Winter 12-13-2019

Equivalence-Based Instruction to Teach Single-Subject Designs in Higher Education

Griffin DeCuir Williams
Pittsburg State University, gdwilliams@gus.pittstate.edu

EQUIVALENCE-BASED INSTRUCTION TO TEACH SINGLE-SUBJECT DESIGNS
IN HIGHER EDUCATION

A Thesis Submitted to the Graduate School
in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Psychology

Griffin DeCuir Williams

Pittsburg State University

Pittsburg, Kansas

December, 2019

EQUIVALENCE-BASED INSTRUCTION TO TEACH SINGLE-SUBJECT DESIGNS IN HIGHER EDUCATION

Griffin DeCuir Williams

APPROVED:

Thesis Advisor ______________________________
Dr. Ryan Speelman, Psychology and Counseling

Committee Member ______________________________
Dr. Julie Allison, Psychology and Counseling

Committee Member ______________________________
Dr. Jamie Wood, Psychology and Counseling

Committee Member ______________________________
Dr. Brian Sims, Teaching and Leadership

It is my genuine gratefulness and warmest regard that I dedicate this thesis to Dr. Murray Sidman whose early work made the current study possible.

Dr. Murray Sidman

(1923 – 2019)

EQUIVALENCE-BASED INSTRUCTION TO TEACH SINGLE-SUBJECT DESIGNS IN HIGHER EDUCATION

An Abstract of the Thesis by
Griffin DeCuir Williams

With the increasing prevalence of online courses and instruction, advanced methods to teach technical concepts for students in higher education are of value. Equivalence-Based Instruction (EBI) is an effective, efficient, and empirically validated teaching methodology. This study developed a match-to-sample EBI protocol embedded in CANVAS to teach four common single-subject designs to graduate students. Results were compared between an equivalence group receiving a reduced-intensity EBI protocol and a group receiving a traditional video lecture to serve as the control. The ability of the participants to generalize the trained relations to novel stimuli was also evaluated. Results indicate that the EBI procedure implemented by the current authors significantly increased pretest-to-posttest scores and allowed for generalization to novel exemplars. EBI streamlines the teaching of intricate concepts, may be shared across disciplines, and may allow students to gain a minimum competency prior to attending a lecture.

Keywords: stimulus equivalence, EBI, higher education, distance learning

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

Introduction

Applied behavior analysis (ABA) is a scientific methodology whose dimensions were first defined by Baer, Wolf, and Risley (1968). The core dimension of ABA has involved a strong emphasis on objective measurement of observable behavior with the goal of applying basic research principals to improve all socially significant behavior. All treatment methods are built from empirically validated principles of behavior and are completely described in such detail that a casually trained implementor may replicate the results. In addition, target behaviors of interest must show meaningful improvement for an extended period of time, in additional contexts, or influence functionally similar behavior (Baer, Wolf, & Risley, 1968). These dimensions continue to be a guiding force in the development of effective, efficient, and empirically validated behavioral interventions. Behavior analytic procedures have the potential to improve instructional methods that promote the learning of young people, who "are by far the most important natural resource of a nation, and the development of that resource is assigned to education" (Skinner, 1984, p. 953).

Behaviorally speaking, an educator is a professional who arranges contingencies to best promote student learning and overcome challenges facing the learner. First it is useful to conceptualize the academic learning environment in common behavioral

terminology (i.e., antecedents, consequences, and motivational variables). Antecedents are stimuli which signal the differential availability of rewards or punishers that impact behavior (Cooper, Heron, & Heward, 2014). For example, the last week of an academic semester signals final examinations. Students' fearful of failing may engage in studying behavior to avoid the potential threat of failure. Instructional antecedents include lectures, response prompts, supplemental readings, and syllabi which outline a plethora of response expectations and associated consequences. A consequence is a stimulus change that occurs following an emitted behavior that increases (i.e., reinforcing consequence) or decreases (i.e., punishing consequence) the future occurrences of that behavior (Cooper, Heron, & Heward, 2014). For example, a student who studied dutifully for a final exam received a passing letter grade. This rewarding consequence will likely increase the future occurrences of the studying behavior. Instructional consequences include corrective feedback, letter grades, praise, and ultimately a degree/career. Educators can contrive learner motivation by highlighting intrinsic motivators (e.g., desire to gain knowledge, excelling in a career/program, personal growth/development, etc.) and external motivators (e.g., appeasing professors/peers/staff, attaining desired employment, monetary gains, letter grades, etc.) for individual learners. Given this framework, behavior analytic treatment packages that include proper examination of antecedents and consequences, combined with an evaluation of individual motivating variables, may increase instruction effectiveness.

One such treatment package is Personalized System of Instruction (PSI) designed by Keller (1968). PSI fosters direct contingencies that promote mastery learning by establishing antecedent control (i.e., unit mastery requirement and use of

lectures/demonstrations for motivational purposes), a focus on written communication, and immediate consequences (i.e., self-paced unit mastery tracking). Unlike traditional teaching methods, PSI allows student behavior to directly control advancement through academic tasks. Another behavior analytic methodology, Precision Teaching, provides a framework for educators to evaluate the effectiveness of an instructional method (Austin & Carr, 2000). This goal is achieved by emphasizing the measurement of directly observable behavior, as well as frequency rates as the measure of performance (i.e., the average number of total responses during each interval of an assessment period) charted using standard celebration (Lindsley, 1990, 1992).

Additional interventions based in ABA have included; guided notetaking procedure to maximize student learning (Barbetta & Skaruppa, 1995; Neef, McCord, & Ferreri, 2006), Situational Awareness Training to decrease nervous habits such as use of "like" and tongue clicks during public speaking (Spieler & Miltenberger, 2017), intermittent contingent deadlines with penalties to reduce procrastination (Dillon, Kent, & Malott, 1980; Johnson, Perrin, Salo, Deschaine, & Johnson, 2016), as well as Behavior Skills Training (i.e., instructions, modeling, rehearsal with feedback) to improve common interview skills such as smiling and sitting up straight (Stocco, Thompson, Hart, & Soriano, 2017). These studies demonstrate the efficacy of the ABA methodology in addressing challenges students experience in higher education. A common theme is frequent measurement and consistent evaluation of educational practices in conjunction with the guiding principle of the learner knows best; meaning instructional packages acquire value only through demonstrated student efficacy.

CHAPTER II

Literature Review

While the aforementioned instructional procedures have proven to be effective, they rely on direct contingencies to promote learning. Another instructional paradigm rooted in behavior analysis exists where not all related functions or stimuli require direct training. The concept of *stimulus equivalence* has an extensive history in behavioral psychology and may prove to be an effective, efficient tool in the academic setting. Hull (1939) defined stimulus equivalence as a phenomenon in which a stimulus induces a reaction without prior conditioning and concluded that a relationship between a stimulus and other topographically dissimilar stimuli must exist to account for the novel responding. Later, Sidman (1971) observed a disabled youth demonstrate the ability to match animal pictures (A) to their spoken names (B) and verbally name (C) the pictures (i.e., B = A and A = C). Next Sidman taught the youth to match written words (D) with the spoken words (i.e., trained D – B). Following this training, the individual was then observed to match the written words to a picture of each and vice versa (i.e., D – A and A – D), as well as orally state the written words (i.e., D – C) without these relationships being directly trained. This finding, outlined in Appendix A, formally demonstrated Hull's observed phenomenon and the paired relationship between functionally related, topographically dissimilar stimuli was defined as a *stimulus equivalence class*.

Foundations of Stimulus Equivalence

Match-to-Sample

The formation and testing of stimulus equivalence classes is often achieved through a match-to-sample (MTS) procedure (Cooper, Heron, & Heward, 2014). MTS training involves presenting a sample stimulus (e.g., picture of a cat) and corresponding comparison stimuli with distractors (e.g., written cat, dog, bat). The learner is reinforced for correctly matching the sample stimulus with its related comparison stimulus and rejecting the nonmatching stimuli. The trainer may pair the sample stimulus with an additional comparison stimulus (e.g., spoken word cat). To test the emergence of an equivalence class, the trainer would then present the written word cat and evaluate if the learner was able to match with the spoken word cat and vice versa. MTS permits each individual relation within a stimulus equivalence class to be examined by measuring the selection-based responses of the learner.

Early Findings

Sidman and Cresson (1973) replicated the results of Sidman (1971) by using MTS to directly training two youths diagnosed with Down's syndrome to match pictures (B) and printed (C) words to the verbally dictated (A) words (i.e., A = B and A = C). Once these two stimulus relations were established, the participants were able to correctly match the picture and printed words together (i.e., B – C and C – B) without direct training. Sidman and Tailby (1982) used mathematical terminology to label this finding; if A = B and B = C, then A = C. Sidman (1994) recapitulated a key aspect of protocols based in stimulus equivalence; direct training on certain stimuli relationships will result in further untrained relationships emerging, coined *derived stimulus relations*. This

research illustrates how complex novel responses can be attributed to environmental variables that can be experimentally manipulated. By focusing solely on the learning history and measurable behavior of the individual rather than hypothetical constructs, Sidman and colleagues were able to identify three testable conditions for determining if a stimulus equivalence class has formed.

