improving supervisee behavior during DTT?), results showed that overall during baseline, supervisor’s performances remained at basic-levels of supervisory competence, while supervisees’ performances remained between low and medium levels. Specifically, BCBA supervisors primarily demonstrated the following skills during baseline: a) *observing the supervisee*, b) *identifying and documenting relevant supervisee errors*, c) *providing corrective feedback to the supervisee*, d) *coaching and/or modeling appropriate supervisee responses*, and e) *providing positive reinforcement to the supervisee*. Except for *observing the supervisee*, the skills listed above were included in the category of basic-level of supervisory performance. Results suggest that during the study, standard BCBA supervisory practices were not effective in improving
supervisee behavior during DTT. Furthermore, data were indicative of performance limitations, as none of the supervisors demonstrated more than five skills during the baseline condition. This presented a prime opportunity for the introduction of new skills (i.e., the modified BST protocol).

Study results also showed rapid skill acquisition and immediate performance improvements for the BCBA and supervisee participants respectively, after the introduction of the modified BST protocol. Moreover, for all BCBAs and most supervisee participants, these gains were maintained throughout the course of the study. Therefore, the above results appeared to have answered the second research question (i.e., After the modified BST protocol is introduced, will BCBA supervisors demonstrate intermediate levels of supervisory behavior during supervision sessions?), because all BCBA supervisors demonstrated intermediate levels of supervisory performance during all supervision treatment sessions, including protocol implementation and maintenance. As previously stated, intermediate level of competency included the demonstration of the five BST skills, plus five newly-introduced skills. All supervisors missed one or two skills once or twice during the study. Nevertheless, the absence of these skills did not appear to affect supervisee performance. The comparison of overall scores and number of skills demonstrated during the baseline and modified BST conditions were indicative of experimental control, validated by high-average inter-observer agreement values in both conditions. However, it was unclear whether the implementation of the booster session intended to bring scores back up to maximum criteria might have in fact been necessary, versus a simple reminder from the Uber Supervisor to refer back to the list of skills on the protocol often, during supervision sessions with the supervisee.
Supervisee performance also seemed to improve as a result of improvements in supervisory performance. Therefore, results of the study appeared to answer the third research question (i.e., Does the introduction of the modified BST protocol affect the performance of supervisees receiving supervision from BCBAs during the implementation of DTT?). Collateral supervisee data analysis of both baseline and modified BST protocol conditions revealed low to medium levels of performance scores on the first condition with a mean overall value of 45 percent, followed by an immediate increase in performance scores after the implementation of the treatment condition, for four of the five supervisee participants (though all participants showed increases in performance scores). Data also showed that during the modified BST protocol phase, supervisee participants’ performance reached a mean score of 87 percent, indicating improved performance in the application of DTT by supervisees. Therefore, the introduction of the modified BST protocol with supervisors appeared to have positively affected the performance of supervisees during their implementation of DTT. Furthermore, these changes though indicative of behavior under skill acquisition for most of the participants, maintained throughout the course of the study. When comparing overall scores and number of skills demonstrated during the baseline and modified BST conditions, the data presented though less strong, were also indicative of behavior under experimental control, validated by high-average inter-observer agreement values in the treatment condition. Because total inter-observer agreement scores were less than the general requirement for single-case design data (i.e., 80 percent or higher; Bailey & Burch, 2018) during the baseline condition, data on supervisees’ performances should be interpreted with care. Baseline measures lacked overall consistency, which suggests that the findings lacked a certain level of predictability and reliability. This suggests that the behaviors scored by the independent observers (i.e., the primary and reliability
observers) were differently perceived, which continued to be the case even after re-trainings were conducted. It appeared that observers were able to more easily score clear demonstrations of skills during the modified BST condition, than the display of what appeared to be idiosyncratic performance during baseline. This may be one explanation for the less-than-desired observer agreement scores obtained during the baseline phase of the study.

**Interpretation of Findings**

Overall, the study shows that a modified BST-based supervision training protocol was effective in developing the competency and readiness skills of five BCBA supervisors, who regularly engaged in DTT supervision. All supervisors’ data showed rapid skill acquisition after the implementation of the protocol, which maintained throughout the course of the treatment phase, even after the protocol was withdrawn. Four out of the five supervisors demonstrated all the newly acquired skills over three or more sessions, while showing no loss of the skills demonstrated during baseline. Therefore, all supervisors exhibited proficiency past merely completing a training, and validated their supervisory skill performance in vivo via observation, measurement, and evaluation. Moreover, one of the supervisors also exhibited all the skills from the modified BST protocol with a different supervisee. Though more data on maintenance of generalization skills as well as stable supervisee data were needed to support that indeed generalization occurred in the above supervisor’s case, the data obtained showed promising results. Implications of the results indicate that a modified BST protocol may be effectively used to train even experienced BCBAs in competency-based, ethical supervision practices in the area of DTT, as all participants had engaged in the supervision of therapists conducting DTT for well over a year. Additionally, the modified BST protocol appeared to provide structure to the supervision process. The BCBAs seemed more prepared to provide specific feedback with clear
COMPETENCY-BASED SUPERVISION IN ABA

instructions (i.e., a written protocol) to follow, aided by a means to organize and prioritize feedback to the supervisee, as well as a method of supervisee performance evaluation that is DTT-specific and competency-based. The latter produced what appear to be more focused supervision sessions and encouraged the BCBAs to look more closely at their supervisees’ performances, prompting a deeper analysis. It appears that the implementation of the modified BST protocol facilitated the development of antecedent strategies aimed at increasing not only the quality of supervision, but also the quality and number of supervisees’ skills. The structure provided by the modified BST protocol also allowed for a systematic way of fading the supervisors’ assistance after supervisees had mastered all DTT skills. This in turn, enabled a systematic shift from strict supervision to more mentorship-like supervision sessions that focused on maintenance and generalization of the supervisee’s DTT skills. The introduction of the video review (i.e., of the supervisor and supervisee) during Uber supervision sessions appeared to have reinforcing value for the BCBA. The videos did not identify skills that the BCBA was performing incorrectly, but rather provided validation of skills that the BCBA was performing correctly, in a different manner. Supervision sessions conducted in a confidential office setting, while the child was not present, looking at video samples of supervisees’ performances, organized after specific feedback was prioritized, and structured to focus on individualized performance - both correct and in need of improvement, appeared to have increased supervisees’ opportunities to practice DTT skills, in an environment conducive to learning and for the most part, free of distractions. Thus, facilitating a refinement of supervisees’ DTT skills.

Although the modified BST protocol was implemented at the BCBA supervisor level, the study also noticeably shows that supervisee performance improved during the implementation of DTT when supervisors acquired new supervision skills. This may suggest that the
implementation of the independent variable indirectly affected supervisees’ performances. While expected, the results suggest that an intervention provided at the supervisor level, may be effective in improving supervisee behavior, which implies that one important key variable to consider when attempting to change supervisee behavior may be a change/improvement in supervisor behavior via a systematic, competency-based training. In addition, all supervisees acquired skills that were previously not part of their repertoire, thereby demonstrating more complex skills beyond those typically taught during basic DTT training. The acquisition of more complex DTT skills on the part of the supervisees, appeared to have been a result of supervision directly related to the implementation of the modified BST protocol by the supervisor. As predicted, supervisee performance was indicative of behavior under skill acquisition. The main reason for this may have been related to the systematic way in which feedback was provided to the supervisee, prioritizing a small number of skills to address during each session, in a cumulative manner. This also ensured that the supervisees heavily focused on a few skills to improve, while retaining already mastered abilities. Because DTT sessions during treatment took place immediately after the supervision sessions, supervisees were able to recall the skills that had just been practiced. This structure, which is also part of the protocol, may have also been responsible for improvements in supervisee performance. Supervisees had a repertoire of behaviors from which to draw (as a result of having role-played skills during supervision), which allowed them to solve specific previously addressed issues, in the moment (i.e., during the DTT session with the child). The introduction of the video review (i.e., of the supervisees with the children) during supervision sessions in the modified BST condition, appeared to have increased supervisee awareness of their own DTT skills. Their use seemed to serve two purposes: 1) the ability to show supervisee correct/quality performance, which appeared to have reinforcing
value; and 2) the ability for supervisees to see their own mistakes on video, which appeared to have punishing value, and significantly decreased discrepancies during supervision sessions, thereby increasing efficiency, opportunities for practice, and focus. Despite the formerly described promising results, supervisee performance scores should be interpreted with caution. During baseline, mean inter-observer agreement scores for four out of the five participants were below 80 percent, which may suggest that supervisees’ performances during DTT were not universal, but rather idiosyncratic to each therapist, context, session, and client. Consequently, not all skills were shown in all supervisees. Close data inspection revealed that observers appeared to disagree more often when behaviors were not clearly present or did not apply, than when they were distinctly demonstrated during the treatment phase. In addition, as previously mentioned the modified BST protocol was implemented at the BCBA supervisor level. Thus, the independent variable for supervisee behavior was uncontrolled, which added to the arguments above, may account for less than desired inter-observer agreement scores. Nevertheless, apparent dramatic changes in behavior were observed in supervisees’ performances from baseline to treatment, which once again demonstrates, to a lesser level, experimental control. During baseline, though the types of skills demonstrated on each session varied, the supervisees mostly engaged in the following: phrasing the SD as an instruction, differentially reinforcing correct responding, interspersing skills, accurately collecting data, and providing different types of praise. In addition to the aforementioned skills, supervisees consistently demonstrated the following abilities during the modified BST condition: conducting a preference assessment, using different ways to call attention, contriving manding opportunities, shaping, and incorporating previous feedback. The acquisition and maintenance of these skills throughout the
treatment phase of the study, supported the hypothesis that supervisees’ performances improved as a result of improvements in supervisory performance.

On a smaller scale, the results also showed that supervisor performance, while it remained at high levels throughout the modified BST phase, sometimes required one or two booster sessions to bring performance levels back up to maximum criteria. Close data inspection and identification of consistently-missed skills across supervisors showed that the skills the supervisors failed to demonstrate during a few treatment sessions had no bearing on supervisees’ performance improvements. This may suggest that skills like *prompted note-taking, prompted the supervisee to engage in relevant skills* and *scheduled a follow-up appointment* may not have been directly responsible for improvements in supervisee performance, and may not have been strictly necessary for the supervisor to demonstrate.

Results of the social validity questionnaire suggested that all participants appreciated the introduction of new supervisory skills during the treatment phase, as the protocol was helpful in improving their supervisory practice. They also agreed in their future recommendation of the modified BST protocol to colleagues, and the importance of the skills learned in achieving and maintaining excellent therapist performance. However, two of the five BCBA participants indicated that they would not continue to use the full modified BST protocol in their day-to-day supervision. One of the participants alleged that the protocol was difficult to incorporate into their program schedule but would continue to use the *Evaluation of Supervisee Performance Form* (see Appendix D for a copy of this document) to monitor performance of each skill, and provide feedback to supervisees. The second supervisor indicated sporadic use, but not on a daily basis. These comments are important to consider. Failing to explore the reasons why supervisors will not utilize the protocol on their day-to-day supervision, may potentially pose a threat to both
the generalization and technological dimensions of ABA (Baer, Wolf, & Risley, 1968), as well as the potential social validity of the modified BST protocol.