General Background

General equivalence training requires first that the learner formed relations between directly trained stimuli. For instance, a learner is directly trained to verbally respond “woof” (B) and “dog” (C) when presented with a picture (A) of a dog. A direct test would involve evaluating the presence of directly trained stimulus relations by presenting a picture of the dog and having the learner verbally respond with “woof” and then “dog” (i.e., A – B and A – C). Second is a symmetrical test over the directly trained stimuli relations; meaning the learner is presented with a verbal “woof” then “dog” and the ability to match to the picture of the dog (i.e., B – A and C – A) is evaluated. The third and most important test is of transitive/equivalence (i.e., derived) relations; meaning a relationship emerged between the stimuli not presented together in direct training. For the current example, transitivity would be the learner’s ability to verbally respond with “dog” when hearing “woof” and “woof” when hearing “dog” (i.e., B – C and C – B). The learner’s correct responding on direct, symmetry, and transitivity tests demonstrate the formation of a stimulus equivalence class.

Sidman (1994, 2000) reported that a separate process, *stimulus generalization*, promoted class expansion and categorization. Stimulus generalization occurs when prior learning allows an individual to respond to topographically similar, novel stimuli without

direct training. Topographically similar stimuli include those sharing common sensory (e.g., visual, auditory, olfactory, etc.) features. The ability to categorize (i.e., discriminate within classes as well as between classes) is tested to determine whether stimulus generalization has occurred. Using the previous example, a generalization test may involve evaluating if the learner was able to respond with a verbal "woof" and "dog" when shown a novel picture of a dog from a different breed but not when presented a picture of a different mammal (e.g., cat, horse, tiger, goat, etc.). This would demonstrate the ability of individual equivalence class members to encompass novel, relevant stimuli and discriminate between other nonrelated stimuli. Generalization remains an imperative goal of effective and efficient educational instruction.

Respondent Conditioning

While behavioral phenomena are often researched in isolation, there are many studies demonstrating how they interact. For instance, respondent/classical conditioning refers to a learning procedure in which a primary reinforcer (e.g., food, water, and safety) or primary punisher (e.g., pain and extreme temperatures) is paired with a neutral stimulus. Frequent pairings transfer the function of the primary reinforcer/punisher to the neutral stimulus, resulting in a conditioned stimulus (Cooper, Heron, & Heward, 2014). Watson and Rayner (1920) demonstrated respondent conditioning of fear in a landmark study commonly referred to as "The Little Albert Experiment". Albert, an 11-month-old infant, was conditioned to fear a fluffy, white rat (i.e., neutral stimulus). This was accomplished by repeatedly presenting Albert with the rat while simultaneously striking a metal bar with a hammer, producing a very loud noise (i.e., primary punisher). After repeated pairings, the rat became a conditioned punisher; meaning the fear of loud noises

was paired with the rat and fear was elicited in the presence of the rat alone. Further, conditioned fear generalized to topographically related stimuli (e.g., white rabbit, white mask, and fluffy white cotton) through a second process of stimulus generalization described above. Watson's study demonstrated that emotional responses can be conditioned in humans and elicited fear can be brought under experimental control.

Potential clinical applications of relational responding as an underlying process that could cause suffering and interact with elicited fear was conducted by Dougher, Augustson, Markham, Greenway, and Wulfert (1994). This study demonstrated transformation of stimulus functions, namely fear of shock, via relational responding. First two stimulus equivalence classes, each consisting of four arbitrary symbols to avoid prior stimuli pairing with a desired or aversive consequence, were trained using a MTS procedure. Next a single class member from one equivalence class was repeatedly paired via respondent conditioning with a painful electrical shock. When all class members were again presented, the fear of shock had transferred (i.e., derived) to the other three members without their direct pairing with the shock. None of the stimuli members from the second equivalence class received a shock and transfer of fear between the two equivalence classes was not observed.

In a second experiment, Dougher et al. (1994) replicated the first experiment's procedure. Next the specific class member that had been paired with the shock was put on extinction; meaning it was presented repeatedly with no electric shock in order to eliminate the conditioned fear. Once the fear response was removed via extinction from the stimulus, all equivalence class members were re-presented. The previously derived fear was no longer observed in the other three class members, demonstrating that

relational responses may be unconditioned using an extinction procedure. In addition to affirming the observed emergence of a stimulus equivalence class based on functionally equivalent stimuli as Sidman previously described, this study demonstrated that language-based inferences and subjective experiences paired with a single stimulus can be transferred to all other equivalence class members. For example, if a child is bitten by a dog, the spoken and written word "dog" may come elicit a fear response. If the child then learns that "perro" is equivalent to "dog", both the spoken and written word "perro" may come to derive the fear response that "dog" elicited. Basic findings such as these and additional applied studies highlight the ubiquity relational responding, though a full discussion of related clinical applications is beyond the scope of the current study (see Hayes, Barnes-Holmes, & Roche, 2001; Hayes, Strosahl, & Wilson, 2012 for a comprehensive review). Despite this apparent "dark side" of relational responding (Törneke, 2010), many studies have capitalized on this unique ability to promote learning across a variety of settings, including college campuses.

Stimulus Equivalence in Higher Education

A meta-analysis of 28 Equivalence-Based Instruction (EBI) studies in higher education was conducted by Brodsky and Fienup (2018) who found that EBI produced similar learning outcomes when compared to traditional instruction methods such as lectures. However, EBI was found to be more efficient at creating new repertoires in certain circumstances. Few studies included in the meta-analysis implemented EBI outside of highly controlled laboratory settings with smaller populations (e.g., Greville, Dymond, & Newton, 2016; O'Neill, Rehfeldt, Ninness, Muñoz, & Mellor, 2015; Pytte & Fienup, 2012; Varelas & Fields, 2017; Walker & Rehfeldt, 2012). Brodsky and Fienup

(2018) concluded that future research should investigate EBI in large-scale, naturalistic college settings and evaluate the effectiveness of varying dissemination methods of EBI procedures.

In larger scale applications, Fields et al. (2009) examined the ability of 21 undergraduate students to form a four member equivalence class (i.e., name, definition, graph, and textual description) over common statistical interactions. The authors found that EBI resulted in a 35% average increase in outcome between pretest and posttest scores. Further Walker, Rehfeldt, and Ninness (2010) investigated the efficacy of selection-based vocal EBI intraverbal training to develop equivalence classes (i.e., name, definition, primary cause, and treatment) for 12 common disabilities (e.g., cerebral palsy, cystic fibrosis, schizophrenia, etc.). The authors found it was effective at teaching vocal intraverbal relations and those relations may generalize to additional verbal and written intraverbal response conditions. Despite low scores on final written test probe, the authors hypothesized that this issue was due to poor relation maintenance rather than an issue generalizing across interverbal conditions.

To address this prediction, Alter and Borrero (2015) replicated the study with 17 undergraduate students and implemented two dual-section written intraverbal posttests, the first immediately after the EBI training and a second maintenance test approximately 62 days later. The results of the posttest directly after training demonstrated the effectiveness of the procedure (mean: Part 1 = 99% ; Part 2 = 79.4%). In the maintenance test, the scores were significantly lower (mean: Part 1 = 49% ; Part 2 = 39.5%). These results supported the Walker et al. (2010) prediction that a potential weakness of EBI may involve the ability to generalize the acquired skill relations across

time. Nevertheless, generalization and maintenance of skills over time is a common challenge in academic settings. In addition to the issue of relation maintenance, EBI must be systematically compared to other instructional protocols used in higher education to determine its comparative efficacy.

Fienup and Critchfield (2011) examined the learning outcome of conditional discriminations to teach three lessons (i.e., statistical significance, hypothesis decisions, and directions of future research) to 59 undergraduate students. The participants were divided into three groups; a stimulus equivalence group receiving EBI, a complete instruction group directly taught all target relations, and a control group who received no instruction. As expected, the authors found that both the stimulus equivalence and complete instruction groups had significantly higher learning outcomes than the control group. A significant difference between learning outcomes was not observed, however the stimulus equivalence group required significantly less training to reach the mastery criteria. This was the first time greater efficiency of EBI compared to other instruction procedures was demonstrated in higher education. Zinn, Newland, and Ritchie (2015) compared the learning of 32 pairs of proprietary and generic drug name relations with four stimuli each (i.e., the spoken generic name, the written generic name, the spoken proprietary name, and the written proprietary name) between an EBI group and two control groups (i.e., criterion-control and trial-control). The study found that not only was the EBI procedure more efficient at teaching the relations when compared to the control groups, it was also found to be more effective. This finding was significant because a majority of prior studies reported similar learning outcomes between learners receiving EBI and those receiving traditional instruction methods. With relation

maintenance and efficacy compared to traditional instruction considered, the varying implementation methods of EBI itself must be examined.

To do this, Fienup, Mylan, Brodsky, and Pytte (2016) conducted a study in a Behavioral Neuroscience course to evaluate if the presentation order of meaningful stimuli (i.e., picture of the structure, the name of the structure, and the function of the structure) had an effect on the learning outcomes, as Arntzen (2004) found with nonsense stimuli. Additionally, the authors evaluated whether students who voluntarily received EBI had higher classroom examination performance than control students who chose not to receive EBI. No significant difference between learning outcomes and the presentation order of meaningful stimuli was observed for the EBI group. A second finding was that EBI participants scored slightly higher overall and significantly higher on equivalence-based questions compared to the control group. A notable result also observed by the authors was that the time to complete the training was reduced with each repeated exposer to EBI trials. This indicated that the efficiency of EBI will continually increase as learners become more familiar with the procedure's format.

Expanding on these methods, Fienup & Brodsky (2017) evaluated 57 undergraduate students' mastery of neuroanatomy experimental stimuli (i.e., name, picture, function, and description of the result of damage) for the amygdala, cingulate cortex, hippocampus, and mammillary body brain structures. The authors examined if different mastery criterions (i.e., block and rolling) established during EBI training impacted learning outcomes. Participants were divided into three mastery criterion groups (i.e., rolling criteria of 6 or 12 consecutively correct responses and a block criteria of 12 trials per block with 100% score). Results showed no significant difference in

learning outcomes between 12 rolling and 12 block mastery criterion groups. However the authors stated these two more stringent mastery criterion groups demonstrated more efficient learning with less errors, failures, and retraining when compared to the more lenient 6 rolling mastery criteria group. The study indicates that while mastery criteria method does not significantly impact learning outcome, a more stringent mastery criteria will result in greater learning efficiency. In addition to evaluating the possible implementation methods of EBI, the rapidly expanding use of online instruction in higher education requires further consideration.

In the past decade, enrollment in online higher education has substantially increased (i.e., 10% annual growth rate) compared to a less than 1% increase in traditional higher education, with 31% of students taking at least one online course (Allen & Seaman, 2011). Determining how to effectively teach material, maintain contact, and provide feedback to online students present new challenges for educators. In particular, educators in higher education with limited resources must have the ability to present complex and technical curriculum in an effective, efficient manner. EBI procedures may have the potential to overcome these challenges by promoting learning without direct contact nor immediate feedback from the instructor. Supplementing EBI procedures into a course may provide a solid foundational baseline of the material which will grant the educator more time to discuss the subject matter in-depth and promote the development of other useful skills. Additionally online Learning Management Systems (e.g., Blackboard, CANVAS, etc.) adopted by universities allow educators to export and share virtual course materials with other educators and universities. Two studies that examined the implementation of online EBI procedure are examined closely by the current author.

The first was conducted by Lovett, Rehfeldt, Garcia, and Dunning (2011) in the laboratory setting that compared the learning outcomes of four single-subject research designs (i.e., alternating treatments, changing criterion, multiple baseline, and withdrawal) between undergraduate students in an equivalence and control group. The design name (A), definition (B), graphical representation (C), and clinical vignette (D) relations for each design were trained. The equivalence group completed an online EBI training procedure (average of 85 minutes to complete) and the control group viewed a 56 minute online instructional video lecture over the designs. Both groups completed a paper-and-pencil pretest and posttest. The EBI procedure included three direct training phases for each designs' A – B, A – C, and A – D relations that required 11 out of 12 (92% block mastery criteria) correct responses to advance to the next phase. Once 92% mastery was achieved, a symmetry test over each of the three relations (B – A, C – A, and D – A) was completed. If an equivalence group participant failed to achieve 92% mastery on the symmetry test for a relation, they returned to the direct relation training. For instance, if a participant achieved 92% mastery on the directly trained A – B relation but received a score less than 11 of 12 on the subsequent B – A symmetry test, the A – B direct training was repeated. Once the direct training and symmetry phases were completed with 92% mastery, participants completed a mixed symmetry, transitivity, equivalence, generalization, and tact posttest over the relations. No feedback was provided and participants advanced regardless of performance on the tests. Results indicated that the EBI procedure completed by the equivalence group, while not significant, resulted in more effective learning outcomes than the control group.

A second study conducted by Walker and Rehfeldt (2012) built from the previous study by examining the efficacy of an online EBI method over the same four single-subject designs using an online distance learning platform. Graduate students enrolled in an online Behavior Observation and Assessment course participated in this naturalistic, fully online study that examined if trained selection-based intraverbal relations led to derived written topography responses (i.e., short answer). Results indicated that the EBI method was effective and efficient at learning the trained relations and prompting derived responses. However poor generalization and maintenance effects were found. While the authors' statement that written-topography based responding is a more impressive measure of emergent skills than selection-based responding may be accurate, MTS in the form of multiple choice tests remain a more widely implemented examination method in the college setting.

Purpose

With the commonality of selection-based responding considered, the current study sought to evaluate a match-to-sample EBI procedure for an online distance learning college course. The purpose of the current study was to further examine the efficacy of an online EBI procedure to teach college students the stimuli relations of the alternating treatments, changing criterion, multiple baseline, and withdrawal single-subject research designs. The use of appropriate experimental methodology is essential in clinical settings due to the inherent nature of a design. For example, removing an intervention would not be ethical/clinically sound when treating aggression. The current study sought to teach basic design concepts so that future practitioners are able to choose the best evaluative methods given a clinical example. The experiment examined the efficacy of an online

EBI procedure between two group conditions; an equivalence group receiving a reduced-intensity EBI procedure and a control group receiving traditional video lecture instruction. The current study sought to extend previous research in the following ways:

- Evaluate if the reduced exemplars during the current study's EBI training resulted in similar learning outcomes compared the control group as found in previous studies
- Examine if the EBI training promoted generalization to novel exemplars

CHAPTER III

Method

Sixty-seven graduate students enrolled in an online Behavior Analysis and Management course participated in the experiment. Participation requirements were outlined in a consent document at the beginning of the experiment. The subject matter of the experiment was a preexisting aspect of the course curriculum and resulted in course credit equivalent to 2% of the final grade for participation alone, regardless of outcome or assigned group. Group assignment was determined via a random number generator.

Eight participants' denied consent resulting in exclusion from any reporting of their results in accordance to internal review board ethical guidelines. In addition, criterion for exclusion from study was defined by the current authors as follows. First all pretest, training, and posttest phases were required for inclusion. This criteria removed five and six participants in the equivalence and lecture groups respectively. Second, participants exhibiting an interresponse time of four seconds or less between questions were removed. This second criteria removed one participant from the equivalence group. Third, participants scoring 25% or less during the posttest phase were excluded due to suspected guessing (given multiple choice of four options one would expect 25% to occur by chance or random guessing). This third criterion resulted in removal of two and one participant(s) from the equivalence and lecture groups respectively. In addition, one

participant in the lecture group scored 100% on the 24 pretest questions demonstrating mastery of all train and test relations prior to intervention. Thus no need for further evaluation was warranted. The final pool consisted of a total of 45 participants: 22 in the equivalence group, 23 in the lecture group.

Design

The study implemented a pretest – training – posttest sequence for both groups. The equivalence group received a pretest – EBI training – posttest sequence. Training for the lecture group consisted of a pretest – traditional video lecture – posttest sequence (see training). In accordance with the course structure, all participants were provided exactly 10 days to complete all phases of their perspective groups' assigned material beginning the moment the CANVAS module was unlocked. This narrow window controlled for maturation and treatment diffusion that may have occurred through additional exposure to course content across subsequent weeks. The automated, online nature of the experiment controlled for possible instrumentation threats as well as observer drift and any experimenter bias in data collection. The designed methodology provided a useful comparison of learning outcomes between groups both prior to and after intervention.

Stimulus Materials

As with the Lovett et al. (2011) and Walker and Rehfeldt (2012) studies, the four single-subject designs were organized by name (i.e., A stimuli), definition (i.e., B stimuli), graphical representation (i.e., C stimuli), and clinical vignettes (i.e., D stimuli). The B and D stimuli were adapted from information provided by Alberto and Troutman (2009), Cooper, Heron, and Heward (2014) and Kazdin (2011). To increase external validity, the C stimuli included a sample of studies pulled from a comprehensive list of

the most commonly assigned articles in ABA graduate programs (Pastrana et al., 2018). Exemplar stimuli remained consistent throughout pretest, training, and posttest phases with the exception of generalization test items (see Table 1 for training and test exemplars). The exemplar stimuli provided in the Lovett et al. (2011) study for B and D stimuli were adapted as feedback for the definition to name (i.e., B – A) and clinical vignette to name (i.e., D – A) direct training trials. Omission of client names and other salient features from stimulus sets ensured participants were attending to critical design elements and not irrelevant portions (e.g., names changed to 'client', definition exemplars began with 'This design involves' and clinical vignette exemplars began with 'A therapist').

Procedure

Pretest – Posttest

As mentioned, pretest and posttest measures were identical for both groups. The pretest included 24 randomized multiple-choice questions. This array included 12 symmetry relations (A – B, A – C, and A – D), 6 transitivity test trials (D – C, B – D, C – B), and 6 equivalence test trials (C – D, D – B, and B – C) for each of the four designs. The posttest included 24 identical pretest questions and 16 additional tests for generalization. Generalization tests consisted of two novel graphical representation (i.e., C) and clinical vignette (i.e., D) exemplars for each designs' C – A and D – A relations. Taken together, the posttest included a total of 40 randomized items (see Table 2 for test procedures). Participants were provided a total score indicating the number correct following both pretest and posttest phases. Correct/incorrect answers were not provided and the participants could not view the questions after each test was submitted. Once the

equivalence group submitted the pretest, the four EBI training trials (see Appendix B for outlined EBI training procedure) were completed.