**Context of Findings**

Findings from the present study relate to the literature discussed in several ways. First, the study examined supervision practices in the field of ABA, and more specifically in the area of DTT. Though supervision had been examined in a variety of fields, no study to date provided an empirical examination of supervisor behavior in this field, or more specifically in the area of DTT. Second, the findings attempted to fill a gap in the literature by empirically examining a competency-based method of teaching BCBA supervisors how to engage in the ethical and effective practice of supervision. Third, the study added to the behavioral research literature in supervision by answering research questions utilizing one of the most widely-used behavioral research designs: the multiple baseline. Once again, no study to date has employed an empirical, well-established research design to examine behavior analysts’ supervisors’ performances in the area of DTT. Fourth, the study also included a procedural fidelity component for the implementation of the independent variable. This is also the first study of its kind that provides a measure of treatment adherence in the area of behavior analysis supervision, that is specific to DTT. Finally, the current study utilized a modified version of a promising evidence-based method: BST, thereby augmenting the existing literature on BST into a new area: supervision of DTT.

Previous research in the area of supervision revealed diverse discussions and approaches to the study of this topic. One of the first major contributions originated from the fields of psychology and counseling, where supervision was discussed from the viewpoint of multitherapy models (Schoenwald et al., 2000; Schoenwald et al., 2000; and Henggeler et al., 2002), within
the context of school psychology (Welsh et al., 2003), followed by developmental approaches to supervision (Barret & Barber, 2005; Stoltenberg, 2005), as well as the creation of a competency-based supervision tool (Madson et al., 2005). This led to the subsequent development of the APA guidelines for supervision (APA, 2006) and a variety of models for supervision spearheaded by the works of Falender and Shafranske (2007). The aforementioned authors also targeted supervision from a competency-based methodology, as did Gonsalvez and Calvert (2014), and Stein et al. (2015). Because the current study evaluated the use of a competency-based protocol that may provide data-based support of the its use in the supervision of DTT, evidence-based approaches to supervision were also examined. Milne and Reiser (2012) provided a rationale for the need for evidence-based clinical supervision of graduate-level students in psychology.

Similarly, Leffler et al. (2012) described several training modules used to enhance skills and techniques of EBP in psychology trainees.

Other supervision models examined included attachment-informed supervision (Bennett, 2008), couple and family therapy supervision (Celano at al., 2010), supervision within the context of play therapy Hudspeth (2015), and supervision in clinical neuropsychology (Stucky et al., 2010). Ng et al., (2012) developed an international competency-based supervision checklist and encouraged U.S. counseling training programs to consider using the 43 items included as a set of primary indicators to guide them in the propagation of their work on international grounds. One year later, Hatcher et al., (2013) presented a description of the latest version of “The Competencies Benchmarks for Professional Psychology” (p. 84).

Relevant to the current study, previous research in competency-based training yielded several investigations of this topic in the field of medicine. Literature included the use of competency-based training in palliative care (Yuen et al., 1998), the training of physicians
employed full-time in occupational and environmental medicine (Emmett & Green-McKenzie, 2001), the training of residents in psychiatry (Louie et al., 2004), the evaluation of the implementation of post-abortion care services outside of major urban centers in Nepal (Basnet et al., 2004), the training of health professionals on responding to weapons of mass destruction events (Pryor et al., 2006), and the creation of an internationally acceptable competency-based training program for specialists in intensive care medicine (Rubulotta et al., 2009).

Early behavioral research in supervision focused on changing supervisor behavior (Rosenbaum, 1975; Ivancic et al., 1981; and Methot et al., 1996), as well as supervisees’ behaviors (Dillon et al., 1980; and Hundert & Hopkins, 1992). Recent research on supervision in behavior analysis has focused on fundamental recommended practices for the supervision of aspiring behavior analysts (Sellers, et al., 2016), discussing rationales and consequences related to supervision concerns as well as providing guidelines for professional development in supervisory areas within the BACB ethics code (Sellers et al., 2016), a systematic approach to issues that may arise during the supervisory relationship and guidance for supervisors to identify and address barriers to successful supervision (Sellers, et al., 2016), the provision of a practice model and considerations for supervising ABA students (Turner et al., 2016; Hartley et al., 2016; Garza et al., 2018), and how to plan and conduct efficient group supervision for ABA students (Valentino et. al., 2016).

The current study also evaluated the researcher’s adherence to the modified BST protocol proposed. Early research on this topic focused on new possibilities of treatment adherence was spearheaded by the works of (Hogue et al., 1996). In addition, (Hogue, 2008) assessed the fidelity, adherence to evidence-based treatment, competence, and treatment outcomes of adolescents exhibiting substance abuse and behavior challenges. (Sheidow et. al., 2008)
evaluated clinician fidelity to an evidence-based treatment for adolescents engaging in substance abuse. Later, (Hogue et. al., 2013) discussed instrumental shifts in operating practices and procedures aimed at fortifying evidence-based practices and insulating against protocol drift and erosion, which was followed by a study on the use of a multi-method assessment of treatment fidelity to a family therapy approach by front-line therapists in a community behavioral health clinic (Hogue & Dauber, 2013), and assessing the reliability and accuracy of groups of community therapists who reported on their use of FT and motivational (MI/CBT interventions during routine treatment of adolescents with conduct and substance use problems (Hogue et. al., 2015).

BST, an evidenced-based intervention, on which the researcher relied to develop the independent variable for this study, holds a robust place in the behavioral literature. As previously mentioned, BST has been successfully utilized to teach appropriate social behaviors, and social-vocational skills to individuals with intellectual disabilities and behavior challenges (Miltenberger, 2000; Burke et al., 2010). In addition, BST has been successfully employed in the teaching of gun safety, as well as firearm injury and gun play prevention (Jostad, et. al., 2008; Kelso et. al., 2007; Gross et. al., 2007). Moreover, BST is a valuable protocol to utilize when teaching individuals with intellectual disabilities abduction and sexual abuse prevention skills (Johnson, et al., 2005; Egemo-Helm et al., 2007; Lee & Tang, 1998). Lastly, BST has also been applied to train adults with intellectual disabilities to implement DTT-related procedures (Lerman et. al., 2013; Lerman et. al., 2015). A review of the aforementioned literature offered support for the effectiveness of a modified BST approach when applied to the practice of supervision, based on its proven effectiveness for various clinical uses.
Implications of Findings

The current study findings have several implications. Theoretically, the findings suggest that an empirical, behavioral, competency-based protocol could be successfully utilized in changing supervisory behavior in the area of DTT. These findings also extended to supervisees’ performances and hold the potential to also affect the behavior of clients who receive behavior analytic services, specifically discrete trials. Therefore, training programs that rely on behavioral theories and utilize behavioral methodology such as BST have the potential to effectively improve supervisory behaviors, specifically in skill acquisition arenas. What’s more, the results suggest that a carefully crafted behavioral protocol can positively affect the behaviors of trainees when implemented at the supervisor level.

In terms of research, the current study may have further unlocked a new line in the area of organizational behavior/performance management, and more specifically in the realm of competency-based supervision. While fulfilling one of the many gaps in this research topic, this study proposed a new way of operationalizing, measuring and evaluating appropriate supervisor behaviors. This may allow future researchers to conceptualize and operationalize the practice of supervision in a novel way, as well as examine other areas of supervision in behavior analysis perhaps utilizing a similar approach. At the same time, the study evaluated a new way of conducting supervision, and provided empirical support for its efficacy. Future researchers may be interested in replicating this study perhaps in other disciplines outside of behavior analysis. Additionally, this study sought to change supervisee behaviors through the implementation of the independent variable at the supervisor level. Positive findings at the supervisee level suggest that interventions implemented at the supervisor level have the potential to also change supervisee behavior. This was an interesting finding, and one that may have practical implications for
COMPETENCY-BASED SUPERVISION IN ABA

clinical practice. Though it appears evident to assume that improvements in supervisors’ behaviors would also translate to improvements in supervisees’ behaviors, this study provided empirical support for this belief. In addition, it may serve as the foundation for future studies in supervision and training where the independent variable is only applied at the supervisor/manager level. Further investigational implications of the study may include component analyses aimed at identifying the most critical elements of the intervention packet, as well as the development of ways to increase interobserver agreement of supervisees’ behaviors resulting from idiosyncratic performance.

From a practical standpoint, the study offers a way to train and evaluate supervisor behavior that involves objective observation and data collection. This protocol may potentially be utilized by any supervisor with access to a camera and a computer. Therefore, this method could be used for the training of prospective ABA supervisors (i.e., student-analysts), as well as the training and supervision of technicians, parents, lead therapists, or anyone who may oversee the implementation of DTT. Additionally, this study has implications for the practice of remote or distance supervision. The use of video recordings of supervisor and supervisee performances to collect, analyze and evaluate clinical supervisor-supervisee data may allow for more efficient and ethical supervision practices in lieu of in-person supervision. Still, though this may be an exciting implication for the area of distance supervision, it lacks empirical support.

Limitations of the Study

Perhaps the biggest limitation of this study, is that it evaluated a supervision model where the supervision sessions themselves were not conducted during DTT sessions, in the presence of the child, as they typically occur, but rather in an office setting away from the child. Current billing practices only allow for supervision to be conducted while the client is present. Therefore,
COMPETENCY-BASED SUPERVISION IN ABA

this protocol though successful, is not currently supported by funding agencies, which may severely limit implementation by BCBAs in clinical practice. Perhaps the results of this study might provide evidence toward a shift in supervision practices that are more systematic, competency-based, ethical, and in the long run, effective. Similarly, funding agencies might rely on the results of this study to support competency-based supervision practices. In addition, the provision of supervision in the manner conducted in this study might limit the supervisor’s ability to provide feedback in vivo, when unique problems arise during DTT sessions, perhaps related to situations newly encountered by supervisees.

The protocol required supervisors to carve time out of their day to observe, document, prioritize and prepare materials as well as the feedback they would provide to the supervisees just before their upcoming sessions. This suggests an increase in response effort for the supervisor, which may also limit the frequency at which they may implement this protocol. This limitation was further corroborated by two of the five supervisors who completed the Supervision Social Validity Questionnaire (see Appendix H for a copy of this document).

ABA therapy sessions most often occur in the clients’ natural environment (e.g., their home). One limitation of this study is that it is unknown whether the modified BST protocol could be successfully implemented in a natural environment setting, which may discourage BCBA supervisors from using the protocol. Though a potential limitation of the study due to confidentiality issues in environments where there is more than one client and supervisee present in the same room, the use of videotaping may also not be a readily available option to supervisors, and may therefore limit the provision of accurate and effective feedback, as well as opportunities for the supervisees to practice specific skills’ improvements. Therefore, going forward it is unknown what results researchers would obtain if they had less resources.
Another limitation of the study was the IOA scores obtained for supervisee performance. Though supervisees’ IOA scores appeared to improve during the treatment phase, it seemed that idiosyncratic performance may have been responsible for the disagreements and somewhat lower scores throughout the study, especially across the pre-treatment phase. This limits the level of experimental control that may be inferred, the robustness of the intervention, and may question the researcher’s decision to implement treatment with less than desired IOA results. Still, because the study was conducted in clinical settings, the demands of each situation required intervention, in favor of avoiding risking participant withdrawal due to running the study beyond the timeframe that subjects were prepared to commit. In addition, it may be worth taking a closer look at each supervisee’s idiosyncratic performance, to determine variables that may contribute to variability.