Training

Phase 1 training consisted of four exemplar items covering the designs' definition to design name (i.e., B – A) relation. Once submitted, feedback consisted of a pop-up box that presented written praise (e.g., correct, nice work, fantastic, etc.) for correct responses with an additional feedback exemplar adapted from Lovett et al. (2011). Feedback included the definition exemplar stimuli adapted from the previous study for the corresponding incorrect answer. For example, if the correct answer for a B^1 – A^1 trial was *withdrawal* and the hypothetical participant selected *multiple baseline*, the feedback provided the previous study's definition exemplar for the multiple baseline design. This ensured that two exemplars were provided for each designs' B stimuli, regardless of response accuracy. No specific correct answers were delivered contingent upon incorrect responding. Participants were required to achieve 100% mastery criteria (i.e., 4 out of 4) to advance to the next phase.

Phase two training included four exemplar items covering each designs' graphical representation to design name (i.e., C – A) relation. Feedback followed the previous training trial's format; correct responses resulted in written praise and an additional exemplar; incorrect responses were provided a corresponding graphical representation exemplar from Pastrana et al. (2018). These feedback exemplars were selected to be topographically dissimilar to the training exemplar while still meeting the design requirements for generalization purposes. Participants advanced to phase three after meeting 100% mastery criteria (i.e., 4 out of 4).

This third training phase consisted of four exemplar items covering the clinical vignette to design name (i.e., D – A) stimuli relations. Feedback included the corresponding clinical vignette exemplar adapted from Lovett et al. (2011) which removed hypothetical client names to avoid irrelevant features taking stimulus control. The participants advanced once 100% mastery criteria (i.e., 4 out of 4) was achieved. Phase four randomized and combined the previous three trials' exemplars for a total of 12 items. The same feedback exemplar was provided for each response corresponding to the specific relation and selected answer as with the previous training phases. Once 100% mastery (i.e., 12 out of 12) was achieved, the equivalence group participants continued to the posttest.

The lecture group completed the same online CANVAS pretest and posttest as the equivalence group. Instead of the EBI protocol training described above, the lecture group viewed a 22-minute video which included PowerPoint slides with an audio lecture. First the lecture group participants were taught the basics of single-subject designs. Next the design name, definition, graphical representation, and clinical vignette was covered for each of the four designs. The same training and feedback exemplars as the equivalence group was incorporated into the video lecture to maintain the comparison validity. This ensured that the results would solely be a consequence of the instructional procedure received by either group.

Interobserver Agreement

Participant responses provided in CANVAS were exported to Microsoft Excel using the student analysis feature. Point-by-point interobserver agreement on the pretest and posttest for both groups and the training for the equivalence group was calculated

using the Excel data sheet as a permanent product. This interobserver agreement methodology involved examining the sum of agreed upon scores between CANVAS and the Excel file. Then the number of agreements was divided by the sum of agreements and disagreements (i.e., agreements ÷ agreements + disagreements). Agreement between participant scores on CANVAS and the Excel file was calculated at 100%. A high level of agreement ensured the instrumentation validity between the automated CANVAS procedure and data export.

Social Validity Survey

Participants of both groups completed a social validity survey (see Appendix C) following the conclusion of all pretest, training, and posttest experimental phases. Four questions regarding the usefulness of EBI or video lecture protocols were rated using a 7-point Likert scale. These items measured the participants' confidence in their knowledge of the designs, the degree to which they would want to receive the instructional protocol in the future, and whether they felt the time commitment to complete the protocol was appropriate compared to what they had learned. A fourth question measured each participants' level of frustration while completing the instructional protocol.

CHAPTER IV

Results

Results for the pretest and posttest for both equivalence and lecture groups are presented in Table 3 and Figure 1. The equivalence group's mean score on the pretest was 58.52% (*SD* = 0.17) with a range from 29.3% to 88%. On the posttest, the equivalence group's mean score was 90.53% (*SD* = 0.12) with a range from 63% to 100%. The average improvement from pretest to posttest was 32.01% and a paired sample *t* test revealed this change was significant ($M = .32$, $SD = .19$, $t(21) = -8.05$, $p < .001$). The lecture group's mean score on the pretest was 52.17% (*SD* = 0.12) with a range from 13% to 79%. On the posttest, the lecture group's mean score was 88.59% (*SD* = 0.17) with a range from 29% to 100%. The average improvement from pretest to posttest was 36.41% and a paired sample *t* test revealed this was change was also significant ($M = -.36$, $SD = .22$, $t(22) = -7.99$, $p < .001$). A one-way MANOVA revealed that there was no significant difference in the pretest ($M = .55$, $SD = .17$, $p = .204$) and posttest ($M = .90$, $SD = .15$, $p = .654$) scores between the equivalence and lecture participants.

Generalization Items

Results for generalization to novel exemplars across groups are provided in Figure 2. The equivalence group's mean score for the 8 novel graphical representation

exemplars was 86.36% (*SD* = .22, range, 12.5% to 100%). For the 8 novel clinical vignette exemplars, the mean score was 78.98% (*SD* = .27, range, 0% to 100%). One participant demonstrated no generalization to novel exemplars. The lecture group's mean score for the 8 novel graphical representation exemplars was 85.87% (*SD* = .21, range, 25% to 100%). On the 8 novel clinical vignette exemplars, the mean score was 82.07% (*SD* = .21, range, 12.5% to 100%). An independent sample *t* test revealed no significant differences between the groups' mean scores for generalization to novel graphical representation (i.e., $t(43) = .08$, $p = .975$) and clinical vignette (i.e., $t(43) = -.43$, $p = .267$) exemplars.

Equivalence Class Formation

The number of attempts each equivalence group participant required to reach the 100% mastery criteria are provided in Table 4. The first training consisted of the B – A direct relations and the mean number of attempts to reach mastery was 2.14 (range, 1 to 6). The second consisted of the C – A direct relations and the mean number of attempts to reach mastery was 3.05 (range, 1 to 8). The third training consisted of the D – A direct relations and the mean number of attempts to reach mastery was 1.64 (range, 1 to 4). The fourth training combined the previous three trainings and the mean number of attempts to reach mastery was 2.36 (range, 1 to 6).

Social Validity Survey

The mean agreement rating for confidence in knowledge of single subject designs was 4.14 (range, 1 to 6) for the equivalence group and 4.22 (range, 1 to 6) for the lecture group. The mean degree to which participants wanted to receive similar instruction in the future was 3.29 (range, 1 to 6) for the equivalence group and 4.17 (range, 1 to 7) for the

lecture group. The mean degree of frustration while completing the instructional protocol was 5.12 (range, 1 to 7) for the equivalence group and 4.48 (range, 2 to 7) for the lecture group. This item (i.e., item 3) reflects a negative review of the instructional protocol, meaning a higher score is correlated with a more negative review. The mean degree to which participants reported the appropriateness of the time required to complete the instructional protocol compared to what was learned was 4.53 (range, 1 to 6) for the equivalence group and 4.96 (range, 1 to 7) for the lecture group. The overall mean score for positive items (i.e., items 1, 2, and 4) was 11.94 (range, 4 to 18) for the equivalence group and 13.35 (range, 3 to 18) for the lecture group. Positive reviews are correlated with higher ratings on these three items.

CHAPTER V

Discussion

The current study provided four extensions from previous EBI research. The first examined if a reduced-intensity EBI procedure (i.e., one direct training and one feedback exemplar for each relation) resulted in derived relational responding between complex stimuli not directly trained (i.e., Brodsky & Fienup, 2018). Results demonstrate that the EBI procedure significantly improved mastery of complex relations from pretest to posttest. This finding is valuable given the challenges educators face in higher education such as limited time and resources. Educators often must receive continuing education to maintain a credential, conduct research, and provide guidance to advisees. Perhaps the most time consuming is the teaching, organization, and effective incorporation of relevant curriculum for multiple classes, each with a multitude of students. Given these restraints, educational tools that are efficient as well as effective are essential. Reduced-intensity EBI procedures may provide educators an efficient methodology to promote the effective learning of students in higher education.

Learning outcomes of the equivalence group were compared to the lecture group. The video provided to the lecture group consisted of a Microsoft PowerPoint presentation supplemented with an audio lecture and was specifically designed to include the same information over each designs' relations as the equivalence group. The video first briefly

described what makes single-subject designs different than between-subject designs. Next the video discussed key aspects of single-subject designs such as functional relationships, baseline logic, and independent, dependent, and extraneous variables. Lecture group participants were taught the basics of visually inspecting graphs to determine the impact of an intervention from baseline to treatment conditions (i.e., changes in trend, mean, latency, and level). Having established the introductory concepts of single-subject designs, the video then delved into the four designs targeted by the current study (see Appendix D). As mentioned, identical training and feedback exemplars were used across groups to facilitate valid comparisons and isolate the independent effects of the training methods. Lecture group participants were first prompted to read the definition and clinical vignette provided in the PowerPoint slide (i.e., the equivalence group's training exemplar). Verbal instructions provided the same definition and vignette feedback exemplar as the equivalence group participants. Finally, the participants were provided identical training and feedback graphical representation exemplars. Significant improvements were found for both groups across phases while no significant differences were found between groups. These findings suggest both methods are effective. Overall the results are consistent with Lovett et al. (2011) despite the reduced intensity of EBI procedures.

A second extension involved the use of a fully online format using CANVAS. This demonstrates the capability of a university's specific online learning system to include EBI procedures. Relatively few faculty are fluent in visual basic programing (i.e., Lovett et al., 2011; Walker & Rehfeldt, 2012) compared to CANVAS, BlackBoard, and other online systems. That is the current author anticipates that making quizzes in a

preexisting online system is easier and less time consuming than programing procedures in visual basic with the same functionality. This capacity increases the external validity of EBI procedures. While many universities use different Learning Management Systems, the match-to-sample nature of EBI may allow for its implementation across a wide variety systems. Adaptability is a significant factor to be considered by educators selecting potential instructional methodologies.

A third extension was the graphical representation exemplars were selected from the most common articles assigned to Applied Behavior Analysis graduate students provided by Pastrana et al. (2018). While investigators who create graphs specifically for the EBI procedure may demonstrate that learning occurred, this would not necessarily demonstrate its applicability to real-life examples. Incorporating material the students will likely come in contact with in the future does demonstrate this ability and allows for greater external validity of the results. A primary goal of education is provide instruction that will impact the learner beyond the testing measure implemented in the specific course. The current author’s third extension may benefit the pursuit of this goal.

The fourth extension by the current author was the EBI procedure required a stringent 100% block-mastery criteria for each training. Fienup and Brodsky (2017) recommended that EBI procedures implement stringent mastery criteria due to the reduced errors observed during training and increased maintenance of the learned relations. This stringent criteria was required due to the previous authors’ recommendation and the significant reduction of exemplars presented in current study’s EBI training. Previous research is extended by the current study in the former ways and the results demonstrate the effectiveness of the EBI procedure to teach graduate students

the introductory concepts of four single-subject designs commonly implemented in behavior analysis.

Limitations

There were several limitations of the current study. The integrity of the participants' responses was not controlled by the current study. Participants may have used outside material, taken pictures of the EBI trainings/video lecture, and/or collaborated with other participants. However, this limitation extends to online distance learning platforms in general. Future researchers may control for this limitation by implementing screen and video monitoring features to ensure the integrity of responses in online formats. A second limitation was that the total time to complete the procedure was not accessible. CANVAS currently provides the total time each participant spent on each item. However if the CANVAS page was left open, it continued to accumulate time (e.g., Participant 4 in the equivalence group spent 47 hours and 55 minutes on the procedure). Despite this limitation, the current author concluded the EBI procedure implemented in the study was more efficient than previous studies due the significant reduction to the required number of trainings and exemplars.

A third limitation was participants who completed the entire procedure at once may have different learning outcomes compared to those who completed it over a several day period. Participants who completed the procedure in a single sitting may have scored higher on the posttest. However, the spacing effect (Vlach & Sandhofer, 2012) states that learning which is spaced out across time results in stronger long-term maintenance of the material. Future researchers may control for this limitation by requiring participants to complete the procedure in a single sitting. Additionally, future investigations may wish

to examine the learning outcomes and their maintenance between experimental groups finishing the EBI procedure in one session and the other across multiple sessions. A final limitation was tests for maintenance were beyond the scope of the current researcher. Another primary goal of educators is that learned material will generalize across time. Without an evaluation of this goal, the current EBI procedure's efficacy cannot be determined. Both short and long-term maintenance probes should be implemented by future investigators to address this limitation.

Future Directions

Despite the promising results of the current study, further replications by independent investigators are desired to ensure the procedural efficacy of reduced-intensity EBI procedures. Future investigators may design an experiment where one group receives an EBI procedure similar to the current study and the other receives a more traditional EBI procedure (i.e., multiple trainings and exemplars for each relation with less stringent mastery criteria) to allow for a direct comparison of learning outcomes, generalization to novel exemplars, and maintenance of the relations. If the results for both groups do not differ significantly, reduced-intensity procedures may become an applicable EBI methodology and supplement the instructional repertoire of educators across disciplines.

Future researchers and educators may benefit from an additional item that was added to the social validity survey. This subjective item asked the participants to provide questions, comments, and concerns experienced while completing the procedure. Six participants in the equivalence group expressed frustration with the trial-and-error nature of EBI. Along with the stringent 100% mastery criteria, this may contribute to the lower

positive review provided by this group. Additional investigators may wish to revise this procedure to address this finding and mitigate any anxiety. For example, an equivalence group could receive solely EBI and a combined group could receive brief instruction (e.g., article, short video lecture, etc.) followed by the EBI procedure. Experiments of this nature may allow educators to implement effective, practical EBI procedures that are also desirable to the learners.

Conclusion

Sidman (1971) was the first study to demonstrate the phenomena of stimulus equivalence by observing a disabled youth's ability to derive relations not directly paired during training. Since then, investigators have examined the utility of stimulus equivalence as an educational tool (i.e., EBI) for increasingly complex stimuli relations. These investigations led to an examination of EBI in higher education settings, eventually culminating in the meta-analysis provided by Brodsky and Fienup (2018). EBI was determined to be an effective and efficient methodology with similar learning outcomes compared to traditional instruction (e.g., lecture). The current study examined if a shorter, more practical version of EBI would produce similar outcomes and results indicate similar outcomes were achieved. While the efficacy of the current EBI procedure cannot be determined at present, this finding warrants the procedure's future investigation. In the 1960s, Skinner developed the Programmed Learning Theory to allow learners to receive immediate reinforcement and individualized instruction using independent learning machines (McDonald, Yanchar, & Osguthorpe, 2005). These learning machines provided an all-inclusive, response/reward mechanism which divided the learning process into a large number of smalls steps. This automated learning

methodology accommodated each students' individual rate of learning. With the development of modern technology, automated learning procedures such as EBI are available to educators on a much larger scale.

A goal of applied behavior analysis (ABA) is to improve educational methodologies following the seven core dimensions of ABA described by Baer, Wolf, and Risley (1968). The current study accomplished this goal by applying Sidman's basic research on stimulus equivalence to EBI in a higher education setting. The procedure was rooted in empirically validated principles of behavior and was described in such detail that future investigators may replicate its methodology. Results demonstrate the effectiveness of the procedure to train each relation and allow for generalization to novel exemplars. The current author's adherence to the seven core dimensions of ABA allow future investigators to further examine reduced-intensity EBI procedures as a potential improvement to current instructional paradigms.

Table 1

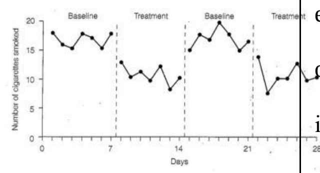
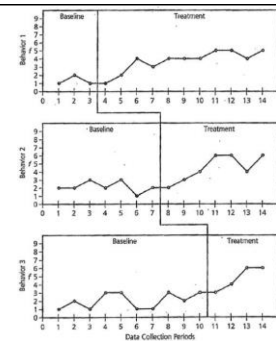
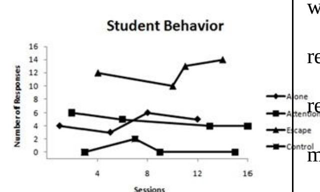
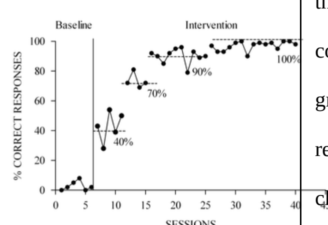
	A	**B**	**C**	**D**
1	Withdrawal	This design involves the repeated implementation and removal of the intervention in order to test its effect on the target behavior.		A therapist wants to evaluate the effectiveness of an intervention on increasing a client's word comprehension by implementing and then removing the treatment.
2	Multiple Baseline	This design involves the introduction of an intervention at staggered times to determine its effect on two or more individuals with the same behavior, multiple behaviors in one individual, or settings.		A therapist wants to evaluate the effectiveness of an intervention on the reduction of three clients' self-injurious behavior to escape an aversive task.
3	Alternating Treatments	This design involved a procedure in which two or more interventions or conditions are interchanged systematically to evaluate each variables' effect on a target behavior.		A therapist wants to evaluate whether a noncontingent reinforcement or a differential reinforcement procedure is more effective at increasing a client's task completion.
4	Changing Criterion	This design involves gradually and successively increasing or decreasing the target response level required for reinforcement until response levels reach a desired terminal goal.		A therapist wants to evaluate the effectiveness of a reading comprehension procedure by gradually increasing the response requirements of a client to reach a predetermined end goal.

Table 1: Exemplar stimuli used during training, pretest, and posttest. Feedback exemplars and generalization items not included.

Table 2

Posttest			
Pretest			
Symmetry	Transitivity	Equivalence	Generalization
A^1-B^1	D^1-C^1	C^3-D^3	A^1-C^1 (2)
A^2-B^2	D^2-C^2	C^4-D^4	A^2-C^2 (2)
A^3-B^3	B^3-D^3	D^1-B^1	A^3-C^3 (2)
A^4-B^4	B^4-D^4	D^2-B^2	A^4-C^4 (2)
A^1-C^1	C^1-B^1	B^2-C^2	A^1-D^1 (2)
A^2-C^2	C^3-B^3	B^4-C^4	A^2-D^2 (2)
A^3-C^3			A^3-D^3 (2)
A^4-C^4			A^4-D^4 (2)
A^1-D^1			
A^2-D^2			
A^3-D^3			
A^4-D^4			

Table 2: Stimuli relations evaluated in the pretest and posttest.