The current study focused on a specific population: children diagnosed with developmental disabilities and who did not engage in severe behavior challenges. It is unknown whether this protocol might be effectively used with different client populations such as teens and adults diagnosed with intellectual disabilities, as well as typically-developing children, and children, teenagers and adults who exhibit severe behavior challenges. Likewise, the study concentrated on ABA supervision in a narrow area of skill acquisition: DTT, which limits the use of this protocol solely to this behavior-change practice.

Similarly, the current investigation did not address supervisors’ use of appropriate social (soft) skills during supervision/feedback sessions. Supervisors’ use of soft skills are essential characteristics of well-rounded, efficient supervisors. Therefore, for seasoned BCBA supervisors, this protocol may appear incomplete.
COMPETENCY-BASED SUPERVISION IN ABA

The current research did not explore the effectiveness of specific types of feedback techniques. This may also limit the effectiveness of the protocol, as ineffective feedback techniques may hinder supervisee performance improvement, and ultimately negatively affect supervisors-supervisees’ relationships.

Finally, the social validity questionnaire was provided to a small sample of participants (i.e., five). The researcher distributed the questionnaire to the participants, as well as collected them and analyzed the results. This means that the results may not be pure, which poses a limitation to the validity of these results.

Future Research Directions

The present study sets the stage for future research in the area of ABA supervision. As previously mentioned, the intervention assessed was designed in response to a call from the BACB to develop effective supervision training systems for newly certified behavior analysis practitioners (BACB, 2012). Additionally, a review of the literature in supervision revealed a very limited number of peer-reviewed studies, which makes it an area ripe for exploration.

There were several phenomena in need of further study which investigators may want to focus on for future research. The researcher found that use of the DTT Performance Feedback Sheet (see Appendix E for a copy of this document) was useful in allowing the BCBAs to identify additional relevant skills (i.e., repeated errors) that were not part of the original Evaluation of Supervisee Performance Form (see Appendix D for a copy of this document). This increased the level of flexibility related to the use of the protocol and provision of feedback to the supervisees. Perhaps future researchers may wish to concentrate on evaluating the current protocol with the addition of relevant DTT skills not currently included on the Evaluation of Supervisee Performance Form (see Appendix D for a copy of this document).
One of the strengths of the study was the use of videotaping to provide feedback to supervisees. As previously mentioned, introducing the video review (i.e., of the supervisees with the children) during supervision sessions in the treatment condition appeared to serve to show supervisees’ correct/quality performance, which seemed to have reinforcing value; and allowed supervisees to see their own mistakes on video, which seemed to have punishing value. Because the use of videotaping may not be a readily available option to supervisors, future researchers may want to replicate this study without the use of the video reviews for supervisees.

Not all ABA therapy sessions occur in a clinic setting. In fact, the majority of them take place in the clients’ natural environment (e.g., their home). Future research may want to replicate this study in natural environment settings where supervision sessions most often take place. It would be interesting to also evaluate the role of technology and financial feasibility on supervisors’ willingness to utilize this protocol in natural environment settings. Similarly, supervisee idiosyncratic performance may increase in every day and community settings, which may simultaneously decrease inter-observer agreement scores. To investigate these phenomena, future research may focus on closely examining each supervisee’s idiosyncratic performance, to determine and better understand variables that may contribute to variability. This may allow researchers to develop ways in which to help clinical supervisors capitalize on supervisee idiosyncratic performance, with the ultimate goal of optimizing clients’ behavior change.

ABA services are also provided to populations other than children diagnosed with intellectual disabilities (i.e., like the participants in this study). Future research should examine whether this protocol might be effectively used to train supervisors in the supervision of appropriate skill acquisition programs for teens and adults diagnosed with intellectual disabilities, typically-developing children, as well as training supervisors in the supervision of
appropriate behavior reduction programs for children, teenagers and adults who exhibit severe behavior challenges.

The current protocol may have laid the foundation for the use of competency-based protocols in the practice of supervision, not only in the ABA field, but also in similar disciplines such as clinical psychology, school psychology, neuropsychology, mental health counseling, and marriage and family therapy, among others. The systematic, empirical way in which supervisor and supervisee behavior change was accomplished in this study, may provide other disciplines with a blueprint from which to draw in the development of their own competency-based supervision protocols. Therefore, future research may focus on the adaptation of this protocol to the training of supervisors in competency-based supervision across different disciplines.

Because the current study concentrated on ABA supervision in a narrow area of skill acquisition: DTT, the use of this protocol is limited solely to this behavior-change practice. This presents an opportunity for the development of competency-based supervision protocols aimed at training current and future supervisors in the supervision of other behavior-change procedures and systems such as behavior contracts, group contingencies, self-management strategies, precision teaching, and functional communication training, among others. Consequently, future research may focus on replicating the structure and systematicity of this protocol to be utilized in the development of competency-based protocols to train supervisors in the supervision different behavior-change systems and procedures.

As previously mentioned, supervisors’ use of soft skills are essential characteristics of well-rounded, efficient leaders. This was an area that the current investigation did not address. Due to the importance of appropriate supervisory social skills in the development and
COMPETENCY-BASED SUPERVISION IN ABA

maintenance of supervisor-supervisee relationships, future investigators may wish to replicate this study while also evaluating BCBAs’ social skills during supervision/feedback sessions.

The current research did not explore the effectiveness of specific types of feedback techniques on supervisor and supervisee performance improvements. Future researchers may wish to conduct a component analysis aimed at identifying and determining specific types of feedback that are most effective, versus ineffective and detrimental feedback techniques. Likewise, the use of a component analysis in future investigations may help ascertain particular elements of the current protocol that are actually responsible for the behavior-changes observed for both supervisors and supervisees.

Ultimately, behavioral research should be aimed at evaluating socially-valid independent variables that positively affect the consumer’s quality of life, independence, as well as access to reinforcement (Bailey & Burch, 2018). An original aim of this investigation was to measure client performance. However, results of the pilot study conducted by the researcher identified a number of confounding variables that threatened the level of experimental control which the results of the current study could soundly claim to achieve, had the researcher also endeavored to measure behavior change at the client level. This was especially true considering that the independent variable was only implemented at the supervisor level, which prevented tight experimental control of the independent variable for supervisee behavior-change. Thus, future replications of this study should focus on developing ways to decrease confounding variables related to the collection of behavior-change data for child participants, to examine whether an intervention implemented at the supervisor level would also positively affect client performance.

As the field of ABA grows, the demand for distance or remote supervision options appears to have also increased. This is an area/supervision modality largely unexplored by
COMPETENCY-BASED SUPERVISION IN ABA

current behavioral researchers. Because the present study offers an empirical, systematic way of training supervisors and changing supervisee behavior in the area of DTT, future research may focus on the replication of this study for remote supervision, first in the area of DTT, and perhaps then in other skill acquisition and behavior-reduction areas.
COMPETENCY-BASED SUPERVISION IN ABA

References


COMPETENCY-BASED SUPERVISION IN ABA


COMPETENCY-BASED SUPERVISION IN ABA


COMPETENCY-BASED SUPERVISION IN ABA


COMPETENCY-BASED SUPERVISION IN ABA

Treatment adherence and differentiation in individual versus family therapy for adolescent substance abuse. *Journal of Counseling Psychology, 45*(1), 104-114.


COMPETENCY-BASED SUPERVISION IN ABA


COMPETENCY-BASED SUPERVISION IN ABA


COMPETENCY-BASED SUPERVISION IN ABA


COMPETENCY-BASED SUPERVISION IN ABA


Appendix A:

Research Supporting Instrument Development
COMPETENCY-BASED SUPERVISION IN ABA

Development of the Evaluation of Supervisor Performance Form

Each skill comprised within the Evaluation of Supervisor Performance form (see Appendix C for a copy of this document) corresponds to either an important feature of supervision (Behavior Analyst Certification Board, 2012) or BST in the following manner: (a) observed trainer (i.e., the first skill listed under the form referenced above) corresponds to scheduled observations before, during, and after training, listed under the second area of the supervisor curriculum (Behavior Analyst Certification Board, 2012); (b) identified and documented relevant trainer errors (i.e., the second skill listed under the form referenced above) corresponds to assess application, listed under the third area of the supervisor curriculum (Behavior Analyst Certification Board, 2012); (c) prioritized corrective feedback (i.e., the third skill listed under the form referenced above) corresponds to providing a rationale for the target skills to be trained, listed under the third area of the supervisor curriculum (Behavior Analyst Certification Board, 2012); (d) prompted note taking (i.e., the fourth skill listed under the form referenced above) corresponds to requiring a written description of the target skills to be trained, listed under the third area of the supervisor curriculum (Behavior Analyst Certification Board, 2012); (e) provided corrective feedback (i.e., the fifth skill listed under the form referenced above) corresponds to providing corrective feedback to the supervisee, listed under the third area of the supervisor curriculum (Behavior Analyst Certification Board, 2012); (f) coached and/or modeled appropriate trainer responses (i.e., the sixth skill listed under the form referenced above) corresponds to demonstrating target skills, listed under the third area of the supervisor curriculum (Behavior Analyst Certification Board, 2012); (g) provided opportunities for role play/practice (i.e., the seventh skill listed under the form referenced above) corresponds to requiring supervisees to rehearse each target skill, listed under the third area of the supervisor curriculum.
compétence de supervision en ABA (Behavior Analyst Certification Board, 2012); (h) provided positive reinforcement (i.e., the eighth skill listed under the form referenced above) corresponds to providing positive feedback to the supervisee, listed under the third area of the supervisor curriculum (Behavior Analyst Certification Board, 2012); (i) prompted trainer regarding engagement in relevant skills (i.e., the ninth skill listed under the form referenced above) corresponds to providing vocal instructions of the target skills, listed under the third area of the supervisor curriculum (Behavior Analyst Certification Board, 2012); and (j) scheduled follow-up appointment (i.e., the tenth skill listed under the form referenced above) corresponds to ongoing scheduled supervision meetings, listed under the second area of the supervisor curriculum (Behavior Analyst Certification Board, 2012). The researcher deemed the aforementioned skills to be the most relevant for assessment of supervisory behaviors in this study.

For the development of this checklist, the researcher selected specific elements listed within the second and third areas of the Supervisor Training Curriculum Outline (Behavior Analyst Certification Board, 2012) entitled “Important Features of Supervision” (Behavior Analyst Certification Board, 2012 p. 2), and “Behavioral Skills Training” (Behavior Analyst Certification Board, 2012 p. 3), respectively. Supervisors’ implementation fidelity will be determined by dividing the number of steps performed correctly by the total number of steps, multiplied by 100.

Checklists have been used in the behavioral literature to improve staff performance. One example includes a study by Bacon, Fulton, and Malott (1983) where the authors provided 14 students with checklists and required them to check off tasks that they completed. Results of the study revealed that the percentage of tasks completed improved by 28.8% when the checklists were used. In another study, Johnson, Welsh, Miller, and Altus (1991) utilized job checklists as
COMPETENCY-BASED SUPERVISION IN ABA

part of a work-sharing system to maintain the performance of seven staff members who worked at a behaviorally run housing cooperative. Results of the study showed that use of the staff management system was effective in increasing the number of tasks performed by six of the staff members. Similarly, Belisle, Rowsey, and Dixon (2016), utilized BST and the “PEAK Implementation Checklist” (p. 73) to improve staff implementation of the “PEAK Relational Training System” (p. 72). Results of the study suggested that the use of BST and the “PEAK Implementation Checklist” (p. 73) improved staff implementation of PEAK.

The third area of The Supervisor Training Curriculum Outline – BST, focuses on this specific protocol. As mentioned in the section above, a number of the skills within BST were selected for this study. The first three skills listed within the BST protocol described by the BACB correspond again to more general features of this protocol that will be addressed during the initial training session. During this first training session, the researcher and/or researcher assistants will provide BCBA participants with a rationale for the implementation of the training, as well as oral and written descriptions of the modified BST protocol. After the aforementioned step has taken place, the researcher will concentrate on the refinement of the skills that have already been taught utilizing the last five features listed within the BST protocol described by the researcher in the above section. The supervision sessions in this study will take place in person, within an ABA clinic setting. Sessions will be conducted in a clinic room, where both supervisor and trainer will be sitting next to each other, facing a computer that will display a video of the trainer engaging in a previously observed (by the BCBA) DTT session filmed by the researcher. The DTT video of the trainer working with the client will be ready to be played and discussed during the supervision session. Prior to this session, the supervisor would have already observed the trainer and documented as well as prioritized relevant errors aided by the DTT Performance
**COMPETENCY-BASED SUPERVISION IN ABA**

*Feedback Sheet* (see Appendix D for a copy of this document). The supervisor will greet the trainer and inform him/her that they will be discussing the trainer’s performance during the last DTT session, supported by video data and notes that the supervisor has collected and prioritized after observation. Following this step, the supervisor will provide the trainer with the selected corrective feedback and prompt the trainer to take notes. In addition, the supervisor and supervisee will review the video sample of the trainer’s performance together during the individual supervision session. As such the supervisor in this study will be utilizing the features of BST described above to improve a fundamental behavior change procedure: DTT. Then, the BCBA participant will coach and/or model appropriate trainer responses. Next, the supervisor will explain to the trainer that role-play will (or is about to) take place. Role-play will then occur as many times as needed until the trainer demonstrates competency of the skills identified by the BCBA participant as needing improvement. The supervisor will provide positive reinforcement for skills performed as previously demonstrated by the BCBA and remind the trainer of the skills that he/she will need to concentrate on, during the following DTT session. This step was designed to ensure that the skill just practiced, stays fresh in the trainer’s mind. Lastly, the supervisor will schedule a follow-up supervision session appointment.

The literature behind prompting and its benefits for skill acquisition and maintenance (as the proposed study will be assessing supervision within DTT) provided a basis and justification for its selection (Morse & Schuster, 2004). Handen and Zane (1987) reviewed 26 studies that used the delayed prompt procedure in applied and laboratory settings with children, adults and others, diagnosed with intellectual disabilities. Results suggested that the delayed prompt procedure was an efficient teaching strategy. More recently, Tekin-Iftar, Olcay-Gul, and Collins (2019) conducted both a descriptive and meta-analysis of simultaneous prompting (SP) studies.
COMPETENCY-BASED SUPERVISION IN ABA

through 2017. They concluded that “there was a substantial body of evidence to support the SP procedure to teach a variety of skills to individuals with disabilities when consistent parameters of the procedure are employed” (p. 309). Also, Aljehany and Bennett (2019) conducted a meta-analysis of video prompting (VP) to teach daily living skills (DLS) to individuals with autism spectrum disorder. They concluded that VP had a moderate effect size on the acquisition of DLS within the population studied.

As deducted from the pilot study’s data, supervisee’s numbers should show a steady upward trend indicating improvements in performance from low to high levels related to the acquisition of new skills, when compared to baseline data, where feedback was not prioritized. This would provide support for prioritizing skills versus overwhelming the supervisee with too much corrective feedback. Finally, during the conceptualization phase of this study, a group of graduate-level students of ABA as well as BCBAs agreed that this would be an important skill to include in the new protocol, and expressed that they would have liked to have experienced the type of supervision where the supervisor prioritized the corrective feedback that they received.

Development of Evaluation of Supervisee Performance Form

The second skill listed under the form referenced above (i.e., uses different ways to call attention), corresponds to “establish ready behavior” (Catania, Almeida, Liu-Constant, & Reed 2009 p. 389), the first step described on the study’s Table 1. In addition, the third skill listed under the form referenced above (i.e., phrases S\textsuperscript{D} as instruction), corresponds to “State S\textsuperscript{D} as specified in the lesson plan” (Catania, Almeida, Liu-Constant, & Reed 2009 p. 389), the fourth step described on the study’s Table 1. Similarly, the fourth and ninth skills listed under the form referenced above (i.e., differentially reinforcing correct responding and uses praise, respectively), correspond to “deliver reinforcer as specified in the lesson plan” (Catania,
COMPETENCY-BASED SUPERVISION IN ABA

Almeida, Liu-Constant, & Reed 2009 p. 389), the sixth step described on the study’s Table 1. Finally, the eighth skill listed under the form referenced above (i.e., accurately collects data), corresponds to “accurately record data” (Catania, Almeida, Liu-Constant, & Reed 2009 p. 389), the ninth step described on the study’s Table 1.

Similar to the current study, the authors selected the percentage of DTT components correctly implemented during five-trial performance probes as the primary dependent variable for their study. These components included: (a) gaining the confederate’s attention; (b) giving the confederate at least one second of eye contact; (c) using the correct instruction; (d) stating the instruction correctly without repeating it within five seconds; (e) using the correct form, sequence, and timing of prompts; (f) delivering the correct form for reinforcement; (g) delivering reinforcement within five seconds of the response; (h) removing the instructional materials within five seconds of delivering the reinforcer; (i) providing no more than 30 seconds of access to the reinforcer; (j) withholding reinforcement for no responses, incorrect responses, and physically prompted responses; and (k) accurately collecting data. For the purposes of this study, skills (a), (c), (d), (f), and (k) were selected in the development of the Evaluation of Supervisee Performance form (see Appendix D for a copy of this document).

The remaining skills, conducts preference assessment (i.e., the first skill listed on the form referenced above), contrives manding opportunities (i.e., the fifth skill listed on the form referenced above), intersperses skills (i.e., the sixth skill listed on the form referenced above), shaping (i.e., the seventh skill listed on the form referenced above), and incorporates previous feedback (i.e., the tenth skill listed on the form referenced above) were selected after surveying several BCBA supervisors in the field, including the researcher, regarding higher level skills within DTT that may make the difference between initially trained skills, and skills that develop
after ongoing supervision. In this way, the researcher aims at differentiating training from supervision, in that supervision is the refinement of the skills that have already been trained. For example, in clinical practice, therapists are trained to select potential reinforcers to use during the DTT session. Because reinforcers are relative, and a function of motivation, a highly skilled DTT trainer should always conduct a preference assessment prior to the start of the DTT session, to ensure the client’s high motivation for that potential reinforcer. Similarly, though most trainers are taught how to conduct manding trials during their initial clinical trainings; and knowledge of mands (Skinner, 1957) as part of verbal behavior is a topic widely taught in graduate school education, contriving mands is considered among clinicians to be a higher-order skill. The same is true for interspersing skills. While trainers are initially taught how to present skills, the interspersing of operants during a DTT trial are, in the researcher’s experience, the result of ongoing supervision and practice. Shaping is another skill that requires clinical practice, as well as the provision of continuous feedback during supervision; according to the researcher’s own clinical experience. Lastly, incorporating previous feedback has been a problematic area often encountered in the researcher’s practice. Whilst trainers may indicate understanding of the feedback provided, their verbal behavior (i.e., saying that they will incorporate the feedback) and their actual execution of the feedback provided may sometimes greatly differ.

**Development of the DTT Data Collection Sheet**

Comprised of seven components, the aforementioned instrument includes: (a) gaining the client’s attention appropriately, (b) providing the correct $S^D$ as specific in the skill acquisition plan, (c) correctly prompting (if necessary) the response, (d) reacting to the response appropriately, (e) providing the correct consequence, (f) accurately collecting data, and (g) removing stimuli following each trial. These components correspond to teaching steps one, three,
COMPETENCY-BASED SUPERVISION IN ABA

four, five, six, seven, nine, and 10 from Catania, Almeida, Liu-Constant, and Reed (2009).

Similarly, the same components listed above and comprised within the *DTT Data Collection Sheet* form (see Appendix D for a copy of this document) also correspond to components (a), (c), as well as (d) through (k) from the Lerman, Hawkins, Hoffman, and Caccavale (2013) study described above. The components are the following: (a) gaining the confederate’s attention; (c) using the correct instruction; (d) stating the instruction correctly without repeating it within five seconds; (e) using the correct form, sequence, and timing of prompts; (f) delivering the correct form for reinforcement; (g) delivering reinforcement within five seconds of the response; (h) removing the instructional materials within five seconds of delivering the reinforcer; (i) providing no more than 30 seconds of access to the reinforcer; (j) withholding reinforcement for no responses, incorrect responses, and physically prompted responses; and (k) accurately collecting data.
Appendix B

Research on Discrete Trials Teaching (DTT)
COMPETENCY-BASED SUPERVISION IN ABA

For the development of this instrument, the researcher initially relied on two research studies on discrete trials teaching (DTT). The first study by Catania, Almeida, Liu-Constant, and Reed (2009) utilized video modeling to train staff on the implementation of this fundamental element of behavior change in ABA. As part of the video utilized to demonstrate the DTT skills, a voiceover script was incorporated, which provided an explanation of each of the 10-modeled teaching steps. Additionally, a list of the skills was also supplied in the form of a table on the publication, from which the researcher selected some of the initial skills on the Evaluation of Supervisee Performance checklist (see Appendix D for a copy of this document). The second study was conducted by Lerman, Hawkins, Hoffman, and Caccavale (2013). In their study, the authors used BST to pilot a vocational program in which they taught adults with high-functioning Autism (HFA) how to conduct DTT. Similar to the proposed study, the authors selected the percentage of DTT components correctly implemented during five-trial performance probes as the primary dependent variable for their study. Additionally, some skills were selected after surveying several BCBA supervisors in the field, including the researcher, regarding higher-level skills within DTT that may make the difference between initially trained skills, and skills that develop after ongoing supervision. In this way, the researcher aims at differentiating training from supervision, in that supervision is the refinement of the skills that have already been trained. For example, in clinical practice, therapists are trained to select potential reinforcers to use during the DTT session. Because reinforcers are relative, and a function of motivation, a highly skilled DTT trainer should always conduct a preference assessment prior to the start of the DTT session, to ensure the client’s high motivation for that potential reinforcer. Similarly, though most trainers are taught how to conduct manding trials during their initial clinical trainings; and knowledge of mands (Skinner, 1957) as part of verbal behavior is a topic
widely taught in graduate school education, contriving mands is considered among clinicians to be a higher-order skill. The same is true for interspersing skills. While trainers are initially taught how to present skills, the interspersing of operants during a DTT trial are, in the researcher’s experience, the result of ongoing supervision and practice. Shaping is another skill that requires clinical practice, as well as the provision of continuous feedback during supervision; according to the researcher’s own clinical experience. Lastly, incorporating previous feedback has been a problematic area often encountered in the researcher’s practice. Whilst trainers may indicate understanding of the feedback provided, their verbal behavior (i.e., saying that they will incorporate the feedback) and their actual execution of the feedback provided may sometimes greatly differ.
Appendix C

Evaluation of Supervisor Performance Form
## EVALUATION OF SUPERVISOR PERFORMANCE

**BCBA Supervisor:** __________________________  **Bln/Tx:** ________________

**Observer:** __________________________  **Student’s initials:** __________

**Trainer:** __________________________  **Duration:** __________

**Directions:** Indicate whether the skills listed below were present during the observation by placing a checkmark (✓) either on the yes or no column on each trial. Place a checkmark under the N/A column if a particular skill did not apply within the observation. Use tally marks under each column when appropriate.