Table 3

Equivalence Group Quiz Scores

Participant	Pretest	Posttest	Pretest to posttest change
1	13	22	9
2	19	22	3
3	15	17	2
4	13	24	11
5	16	15	-1
6	13	22	9
7	16	24	8
8	8	22	14
9	7	23	16
10	13	22	9
11	16	23	7
12	21	23	2
13	14	24	10
14	17	24	7
15	9	19	10
16	16	20	4
17	20	24	4
18	8	15	7
19	16	24	8
20	18	21	3
21	9	24	15
22	12	24	12

Lecture Group Quiz Scores

Participant	Pretest	Posttest	Pretest to posttest change
1	3	22	19
2	13	18	5
3	8	19	11
4	14	23	9
5	17	24	7
6	8	24	16
7	12	21	9
8	15	24	9
9	19	21	2
10	13	24	11
11	17	21	4
12	11	24	13
13	15	23	8
14	7	24	17
15	14	24	10
16	13	24	11
17	11	7	-4
18	10	24	14
19	7	15	8
20	16	23	7
21	15	24	9
22	13	16	3
23	17	20	3

Table 3: Pretest and posttest questions are the same replicated 24 questions.

Table 4

Participant	Training 1	Training 2	Training 3	Training 4
1	1	2	1	4
2	1	3	1	1
3	3	4	1	6
4	3	1	1	1
5	1	4	2	4
6	2	2	1	1
7	2	2	2	3
8	2	8	3	4
9	5	4	2	1
10	1	5	2	2
11	1	2	1	1
12	1	1	1	1
13	2	2	4	2
14	3	2	1	5
15	2	2	2	4
16	1	6	2	1
17	1	2	1	1
18	6	2	1	2
19	2	4	2	3
20	1	2	1	1
21	4	4	2	3
22	2	3	2	1

Table 4: Equivalence group attempts per training to reach criterion.

Figure 1

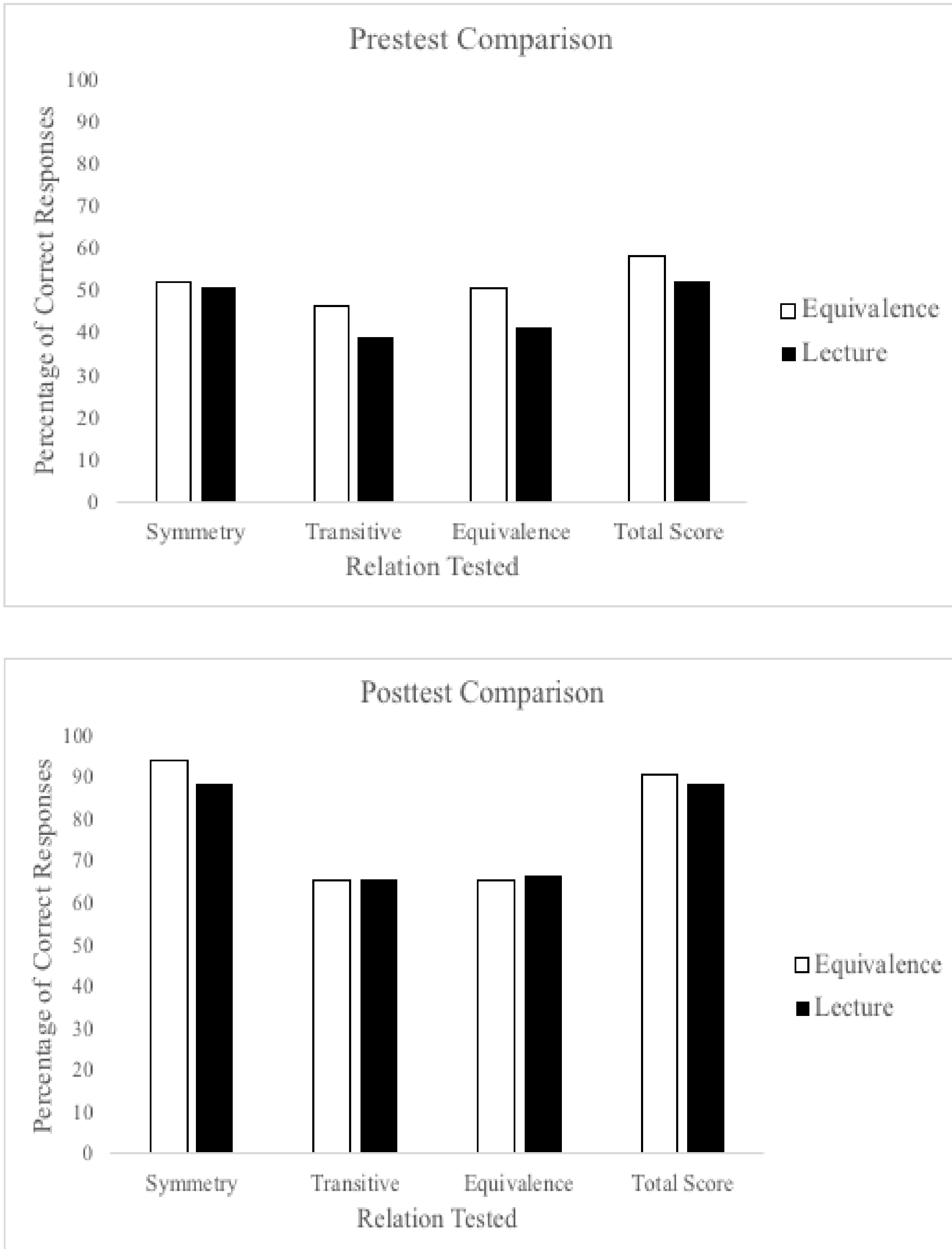


Figure 1: Pretest and posttest total mean comparison between equivalence and lecture groups.

Figure 2

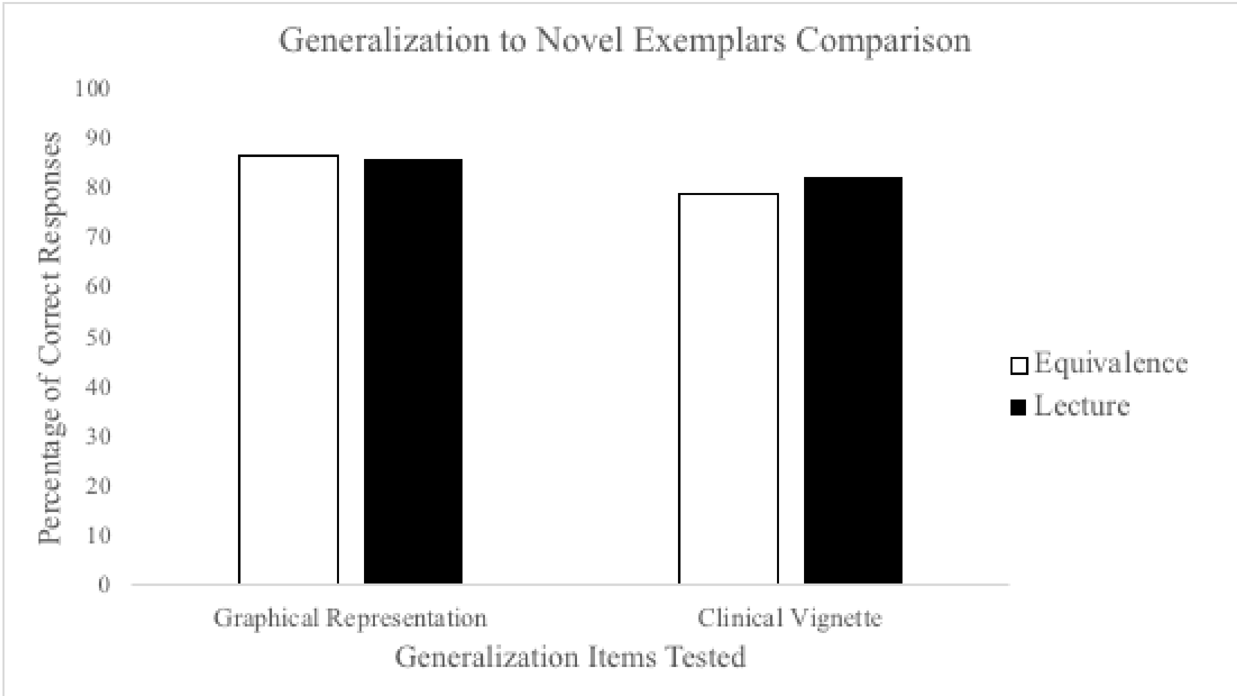


Figure 2: Eight novel graphical representations and clinical vignettes, two for each relation, was added to the posttest.

References

Alberto, P. A., & Troutman, A. C. (2009). *Applied behavior analysis for teachers*. Upper Saddle River, NJ: Merrill/Pearson.

Allen, E. I., & Seaman, J. (2011). *Going the distance: Online education in the United States.* Babson Survey Research Group.

Alter, M. M., & Borrero, J. C. (2015). Teaching generatively: Learning about disorders and disabilities. *Journal of Applied Behavior Analysis*, *48*(2), 376-389.

Arntzen, E. (2004). Probability of equivalence formation: Familiar stimuli and training sequence. *The Psychological Record*, *54*(2), 275-291.

Austin, J., & Carr, J. E. (2000). Behavioral approaches to college teaching. In J. L. Austin (Ed.), *Handbook of Applied Behavior Analysis* (pp. 449-472). Reno, Nev: Context Press.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, *1*(1), 91–97.