<table>
<thead>
<tr>
<th>Dates</th>
<th>Skill</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
<th>Session 5</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed trainer</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>identified &amp; documented relevant trainer errors</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Prioritized corrective feedback</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Prompted note taking</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Provided corrective feedback</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Coached &amp;/or modeled appropriate trainer responses</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Provided opportunities for role play/practice</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Provided positive reinforcement</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Prompted trainer res: Relevant skills</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Scheduled follow-up appointment</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Total # of skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>% Correct</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IOA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix D

Evaluation of Supervisee Performance Form
### EVALUATION OF SUPERVISEE PERFORMANCE

BCBA Supervisor: ____________________  Bin/Tx: ______________
Observer: __________________________  Student’s initials: _______
Trainer’s initials: ___________________  Duration: ____________

**Directions:** Indicate whether the skills listed below were present during the observation by placing a checkmark (✓) either on the yes or no column on each trial. Place a checkmark under the N/A column if a particular skill did not apply within the observation. Use tallies under each column when appropriate.

<table>
<thead>
<tr>
<th>Dates</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
<th>Session 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skill</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Conducts Preference Assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uses diff. ways to call attention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phrases S as instruction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Differentially R correct responding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contrives manding opportunities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intersperses skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shaping</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accurately collects data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Praise</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incorporates previous feedback</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total # of skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Correct</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IOA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix E

DTT Performance Feedback Sheet
## DTT PERFORMANCE FEEDBACK SHEET

<table>
<thead>
<tr>
<th>Trainer’s name:</th>
<th>Date:</th>
</tr>
</thead>
</table>

### CATEGORY 1 – ADDRESS FIRST

<table>
<thead>
<tr>
<th>S1.</th>
<th>Safety related</th>
</tr>
</thead>
<tbody>
<tr>
<td>S2.</td>
<td>Ethical</td>
</tr>
<tr>
<td>S3.</td>
<td>School/setting policy</td>
</tr>
<tr>
<td>S4.</td>
<td>Acquisition plan</td>
</tr>
<tr>
<td>S5.</td>
<td>Repeated errors</td>
</tr>
</tbody>
</table>

### CATEGORY 2 – ADDRESS SECOND

<table>
<thead>
<tr>
<th>S1.</th>
<th>Differential R&lt;sup&gt;+&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>S2.</td>
<td>Preference Assessment</td>
</tr>
<tr>
<td>S3.</td>
<td>S&lt;sub&gt;0&lt;/sub&gt; as question</td>
</tr>
<tr>
<td>S4.</td>
<td>Shaping</td>
</tr>
<tr>
<td>S5.</td>
<td>Previous feedback</td>
</tr>
</tbody>
</table>

### CATEGORY 3 – ADDRESS THIRD

<table>
<thead>
<tr>
<th>S1.</th>
<th>Data collection</th>
</tr>
</thead>
<tbody>
<tr>
<td>S2.</td>
<td>Gaining attention</td>
</tr>
<tr>
<td>S3.</td>
<td>Interspersing skills</td>
</tr>
<tr>
<td>S4.</td>
<td>Praise statements</td>
</tr>
<tr>
<td>S5.</td>
<td>Constriving mands</td>
</tr>
</tbody>
</table>

### CATEGORY 4 – ADDRESS FOURTH

<table>
<thead>
<tr>
<th>S1.</th>
<th>Appropriate affect</th>
</tr>
</thead>
<tbody>
<tr>
<td>S2.</td>
<td>Generalization of skills</td>
</tr>
<tr>
<td>S3.</td>
<td>Maintenance of skills</td>
</tr>
<tr>
<td>S4.</td>
<td>Problem Solving</td>
</tr>
<tr>
<td>S5.</td>
<td></td>
</tr>
</tbody>
</table>
Appendix F

Uber Supervisor Training Protocol Fidelity Checklist
## Uber Supervisor Training Protocol Fidelity Checklist

**BCBA Supervisor:** 
**Observer:** 
**Trainer:** 

**Date:** 
**Uber:** 
**Duration:**

**Directions:** Indicate whether the skills listed below were present during the observation by placing a checkmark (✓) either on the year or no column on each trial. Place a checkmark under the N/A column if a particular skill did not apply within the observation.

<table>
<thead>
<tr>
<th>Skill</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
<th>Session 5</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed BCBA engaging in previous supervision session identified &amp; documented relevant BCBA errors</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
</tr>
<tr>
<td>Waited from 5-30 secs. Prioritized corrective feedback for BCBA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prompted note taking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Waited from 5-30 secs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Provided corrective feedback to BCBA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Waited from 5-30 secs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coached &amp;/or modeled appropriate BCBA responses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Waited from 5-30 secs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Provided opportunities for BCBA to role play/practice</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Waited from 5-30 secs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Provided positive reinforcement to BCBA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prompted BCBA re: Relevant skills during supervision</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Waited from 5-30 secs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scheduled follow-up appointment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Waited from 5-30 secs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total # of skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Correct</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IOA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix G

DTT Data Collection Sheet
### DTT Data Collection Sheet – Please tally the frequency of the response

**Gaining the Client’s Attention Appropriately** – “Yes” left column, “No” right column
- Gains attention
- Timey manner
- Appropriate way
- Other (e.g., preferred item used)

**Providing the Correct S as Specified in Skill Acquisition Plan** – “Yes” left column, “No” right column
- Form
- Stimuli
- Tone
  - Clear
  - Concise
  - S’ Once
  - Cues
  - Positioning
  - Type/Variied
  - Number
  - Appropriate Volume
  - Natural Voice

**Correctly Prompting (If necessary) the Response** – “Yes” left column, “No” right column
- Type
- Timeliness
- Application
  - Hierarchy
  - Other:
  - On-time
  - Safe Handling
  - Single prompt
  - Neutral voice
  - Waits for independence

**Reacting to the Response Appropriately** – “Yes” left column, “No” right column
- Correct response
- Focused
- Correct # of responses
- Interrupted scrolling
- Reinfr. 1st Indep. trial

**Providing the Correct Consequence** – “Yes” left column, “No” right column
- Instructional Control
- Timeliness
- Application
  - Maintained schedule of R+
  - On-time
  - Time w/ R+
  - Scrolling
  - Behavior
  - Praise

**Accurately Collecting Data** – “Yes” left column, “No” right column
- Timeliness
- Application
  - On-time
  - Accurate Recording

**Removing Stimuli Following Each Trial** – “Yes” left column, “No” right column
- Timeliness
  - On-time Removal
  - Stimuli Removed
  - Added Other Stimuli
Appendix H

Supervision Social Validity Questionnaire
Supervision Social Validity Questionnaire

Please answer the following questions by indicating “yes” or “no”.

1. Did you think the intervention was helpful in improving your supervisory practice? [ ] YES [ ] NO

2. Will you continue to use the protocol in your day-to-day supervision? [ ] YES [ ] NO

3. Would you recommend this intervention to other BCBA supervisors? [ ] YES [ ] NO

4. Based on your experience with the protocol, were all the skills important to achieving and maintaining excellent therapist performance? [ ] YES [ ] NO

5. Please list any skills that do not seem important for a supervisor to demonstrate:

THANK YOU!
Appendix I

Recruitment Letter for ABA Companies
Dear staff members and parents,

Our company has agreed to serve as the research site for a study in ABA supervision. The study will be conducted by Yulema Cruz, M.S., BCBA, in fulfillment of her doctoral degree requirements. Please find information about the study below. If you are interested in participating, please contact Yulema at Yulema-cruz@hotmail.com or 305.490.5603.

Title of Research Study
Competency-Based Supervision in Applied Behavior Analysis (ABA)

Investigators
Yulema Cruz, M.S., BCBA, Doctoral Candidate at Nova Southeastern University

Purpose
The proposed study aims at providing a formal and systematic way of training and evaluating BCBAs who supervise DTT sessions. The primary participants of interest in this study are the staff members that are supervising the therapists working with the children (not the children).

Procedures
Specifically, this study will assess the use of a modified behavioral skills protocol (BST), in the training of master-level BCBAs and ABA therapists. The researcher will evaluate BCBA and therapist performance before and after receiving training on the modified BST supervision protocol. Data will be collected on the number of skills performed correctly by the BCBA and therapist.

Risks
There are no significant risks to this study. A reaction to the recording equipment (a small HD video recording device) used to collect data may occur which will be monitored closely to ensure that this does not affect the children’s treatment sessions. No identifying information (i.e., such as participants’ names) will be used at any time. The length and frequency of therapy sessions will not be affected by this study. Data will be collected during the children’s regularly scheduled skill acquisition sessions, as well as supervision sessions.

Benefits
We hope that the staff members will benefit from systematic training and evaluation methods aimed at increasing the quality of treatment, thereby further optimizing children’s treatment outcomes.
Videotaping

Supervision and DTT sessions will be video recorded. The purpose of videotaping is simply to review the implementation of the procedures, and to review data recorded by people working with the study. Videos will be collected about halfway through the children’s skill acquisition sessions for a duration of approximately five minutes, and for approximately 10 minutes during supervision sessions, once a week, every week for approximately 16 weeks.

Data Collection & Storage

All the video recordings and results will be kept confidential and secure and only the people working with the study will see the videos and data, unless required by law. Video recordings of participants will not be shown to anyone that is not directly involved with the study without expressed written consent.
Appendix J

Recruitment Email
Good afternoon,

I hope this finds you well. I received your contact information from _________. I am reaching out as I am about to start my dissertation research on Competency-Based Supervision in ABA and was wondering if you and other behavior analysts at your clinic would be interested in participating. The research entails the implementation of a brand-new protocol that Dr. Jon Bailey and I developed to train BCBAs on how to become effective supervisors. If you or any of the BCBAs on your team are interested, I would be happy to schedule a phone call to tell you more about it. I look forward to hearing back from you!

Kind regards,

Yulema Cruz, M.S., BCBA
yulema-cruz@hotmail.com

Privacy Statement: This e-mail, including any attachments, may include confidential and/or proprietary information, and may be used only by the person or entity to which it is addressed. If the reader of this e-mail is not the intended recipient or his/her authorized agent, the reader is hereby notified that any dissemination, distribution, or copying of this e-mail is prohibited. If you have received this e-mail in error, please reply to the sender and delete/shred it immediately.
Appendix K

Site Approval Letter
SITE APPROVAL LETTER

Nova Southeastern University
3301 College Avenue
Fort Lauderdale, FL 33314-7796

Subject: Site Approval Letter

To whom it may concern:

This letter acknowledges that I have received and reviewed a request by Yulema Cruz to conduct a research project entitled “Competency-based Supervision in Applied Behavior Analysis” at company name, Inc. and I approve of this research to be conducted at our facility.

When the researcher receives approval for his research project from the Nova Southeastern University’s Institutional Review Board/NSU IRB, I agree to provide access for the approved research project. If we have any concerns or need additional information, we will contact the Nova Southeastern University’s IRB at (954) 262-5369 or irb@nova.edu.

Sincerely,

Jane Doe
Executive Director
Phone: ***-****
Email

Company’s logo would go here

Company’s name, Inc.
Company’s address including suite number
City, state, zip code
Phone number
Appendix L

Supervisor General Informed Consent Form
General Informed Consent Form

NSU Consent to be in a Research Study Entitled
Competency-Based Supervision in Applied Behavior Analysis

Who is doing this research study?
College: Abraham S. Fischler College of Education
Principal Investigator: Yulema Cruz, MS, BCBA
Faculty Advisor/Dissertation Chair: Barbara Garcia-Lavin, PhD
Co-Investigator(s): Jon S. Bailey, PhD, BCBA-D, Dana S. Mills, PhD
Site Information: ____________________________________________________
Funding: Unfunded

What is this study about?
This is a research study, designed to test and create new ideas that other people can use. The purpose of this research study is to provide a systematic way of training and evaluating supervisors in Applied Behavior Analysis. Specifically, this study will assess the use of a modified behavioral skills protocol (BST) in the training of master's level Board Certified Behavior Analysts (BCBAs). Included in the supervisor-training curriculum issued by the Behavior Analyst Certification Board (BACB) in 2012, BST is an evidence-based protocol whose efficacy has been established primarily in the area of staff training. It is hypothesized that the use of this protocol for the training of BCBAs will also have a greater positive effect on supervisees' as well as clients' performances, when compared to current supervisory and clinical practices.

Why are you asking me to be in this research study?
You are being asked to be in this research study because you are a Board Certified Behavior Analyst (BCBA) supervisor, who is experienced in the implementation and oversight of discrete trials teaching (DTT) sessions.

This study will include about 20 people. It is expected that 10 people will be from this location.

What will I be doing if I agree to be in this research study?
While you are taking part in this research study, you will participate in approximately 20 sessions. This commitment will encompass 1-2 sessions per week (i.e., one supervision, and one Uber/training supervision session), every week, for approximately 16 weeks.

You may have to come back to the _________ Clinic every week for approximately 16 weeks.

Research Study Procedures - as a participant, this is what you (the BCBA) will be doing:
During baseline sessions

Baseline sessions will take place as they typically would in clinical practice, where the supervisee is engaged in teaching various skills to a child, in the presence of the BCBA. The BCBA supervisor will observe the trainer and provide feedback during the DTT session, or after the session has ended in a separate room. A research assistant and/or the researcher will collect a digital video sample of 20 minutes for each session, and code the video data focusing on the supervisor’s performance. Baseline data will be collected once a week for approximately five weeks, or until the graphic display of the data show a certain degree of stability. After these five weeks, the treatment phase will be implemented.

During Uber training sessions

These sessions will encompass the actual implementation of the independent variable (i.e., the Modified BST Protocol) and will take place between the research assistant (or the researcher) and the BCBA participants, on the same day as the regularly scheduled DTT sessions, two hours prior, in a room designated for the purposes of this training/meeting. During this time, BCBA participants will be trained on the modified BST protocol and taught how to use the Evaluation of Supervisee Performance checklist, as well as the DTT Performance Feedback Sheet form. These training sessions will last and be filmed for approximately 20 minutes each, for purposes of treatment integrity.

During modified BST sessions

These sessions will occur an hour before each scheduled DTT session in the same room as the Uber training sessions, and will include the BCBA supervisor and the supervisee. The BCBA supervisor will implement the modified BST protocol during the supervision session with the supervisee. Data will be collected on the number of modified BST protocol skills correctly performed by the BCBA. Like the Uber sessions, the modified BST supervision sessions will also have a duration of approximately 20 minutes.

Could I be removed from the study early by the research team? There are several reasons why the researchers may need to remove you from the study early. Some reasons may include:

1. Your performance may no longer meet inclusion criteria
2. Your performance demonstrates a failure to follow the study’s interventions
3. More than 2 absences are reported during the duration of the study

Are there possible risks and discomforts to me?

This research study involves minimal risk to you. To the best of our knowledge, the things you will be doing have no more risk of harm than you would have in everyday life.

- Physical risks: none
- Psychological risks: none
- Privacy risks: minimal, low magnitude, short duration (approximately 4 months). Risk will be minimized in the following manner: Videotaped sessions will only be viewed by the
researchers, researcher assistants, BCBA participants (DTT videos, their own), and trainers (their own videos).

- Breach of confidentiality: minimal, low magnitude, short duration (approximately 4 months). Risk will be minimized in the following manner: All data will be kept confidential until the completion of the study. De-identified data may be shared only with research assistants and Co-PIs. Data will be kept in the researcher’s computer, which is password protected, within a password protected file, in the researcher’s home.
- Legal risks: none
- Social risks: none
- Economic risks: none
- Group or community risks: none

**What happens if I do not want to be in this research study?**

You have the right to leave this research study at any time or refuse to be in it. If you decide to leave or you do not want to be in the study anymore, you will not get any penalty or lose any services you have a right to get. If you choose to stop being in the study before it is over, any information about you that was collected before the date you leave the study will be kept in the research records for 36 months from the end of the study and may be used as a part of the research.

**What if there is new information learned during the study that may affect my decision to remain in the study?**

If significant new information relating to the study becomes available, which may relate to whether you want to remain in this study, this information will be given to you by the investigators. You may be asked to sign a new Informed Consent Form, if the information is given to you after you have joined the study.

**Are there any benefits for taking part in this research study?**

The possible benefit of your being in this research study is improvement of your supervisory skills in the area of DTT. Additionally, other potential benefits may include increasing (your) supervisor confidence, enthusiasm, and ability to provide effective/ethical feedback. There is no guarantee or promise that you will receive any benefit from this study. We hope the information learned from this research study will benefit other people with similar conditions in the future.

**Will I be paid or be given compensation for being in the study?**

You will not be given any payments or compensation for being in this research study.

**Will it cost me anything?**

There are no costs to you for being in this research study.

However, you may be asked to come into the clinic a couple of hours prior to the scheduled DTT/supervision session.

**How will you keep my information private?**
Information we learn about you in this research study will be handled in a confidential manner, within the limits of the law and will be limited to people who have a need to review this information. All data will be kept confidential until the completion of the study. De-identified data may be shared only with research assistants and Co-PIs. These data will be available to the researcher, the Institutional Review Board and other representatives of this institution, and any regulatory and granting agencies (if applicable). If we publish the results of the study in a scientific journal or book, we will not identify you. All confidential data will be kept securely in the researcher’s computer, which is password protected, within password-protected and/or encrypted cloud files; in the researcher’s home. All data will be kept for 36 months and destroyed after that time by deleting all data contained within camera memory cards, cloud storage, or researcher’s hard drive password-protected files.

**Will there be any Audio or Video Recording?**

This research study involves video recording. This recording will be available to the researcher, the Institutional Review Board and other representatives of this institution. The recording will be kept, stored, and destroyed as stated in the section above. Because what is in the recording could be used to find out that it is you, it is not possible to be sure that the recording will always be kept confidential. The researcher will try to keep anyone not working on the research from listening to or viewing the recording.

**Whom can I contact if I have questions, concerns, comments, or complaints?**

If you have questions now, feel free to ask us. If you have more questions about the research, your research rights, or have a research-related injury, please contact:

**Primary contact:**
Yulema Cruz can be reached at 305-490-5603

If primary is not available, contact:
Jenna Kriss can be reached at 954-829-5774

**Research Participants Rights**

For questions/concerns regarding your research rights, please contact:

Institutional Review Board
Nova Southeastern University
(954) 262-5369 / Toll Free: 1-866-499-0790
IRB@nova.edu

You may also visit the NSU IRB website at [www.nova.edu/irb/information-for-research-participants](http://www.nova.edu/irb/information-for-research-participants) for further information regarding your rights as a research participant.

All space below was intentionally left blank.
Research Consent & Authorization Signature Section

Voluntary Participation - You are not required to participate in this study. In the event you do participate, you may leave this research study at any time. If you leave this research study before it is completed, there will be no penalty to you, and you will not lose any benefits to which you are entitled.

If you agree to participate in this research study, sign this section. You will be given a signed copy of this form to keep. You do not waive any of your legal rights by signing this form.

SIGN THIS FORM ONLY IF THE STATEMENTS LISTED BELOW ARE TRUE:

- You have read the above information.
- Your questions have been answered to your satisfaction about the research.

Adult Signature Section

I have voluntarily decided to take part in this research study.

Printed Name of Participant ___________________________ Signature of Participant ___________________________ Date ___________________________

Printed Name of Person Obtaining Consent and Authorization ___________________________ Signature of Person Obtaining Consent & Authorization ___________________________ Date ___________________________
Appendix M
Supervisee General Informed Consent Form
General Informed Consent Form
NSU Consent to be in a Research Study Entitled
Competency-Based Supervision in Applied Behavior Analysis

Who is doing this research study?
College: Abraham S. Fischler College of Education
Principal Investigator: Yulema Cruz, MS, BCBA
Faculty Advisor/Dissertation Chair: Barbara Garcia-Lavin, PhD
Co-Investigator(s): Jon S. Bailey, PhD, BCBA-D, Dana S. Mills, PhD
Site Information: ___________________________________________________________
Funding: Unfunded

What is this study about?
This is a research study, designed to test and create new ideas that other people can use. The purpose of this research study is to provide a systematic way of training and evaluating supervisors in Applied Behavior Analysis. Specifically, this study will assess the use of a modified behavioral skills protocol (BST) in the training of master's level Board Certified Behavior Analysts (BCBAs). Included in the supervisor-training curriculum issued by the Behavior Analyst Certification Board (BACB) in 2012, BST is an evidence-based protocol whose efficacy has been established primarily in the area of staff training. It is hypothesized that the use of this protocol for the training of BCBAs will also have a greater positive effect on supervisees’ as well as clients’ performances, when compared to current supervisory and clinical practices.

Why are you asking me to be in this research study?
You are being asked to be in this research study because you are an ABA therapist, who engages in the implementation of discrete trials teaching (DTT) sessions.

This study will include about 20 people. It is expected that 10 people will be from this location.

What will I be doing if I agree to be in this research study?
While you are taking part in this research study, you will participate in approximately 20 sessions. This commitment will encompass 1-2 sessions per week (i.e., one supervision, and one DTT session), every week, for approximately 16 weeks.

You may have to come back to the __________ Clinic every week for approximately 16 weeks.

Research Study Procedures - as a participant, this is what you (supervisee) will be doing:
During DTT sessions

Baseline and treatment DTT sessions will take place as they typically would in clinical practice, where the supervisee is engaged in teaching various skills to a child, in the presence and/or absence of the BCBA. During baseline, the BCBA supervisor will observe the trainer and provide feedback during the DTT session, or after the session has ended in a separate room. During treatment, meetings with the supervisor will take place as described below under Modified BST Sessions. A research assistant and/or the researcher will collect a digital video sample of 20 minutes for each baseline session, and 5 minutes during each treatment DTT session, and code the video data focusing on the supervisee’s performance. Baseline data will be collected once a week for approximately five weeks, or until the graphic display of the data show a certain degree of stability. After these five weeks, the treatment phase (described below) will be implemented.