Barbetta, P. M., & Skaruppa, C. L. (1995). Looking for a way to improve your behavior analysis lectures? Try guided notes. *The Behavior Analyst*, *18*, 155-160.

Brodsky, J., & Fienup, D. M. (2018). Sidman Goes to College: A Meta-Analysis of Equivalence-Based Instruction in Higher Education. *Perspectives on Behavior Science, 41*(1), 95-119.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2014). *Applied behavior analysis*. Harlow: Pearson Education.

Dillon, M. J., Kent, H. M., & Malott, R. W. (1980). A supervisory system for accomplishing long-range projects: An application to master's thesis research. *Journal of Organizational Behavior Management*, *2*(3), 213-227.

Dougher, M. J., Augustson, E., Markham, M. R., Greenway, D. E., & Wulfert, E. (1994). The transfer of respondent eliciting and extinction functions through stimulus equivalence classes. *Journal of the Experimental Analysis of Behavior*, *62*(3), 331-51.

Fields, L., Travis, R., Roy, D., Yadlovker, E., de Aguiar-Rocha, L., & Sturmey, P. (2009). Equivalence class formation: A method for teaching statistical interactions. *Journal of Applied Behavior Analysis*, *42*(3), 575-593.

Fienup, D. M., & Brodsky, J. (2017). Effects of mastery criterion on the emergence of derived equivalence relations. *Journal of Applied Behavior Analysis*, *50*(4), 843-848.

Fienup, D. M., & Critchfield, T. S. (2011). Transportability of equivalence-based programmed instruction: Efficacy and efficiency in a college classroom. *Journal of Applied Behavior Analysis*, *44*(3), 435-450.

Fienup, D. M., Mylan, S. E., Brodsky, J., & Pytte, C. (2016). From the laboratory to the classroom: The effects of equivalence-based instruction on neuroanatomy competencies. *Journal of Behavioral Education*, *25*(2), 143-165.

Greville, W. J., Dymond, S., & Newton, P. M. (2016). The student experience of applied equivalence-based instruction for neuroanatomy teaching. *Journal of Educational Evaluation for Health Professions*, *13*, 32.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: a post-Skinnerian account of human language and cognition*. New York: Kluwer Academic/Plenum Publishers.

Hayes, S. C., Strosahl, K., 1950, & Wilson, K. G. (2012). *Acceptance and commitment therapy: The process and practice of mindful change* (2nd ed.). New York: Guilford Press.

Hull, C. L. (1939). The problem of stimulus equivalence in behavior theory. *Psychological Review, 46*(1), 9-30.

Johnson, P. E., Perrin, C. J., Salo, A., Deschaine, E., & Johnson, B. (2016). Use of an explicit rule decreases procrastination in university students. *Journal of Applied Behavior Analysis*, *49*(2), 346-358.

Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings*. New York: Oxford University Press, 2011.

Keller, F. S. (1968). Good-bye, teacher. *Journal of Applied Behavior Analysis*, *1*, 79-89.

Kennedy, C. H., Itkonen, T., & Lindquist, K. (1994). Nodality effects during equivalence class formation: An extension to sight-word reading and concept development. *Journal of Applied Behavior Analysis*, *27*(4), 673-83.

Lindsley, O. R. (1990). Precision teaching: By teachers for children. *Teaching Exceptional Children*, *22*(3), 10-15.

Lindsley, O. R. (1992). Precision teaching: Discoveries and effects. *Journal of Applied Behavior Analysis, 25*, 51-57.

Lovett, S., Rehfeldt, R. A., Garcia, Y., & Dunning, J. (2011). Comparison of a stimulus equivalence protocol and traditional lecture for teaching single subject designs. *Journal of Applied Behavior Analysis*, *44*(4), 819-33.

McDonald, J. K., Yanchar, S. C., & Osguthorpe, R. T. (2005). Learning from programmed instruction: Examining implications for modern instructional technology. *Educational Technology Research and Development*, *53*(2), 84-98.

Neef, N. A., McCord, B. E., & Ferreri, S. J. (2006). Effects of guided notes versus completed notes during lectures on college students' quiz performance. *Journal of Applied Behavior Analysis*, *39*(1), 123-130.

Ninness, C., Rumph, R., McCuller, G., Harrison, C., Ford, A. M., & Ninness, S. K. (2005). A functional analytic approach to computer-interactive mathematics. *Journal of Applied Behavior Analysis*, *38*, 1–22.

O'Neill, J., Rehfeldt, R. A., Ninness, C., Muñoz, B. E., & Mellor, J. (2015). Learning Skinner's verbal operants: comparing an online stimulus equivalence procedure to an assigned reading. *The Analysis of Verbal Behavior*, *31*(2), 255.

Pastrana, S. J., Frewing, T. M., Grow, L. L., Nosik, M. R., Turner, M., & Carr, J. E. (2018). Frequently assigned readings in behavior analysis graduate training programs. *Behavior Analysis in Practice*, *11*(3), 267-273.

Pytte, C. L., & Fienup, D. M. (2012). Using equivalence-based instruction to increase efficiency in teaching neuroanatomy. *Journal of Undergraduate Neuroscience Education*, *10*(2), A125-131.

Sella, A. C., Mendonça Ribeiro, D., & White, G. W. (2014). Effects of an online stimulus equivalence teaching procedure on research design open-ended questions

performance of international undergraduate students. *The Psychological Record*, *64*(1), 89-103.

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research, 14*(1), 5-13.

Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior*, *74*(1), 127-146.

Sidman, M. (1994). *Equivalence relations and behavior: a research story*. Boston: Authors Cooperative.

Sidman, M., & Cresson, O. (1973). Reading and crossmodal transfer of stimulus equivalence in severe retardation. *American Journal of Mental Deficiency*, 77, 515–523.

Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior*, *37*(1), 5-22.

Skinner, B. F. (1984). The shame of American education. *American Psychologist, 39*, 947-954.

Spieler, C., & Miltenberger, R. (2017). Using awareness training to decrease nervous habits during public speaking. *Journal of Applied Behavior Analysis*, *50*(1), 38-47.

Stocco, C. S., Thompson, R. H., Hart, J. M., & Soriano, H. L. (2017). Improving the interview skills of college students using behavioral skills training. *Journal of Applied Behavior Analysis, 50*(3), 495-510.

Törneke, N., & ProQuest (Firm). (2010). *Learning RFT: An introduction to relational frame theory and its clinical applications*. Oakland, Calif: New Harbinger Publications.

Varelas, A., & Fields, L. (2017). Equivalence based instruction by group based clicker training and sorting. *The Psychological Record*, *67*(1), 71.

Vlach, H. A., & Sandhofer, C. M. (2012). Distributing learning over time: The spacing effect in children's acquisition and generalization of science concepts. *Child Development*, *83*(4), 1137-1144.

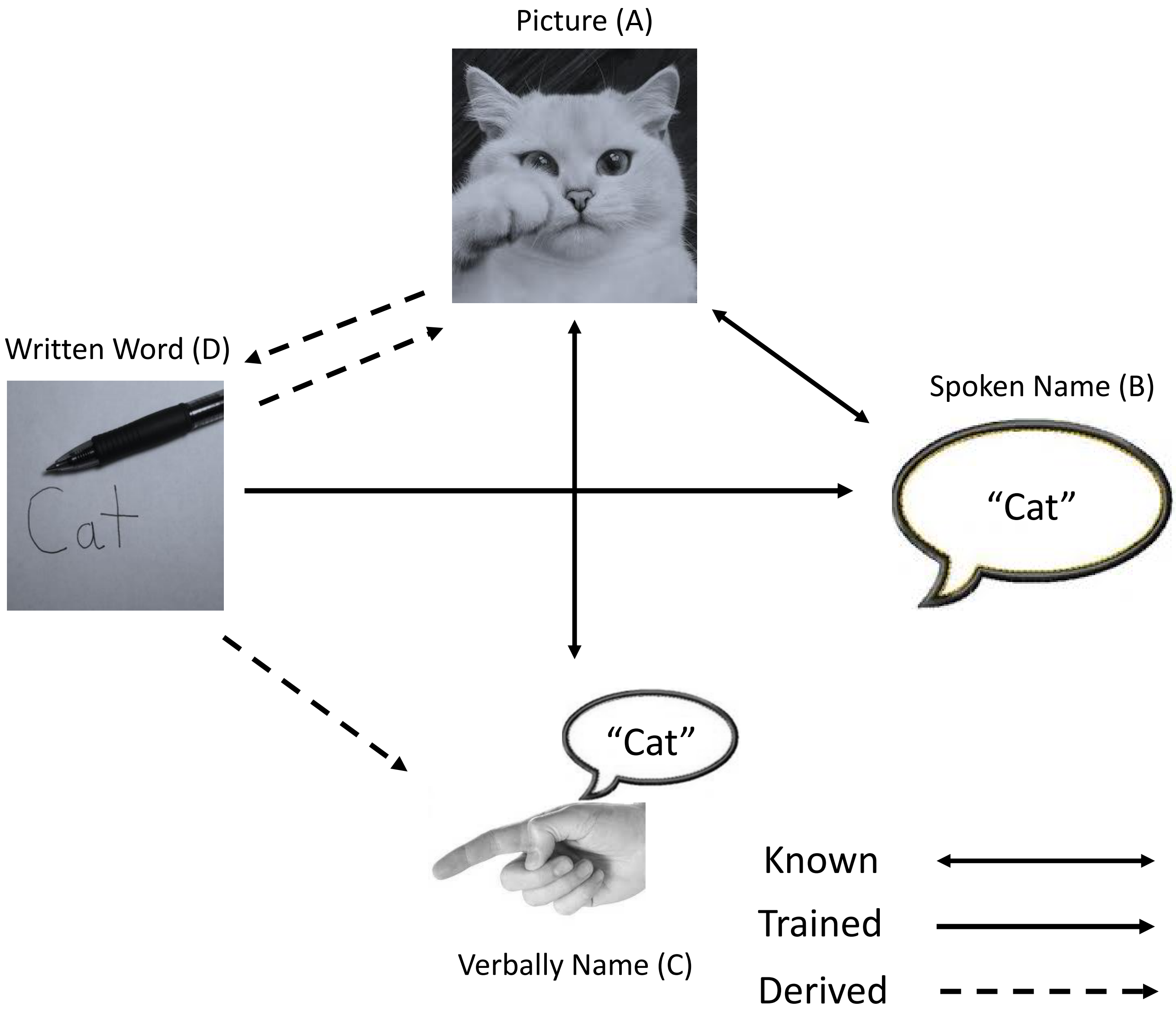
Walker, B. D., & Rehfeldt, R. A. (2012). An evaluation of the stimulus equivalence paradigm to teach single-subject design to distance education students via blackboard. *Journal of Applied Behavior Analysis*, *45*(2), 329-344.