During modified BST sessions

These sessions will occur an hour before each scheduled DTT session in the same room as the Uber training sessions and will include the BCBA supervisor and the supervisee. The BCBA supervisor will implement the modified BST protocol during the supervision session with the supervisee. Data will be collected on the number of modified BST protocol skills correctly performed by the BCBA. Like the Uber sessions, the modified BST supervision sessions will also have a duration of approximately 20 minutes.

Could I be removed from the study early by the research team? There are several reasons why the researchers may need to remove you from the study early. Some reasons may include:

1. You engage in unethical behavior and/or have had complaints filed against you with the Behavior Analyst Certification Board (BACB) or other regulatory entity (e.g., insurance company).
2. You are unjustifiably absent 3 or more times throughout the duration of the study.
3. You verbally express refusal to implement the feedback provided by your BCBA supervisor.
4. You are simultaneously employed by another ABA company different from the one in which the study will be conducted.

Are there possible risks and discomforts to me?

This research study involves minimal risk to you. To the best of our knowledge, the things you will be doing have no more risk of harm than you would have in everyday life.

- Privacy risks: minimal, low magnitude, short duration (approximately 4 months). Risk will be minimized in the following manner: Videotaped sessions will only be viewed by the researcher, researcher assistants, BCBA participants (DTT videos, their own), and trainers (their own videos).
- Breach of confidentiality: minimal, low magnitude, short duration (approximately 4 months). Risk will be minimized in the following manner: All data will be kept confidential until the completion of the study. De-identified data may be shared only with research assistants and Co-PIs. Data will be kept in the researcher’s computer, which is password protected, within a password protected file, in the researcher’s home.
What happens if I do not want to be in this research study?

You have the right to leave this research study at any time or refuse to be in it. If you decide to leave or you do not want to be in the study anymore, you will not get any penalty or lose any services you have a right to get. If you choose to stop being in the study before it is over, any information about you that was collected before the date you leave the study will be kept in the research records for 36 months from the end of the study and may be used as a part of the research.

What if there is new information learned during the study that may affect my decision to remain in the study?

If significant new information relating to the study becomes available, which may relate to whether you want to remain in this study, this information will be given to you by the investigators. You may be asked to sign a new Informed Consent Form, if the information is given to you after you have joined the study.

Are there any benefits for taking part in this research study?

The possible benefit of your being in this research study is improvement of your implementation of DTT, including the refinement of your skills. Additionally, participation in the study may result in potential improvements on the manner in which you conduct the DTT sessions, your affect, and your enjoyment of the sessions, which may decrease the number of sessions canceled, and increase your confidence. There is no guarantee or promise that you will receive any benefit from this study. We hope the information learned from this research study will benefit other people with similar conditions in the future.

Will I be paid or be given compensation for being in the study?

You will not be given any payments or compensation for being in this research study.

Will it cost me anything?

There are no costs to you for being in this research study.

However, you may be asked to come into the clinic an hour prior to the scheduled DTT/supervision session.

How will you keep my information private?

Information we learn about you in this research study will be handled in a confidential manner, within the limits of the law and will be limited to people who have a need to review this information. All data will be kept confidential until the completion of the study. De-identified data may be shared only with research assistants and Co-PIs. These data will be available to the researcher, the Institutional Review Board and other representatives of this institution, and any regulatory and granting agencies (if applicable). If we publish the results of the study in a scientific journal or book, we will not identify you. All confidential data will be kept securely in the researcher’s computer, which is password protected, within password-protected and/or encrypted cloud files; in the researcher’s home. All data will be kept for 36 months and
destroyed after that time by deleting all data contained within camera memory cards, cloud storage, or researcher's hard drive password-protected files.

**Will there be any Audio or Video Recording?**

This research study involves video recording. This recording will be available to the researcher, the Institutional Review Board and other representatives of this institution. The recording will be kept, stored, and destroyed as stated in the section above. Because what is in the recording could be used to find out that it is you, it is not possible to be sure that the recording will always be kept confidential. The researcher will try to keep anyone not working on the research from listening to or viewing the recording.

**Whom can I contact if I have questions, concerns, comments, or complaints?**

If you have questions now, feel free to ask us. If you have more questions about the research, your research rights, or have a research-related injury, please contact:

Primary contact:
Yulema Cruz can be reached at 305-490-5603

If primary is not available, contact:
Jenna Kriss can be reached at 954-829-5774

**Research Participants Rights**
For questions/concerns regarding your research rights, please contact:

Institutional Review Board
Nova Southeastern University
(954) 262-5369 / Toll Free: 1-866-499-0790
IRB@nova.edu

You may also visit the NSU IRB website at www.nova.edu/irb/information-for-research-participants for further information regarding your rights as a research participant.

All space below was intentionally left blank.
Research Consent & Authorization Signature Section

Voluntary Participation - You are not required to participate in this study. In the event you do participate, you may leave this research study at any time. If you leave this research study before it is completed, there will be no penalty to you, and you will not lose any benefits to which you are entitled.

If you agree to participate in this research study, sign this section. You will be given a signed copy of this form to keep. You do not waive any of your legal rights by signing this form.

SIGN THIS FORM ONLY IF THE STATEMENTS LISTED BELOW ARE TRUE:
- You have read the above information.
- Your questions have been answered to your satisfaction about the research.

Adult Signature Section

I have voluntarily decided to take part in this research study.

Printed Name of Participant                  Signature of Participant                  Date

Printed Name of Person Obtaining Consent and Authorization                  Signature of Person Obtaining Consent & Authorization                  Date
Appendix N

Parent/Guardian or Legally Authorized Representative (LAR)
Informed Consent and Adolescent Assent Form
Parent/Guardian or Legally Authorized Representative (LAR)
Informed Consent and Adolescent Assent Form
NSU Consent/Assent to be in a Research Study Entitled
Competency-Based Supervision in Applied Behavior Analysis

Who is doing this research study?

College: Abraham S. Fischler College of Education
Principal Investigator: Yulema Cruz, MS, BCBA
Faculty Advisor/Dissertation Chair: Barbara Garcia-Lavin, PhD
Co-Investigator(s): Jon S. Bailey, PhD, BCBA-D, Dana S. Mills, PhD

Site Information: please insert info here
Funding: Unfunded

What is this study about?

This is a research study, designed to test and create new ideas that other people can use. The purpose of this research study is to provide a systematic way of training and evaluating supervisors in Applied Behavior Analysis. Specifically, this study will assess the use of a modified behavioral skills protocol (BST) in the training of master's level Board Certified Behavior Analysts (BCBAs). Included in the supervisor-training curriculum issued by the Behavior Analyst Certification Board (BACB) in 2012, BST is an evidence-based protocol whose efficacy has been established primarily in the area of staff training. It is hypothesized that the use of this protocol for the training of BCBAs will also have a greater positive effect on supervisees' as well as clients' performances, when compared to current supervisory and clinical practices.

Why are you asking me to be in this research study?

We are not asking you to participate in this study. However, we are requesting your consent for your child to participate in this research study, because you are the parent of a child who receives ABA therapy and participates in discrete trials teaching (DTT) sessions.

This study will include about 20 people. It is expected that 10 people will be from this location.

What will I be doing if I agree to be in this research study?

You will not be a study participant. You would be allowing us to review your child's data to ensure that your child’s age is between two and 10, that he/she has a diagnosis of an intellectual disability, participates in DTT for at least six hours a week, and does not engage in self-injury, aggression, property destruction, elopement, or excessive crying more than three times per week. However, no data will be collected on your child’s behaviors. In addition, you
would be allowing us to videotape your child during his/her regularly scheduled DTT sessions (described below).

**During DTT sessions**

Baseline and treatment DTT sessions will take place as they typically would in clinical practice, where the supervisee is engaged in teaching various skills to a child, in the presence and/or absence of the BCBA. During baseline, the BCBA supervisor will observe the trainer and provide feedback during the DTT session, or after the session has ended in a separate room. During treatment, the supervisee will meet alone with the child. A research assistant and/or the researcher will collect a digital video sample of 20 minutes for each baseline session, and 5 minutes during each treatment DTT session, and code the video data focusing on the supervisee’s performance. Baseline data will be collected once a week for approximately five weeks, or until the graphic display of the data show a certain degree of stability. After these five weeks, the treatment phase will be implemented.

**Could I be removed from the study early by the research team?** There are several reasons why the researchers may need to remove your child from the study early. Some reasons are:

1. Your child is unjustifiably absent 3 or more times throughout the duration of the study.
2. Your child’s frequency/duration/intensity of challenging behaviors increases during the study.

**Are there possible risks and discomforts to me?**

This research study involves minimal risk to your child. To the best of our knowledge, the things your child will be doing have no more risk of harm than you would have in everyday life.

- Privacy risks: minimal, low magnitude, short duration (approximately 4 months). Risk will be minimized in the following manner: Videotaped sessions will only be viewed by the researcher, researcher assistants, BCBA participants (DTT videos, their own), and trainers (their own videos).
- Breach of confidentiality: minimal, low magnitude, short duration (approximately 4 months). Risk will be minimized in the following manner: All data will be kept confidential until the completion of the study. De-identified data may be shared only with research assistants and Co-PIs. Data will be kept in the researcher’s computer, which is password protected, within a password protected file, in the researcher’s home.

**What happens if I do not want to be in this research study?**

You have the right to have your child leave this research study at any time or refuse to let your child be in it. If you decide it is time for our child to leave or you do not want your child to be in the study anymore, you and your child will not get any penalty or lose any services you have a right to get. If you choose for your child to stop being in the study before it is over, any information about your child that was collected before the date you leave the study will be kept in the research records for 36 months from the end of the study and may be used as a part of the research.
What if there is new information learned during the study that may affect my decision to remain in the study?

If significant new information relating to the study becomes available, which may relate to whether you want to remain in this study, this information will be given to you by the investigators. You may be asked to sign a new Informed Consent Form, if the information is given to you after you have joined the study.

Are there any benefits for taking part in this research study?

The possible benefit of your child being in this research study is improvement of his/her skill acquisition rates during DTT. This may include faster skill acquisition rates, as well as potential improvements in his/her attendance, affect, active participation in therapy, and compliance. There is no guarantee or promise that you will receive any benefit from this study. We hope the information learned from this research study will benefit other people with similar conditions in the future.

Will I be paid or be given compensation for being in the study?

You will not be given any payments or compensation for being in this research study.

Will it cost me anything?

There are no costs to you for being in this research study.

How will you keep my information private?

Information we learn about your child in this research study will be handled in a confidential manner, within the limits of the law and will be limited to people who have a need to review this information. All data will be kept confidential until the completion of the study. De-identified data may be shared only with research assistants and Co-PIs. These data will be available to the researcher, the Institutional Review Board and other representatives of this institution, and any regulatory and granting agencies (if applicable). If we publish the results of the study in a scientific journal or book, we will not identify your child. All confidential data will be kept securely in the researcher’s computer, which is password protected, within password-protected and/or encrypted cloud files; in the researcher’s home. All data will be kept for 36 months and destroyed after that time by deleting all data contained within camera memory cards, cloud storage, or researcher’s hard drive password-protected files.

Will there be any Audio or Video Recording?