Walker, B. D., Rehfeldt, R. A., & Ninness, C. (2010). Using the stimulus equivalence paradigm to teach course material in an undergraduate rehabilitation course. *Journal of Applied Behavior Analysis, 43*(4), 615-33.

Zinn, T. E., Newland, M. C., & Ritchie, K. E. (2015). The efficiency and efficacy of equivalence-based learning: A randomized controlled trial. *Journal of Applied Behavior Analysis, 48*(4), 865-8.

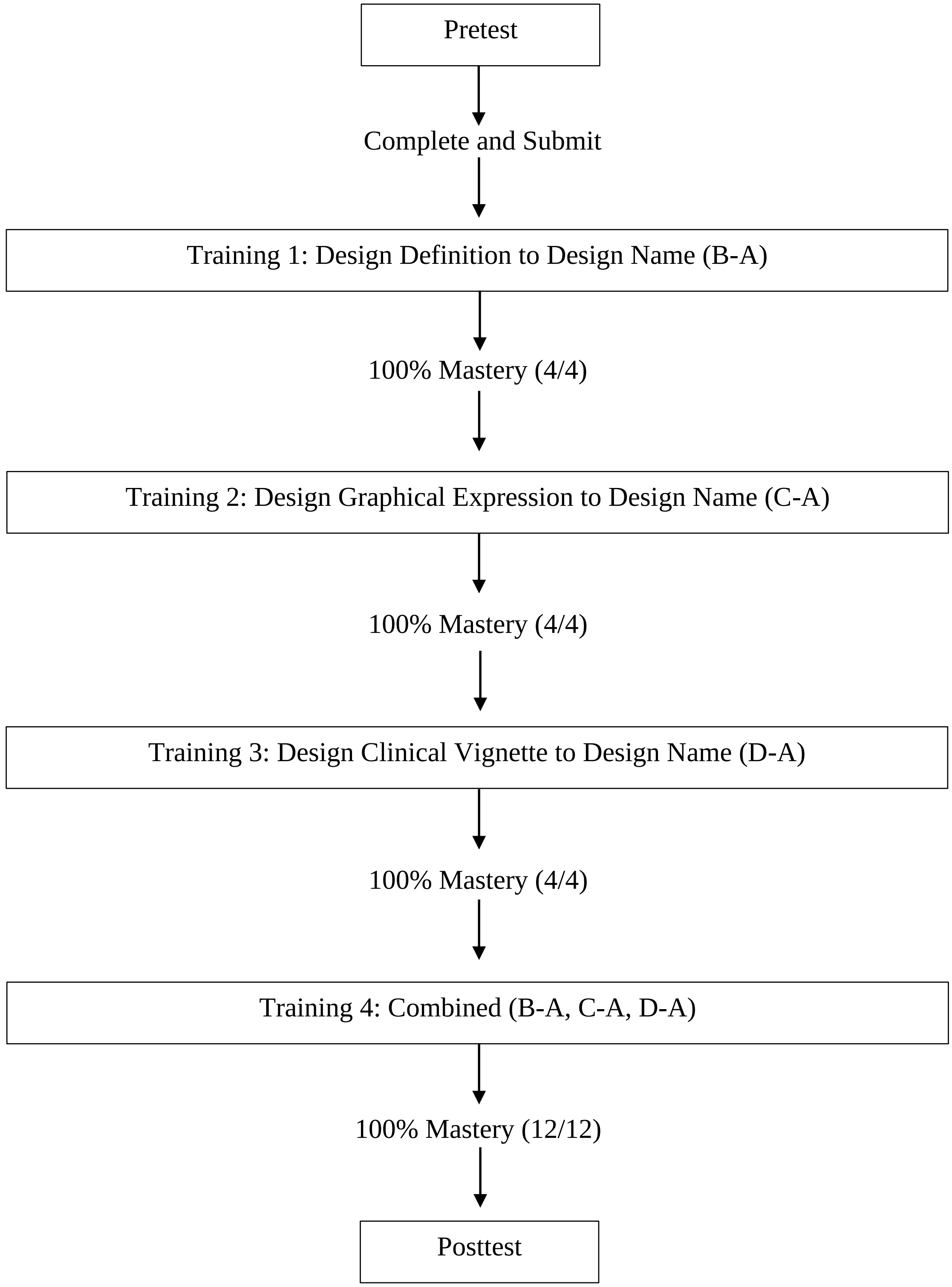
APPENDIX

Appendix A



Appendix A: Results from Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research*, *14*(1), 5-13.

Appendix B



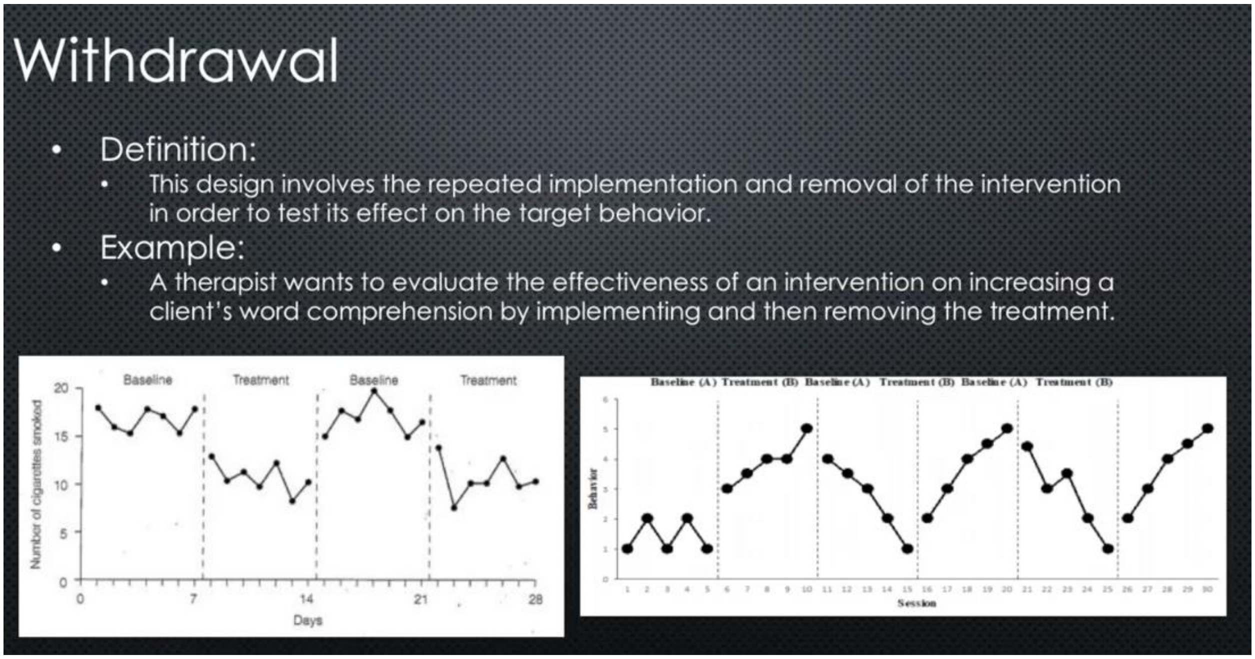
Appendix B: EBI training procedure

Appendix C

I now feel confident in my knowledge of single-case designs.						
1	2	3	4	5	6	7
Strongly disagree			Neutral			Strongly agree
I would like to learn new material using this instructional method in the future.						
1	2	3	4	5	6	7
Strongly disagree			Neutral			Strongly agree
I felt frustrated at times while using this instructional method (pretest, posttest excluded).						
1	2	3	4	5	6	7
Strongly disagree			Neutral			Strongly agree
I felt that the time commitment for this instructional method was appropriate in relation to the amount I feel I've learned.						
1	2	3	4	5	6	7
Strongly disagree			Neutral			Strongly agree

Appendix C: Social validity survey

Appendix D



Appendix D: Example of how the relations were presented to the lecture group in the video.