This research study involves video recording. This recording will be available to the researcher, the Institutional Review Board and other representatives of this institution. The recording will be kept, stored, and destroyed as stated in the section above. Because what is in the recording could be used to find out that it is you, it is not possible to be sure that the recording will always be kept confidential. The researcher will try to keep anyone not working on the research from listening to or viewing the recording.
What Student/Academic Information will be collected and how will it be used?

The following information will be collected from your child’s program book and portal data: client, clinical information, and other relevant documents. These records will be used for inclusion criteria to ensure that your child’s age is between two and 10, that he/she has a diagnosis of an intellectual disability, participates in DTT for at least six hours a week, and does not engage in self-injury, aggression, property destruction, elopement, or excessive crying more than three times per week. These records will be given to the Principal Investigator by the supervising BCBA, the therapist, and the case-management department. Please note that no data will be collected on your child’s behaviors during the study.

Whom can I contact if I have questions, concerns, comments, or complaints?

If you have questions now, feel free to ask us. If you have more questions about the research, your research rights, or have a research-related injury, please contact:

Primary contact:
Yulema Cruz can be reached at 305-490-5603

If primary is not available, contact:
Please insert your info (Hana’s info) here.

Research Participants Rights
For questions/concerns regarding your research rights, please contact:

Institutional Review Board
Nova Southeastern University
(954) 262-5369 / Toll Free: 1-866-499-0790
IRB@nova.edu

You may also visit the NSU IRB website at www.nova.edu/irb/information-for-research-participants for further information regarding your rights as a research participant.

All space below was intentionally left blank.
**Research Consent & Authorization Signature Section**

Voluntary Participation - You are not required to participate in this study. In the event you do participate, you may leave this research study at any time. If you leave this research study before it is completed, there will be no penalty to you, and you will not lose any benefits to which you are entitled.

If you agree to participate in this research study, sign this section. You will be given a signed copy of this form to keep. You do not waive any of your legal rights by signing this form.

**SIGN THIS FORM ONLY IF THE STATEMENTS LISTED BELOW ARE TRUE:**
- You have read the above information.
- Your questions have been answered to your satisfaction about the research.

**Parental/Guardian or Legally Authorized Representative (LAR) Signature Section**

I am voluntarily giving my consent for another person to participate in this study because I believe this person would want to take part if able to make the decision and I believe it is in this person’s best interest.

*Person giving Consent must select whether they are a Parent/Guardian or a LAR

<table>
<thead>
<tr>
<th>Printed Name of Participant</th>
<th>Signature of Participant, indicating Assent for Adults and Children over the age of 13 (Children under the age of 13 must sign the Child Assent Form)</th>
<th>Date</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Printed Name of Person Giving Consent &amp; Authorization for Participant</th>
<th>Signature of Person Giving Consent &amp; Authorization*</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Parent/Guardian</em></td>
<td>☐Parent/Guardian ☐LAR</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Printed Name of Person Obtaining Consent and Authorization</th>
<th>Signature of Person Obtaining Consent &amp; Authorization</th>
<th>Date</th>
</tr>
</thead>
</table>
Appendix O

Pilot Study Summary
COMPETENCY-BASED SUPERVISION IN ABA

During the months of February through April of 2015, a single-case AB design pilot study was conducted to evaluate the effectiveness of the modified BST protocol on improving the supervisory behaviors of one BCBA, who had at least one year of experience supervising trainers conducting DTT sessions with children with intellectual disabilities. The researcher recruited three participants from a local ABA company located in Broward County, Florida. One of the participants was a BCBA. The second participant was a DTT trainer with less than six months of experience in the implementation of DTT with children diagnosed with Autism Spectrum Disorder (ASD). Both trainer and BCBA had undergone agency-specific training on DTT, and met after initial assessments were conducted, inclusion criteria for participation in the study. Additionally, supervisor and supervisee had signed consents for participation in research provided by the local ABA company at which they were both employed. The third participant in the study was a six-year-old child diagnosed with ASD receiving ABA therapy as part of early intervention programs since the age of three. Consent for the child’s participation in the study was obtained from his mother by the agency administration, the BCBA and the researcher. His mother observed all sessions in a separate room via a monitor, which fed live video from the child’s therapy room. After initial reviews of data were conducted, the client also met inclusion criteria for the study, including: a formal diagnosis of an intellectual disability; a current age between two and 10 years old; demonstrated prior success with the acquisition of new skills via the implementation of DTT for an average of at least 10 hours a week, every week, for a period of four months of longer; and engaged in the following behaviors at a frequency of two times per week or less: self-injurious behavior, aggression, property destruction, elopement, or excessive crying (lasting for a period of 10 minutes or more) during each DTT session. Lastly, at the time
of the pilot, the client was an active consumer of ABA services at the aforementioned local ABA company, working with the first and second participants referenced above.

The study was conducted at a local ABA company’s clinic that was equipped with therapy rooms and materials, as well as computers, cameras, speakers, monitors, and other necessary resources. DTT sessions were videotaped during the client’s regularly scheduled therapy sessions at the clinic. During baseline, supervision sessions between the BCBA and the trainer were conducted in the free-play room at the clinic, during the client’s recess time approximately mid-way through the DTT sessions, or after the sessions had concluded. In contrast, supervision sessions during treatment took place in a separate office at the clinic, one hour before the DTT session between the child and the therapist was scheduled to begin. All DTT sessions were videotaped at approximately halfway through the session, for about five minutes. Supervision sessions typically lasted for about 20 minutes, which also corresponded to the duration of the films. Sessions were filmed using a digital video camera with a built-in microphone, and sometimes with an iPad. The researcher videotaped all the sessions for the purposes of data collection, coding/analysis of the data, and inter-observer agreement (reliability) measures.

The independent variable in the study was the modified BST protocol that included a checklist for training and evaluating the supervisory behaviors of the BCBA titled *Evaluation of Supervisor Performance* (see Appendix C for a copy of this document). A second checklist for evaluating the behaviors of the trainer during DTT sessions named *Evaluation of Supervisee Performance* (see Appendix D for a copy of this document) was used for collateral data collection. Additionally, a job aid for the supervisor to use when providing feedback to the supervisee during supervision sessions termed *DTT Performance Feedback Sheet* (see Appendix
Baseline data on current supervisory practices were collected for five weeks from February second through March 16, 2015. During this phase, the researcher first videotaped the DTT sessions between the child and the therapist (as described above), and supervision sessions were filmed either sometime after that, or at the end of the DTT session. All data were collected and analyzed by the researcher along with research assistants who conducted reliability measures. For the completion of reliability checks during this phase, the researcher chose three individuals (research assistants) who had never viewed the videos before, to code all five videos using the two checklists referenced above under dependent measures (i.e., the Evaluation of Supervisor Performance; see Appendix C for a copy of this instrument, and Evaluation of Supervisee Performance; see Appendix D for a copy of this document, checklists). Aside from the checklists, the research assistants were each provided with documents that included definitions for each of the categories/behaviors contained within each of the checklists, for referencing purposes while coding. Moreover, each research assistant was trained on the use of the data collection checklists prior to the beginning of reliability checks, and when necessary, additional training sessions were conducted.

Treatment data were also collected for five weeks from March 30 through April 28, 2015. During this phase, the researcher first met with the BCBA, two hours before the regularly scheduled DTT sessions between the child and the therapist were due to begin, for a total of two times. The purpose of these sessions was to train the BCBA supervisor on the use of the new modified BST protocol that she was asked to use with the trainer (supervisee) during the
COMPETENCY-BASED SUPERVISION IN ABA

subsequent supervision sessions between herself (the BCBA) and the trainer. Training sessions lasted for approximately 20 minutes and were filmed for purposes of treatment integrity. All through the training sessions, the researcher taught the BCBA how to use this protocol while adhering to and using the modified BST protocol herself. In addition, the researcher provided the BCBA with a copy of the job aid (i.e., the *DTT Performance Feedback Sheet*; see Appendix D for a copy of this document). Furthermore, the researcher also discussed with the BCBA her supervisory performance on the latest supervision session video between the BCBA and the supervisee. For the purposes of this study, training sessions between the researcher and the BCBA supervisor were conceived as “uber supervision sessions”. While a definition of *Uber Supervision* has not been found in the ABA literature, this refers to the supervision session between a BCBA who has extensive experience in the area of supervision or an *Uber Supervisor*: a superlative example of its kind (Merriam Webster, 2016), and a BCBA with little or no experience in the area of supervision.

An hour before the DTT sessions began, the BCBA met with the supervisee for supervision. These sessions also lasted and were videotaped for approximately 20 minutes. Throughout these supervision sessions, the BCBA supervisor implemented the modified BST protocol, along with the job aid, and a discussion of the supervisee’s latest DTT session video. Lastly, DTT sessions took place as they normally would at their scheduled time, with the exception of the supervisee being asked to implement the feedback addressed during supervision.

After analyzing the data from the *Evaluation of Supervisor Performance* and the *Evaluation of Supervisee Performance* forms, results of the study revealed correct performance scores for the supervisor’s behavior during baseline equal to 50 percent, for all five sessions. For the supervisee, correct performance scores ranged from 10 to 20 percent. Following the
implementation of the modified BST protocol for the BCBA supervisor, her corrected
performance scores increased to 90 and then 100 percent for four consecutive sessions.
Similarly, after treatment implementation trainer scores immediately increased to 40 percent
during the first two sessions and reached a maximum of 70 and 80 percent during the last three
sessions. Figures one and two below show a graphical representation of the results.

Inter-observer agreement (IOA) was calculated for all sessions. Three research assistants
independently coded all the videos and submitted their results to the researcher who then used
the total number agreements divided by agreements plus disagreements times one hundred, as the
formula to calculate reliability. Baseline IOA scores ranged from 90 to 100 percent for the
supervisor’s performance scored and remained at a steady 100 percent for the supervisee’s
performance scores. Treatment IOA scores ranged from 80 to 100 percent for the data pertaining
to the supervisor’s correct performance, and from 90 to 100 percent for the data pertaining to the
supervisee’s correct performance.

As previously mentioned, client progress data were also collected and analyzed. After
close examinations of the data were conducted, the researcher developed a set of criteria that
were met in order to increase the regularity of the data that was collected for examination. The
first criterion entailed controlling for skill selection during the implementation of DTT, so that
the same skills were taught during baseline and treatment sessions. The second criterion involved
the absence of another therapist working with the same client, on the same skill acquisition
programs. In other words, children participating in this study only worked with one therapist for
the duration of the study. The third criterion concerned the introduction of new skills during the
baseline phase. Accordingly, no new skills were introduced during the treatment phase. Because
client performance data collected during the pilot contained too many irregularities that rendered
it unusable, these criteria were developed as a means to prevent abnormalities within the data that could be collaterally collected in a future study.

As an added measure, the researcher informally interviewed the BCBA and supervisee participants regarding their satisfaction with the study. Results of these informal interviews revealed that both participants self-reported significant skills’ improvement in their clinical practice. The supervisee also reported that what she liked about this new supervision model was mostly the video, the pace of the supervision, and the opportunities to role-play the skill, which embedded opportunities for positive reinforcement.

Results and data analysis of the pilot study are presented below in Figures 1 and 2.

![Figure 1. Supervisor’s scores during the baseline and treatment phases of the pilot study](image-url)
Figure 2. Supervisee’s scores during the baseline and treatment phases of the pilot study.