

ESTABLISHING QUALITY STANDARDS FOR APPLIED BEHAVIOR ANALYTIC SKILL-
ACQUISITION INTERVENTIONS: A TRANSLATIONAL MODEL WITH
UNDERGRADUATE STUDENTS

by

Samantha Bergmann

A Dissertation Submitted in
Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
in Psychology

at

The University of Wisconsin-Milwaukee

May 2018

ABSTRACT

ESTABLISHING QUALITY STANDARDS FOR APPLIED BEHAVIOR ANALYTIC SKILL-ACQUISITION INTERVENTIONS: A TRANSLATIONAL MODEL WITH UNDERGRADUATE STUDENTS

by

Samantha Bergmann

The University of Wisconsin-Milwaukee 2018
Under the Supervision of Professor Tiffany Kodak

Treatment integrity is the extent to which components of an intervention are implemented as intended (Gresham, 1989). Recent behavior-analytic literature has begun to evaluate the effects of impaired treatment integrity on efficacy and efficiency of skill-acquisition interventions. We extended current literature on the effects of errors of omission and commission of reinforcement on the acquisition of conditional discriminations. We used a translational research model to replicate and extend Hirst and DiGennaro Reed (2015) to investigate the effects of impaired treatment integrity with undergraduate students. We compared the efficacy and efficiency of instruction implemented with varying degrees of integrity in a parametric analysis using a randomized-control group design. We used a computer program, which erred on 0% to 50% of trials, to approximate procedures used to teach conditional discriminations in behavior analytic skill-acquisition interventions. The purpose was to identify a level of error at which most participants could still acquire the task. Greater than 80% of participants assigned to integrity levels at or above 85% acquired the skill; therefore, errors of reinforcement occurring on 15% or fewer trials did not hinder or slow acquisition for most participants. These results could inform future research with children diagnosed with autism spectrum disorder.

TABLE OF CONTENTS

LIST OF FIGURES	IV
LIST OF TABLES	V
LIST OF ABBREVIATIONS	VI
ACKNOWLEDGMENTS	VII
INTRODUCTION	1
METHOD	9
Participants and Setting	9
Materials	10
Dependent Variables and Response Measurement	11
Interobserver Agreement and Treatment Integrity	12
Experimental Design	12
Procedure	13
Data Analyses	15
RESULTS	18
DISCUSSION	23
REFERENCES	47
Appendix A: Participant Instructions and Screenshots of Microsoft® Visual Basic .Net	52
Appendix B: Random Assignment of Participants to Conditions	54
Appendix C: Consent Packet	55
Appendix D: Debriefing Form	59
Appendix E: Cumulative Records	60
Curriculum Vitae	67

LIST OF FIGURES

Figure 1. Efficacy: Proportion of participants at mastery by condition	34
Figure 2. Efficiency: Mean trials to criterion by condition	35
Figure 3. Median cumulative correct responses by condition	36
Figure 4. Exemplars of cumulative records 100% and 95% integrity	37
Figure 5. Exemplars of cumulative records 90% integrity	38
Figure 6. Exemplars of cumulative records 85% integrity	39
Figure 7. Exemplars of cumulative records 80% integrity	40
Figure 8. Exemplars of cumulative records 75% integrity	41
Figure 9. Exemplars of cumulative records 50% integrity	42
Figure 10. Proportion standardized AuC before and after phase change	43
Figure 11. Conditional probabilities of incorrect responses by condition	44

LIST OF TABLES

Table 1. Self-reported Demographic Data as a Percentage of the Sample	45
Table 2. Textual and Visual Stimuli Taught in AMTS Task	45
Table 3. Trials Implemented with Integrity and Errors across Conditions	46
Table 4. Mastery by Condition and Trials to Criterion	46

LIST OF ABBREVIATIONS

ABA	Applied behavior analysis
AMTS	Arbitrary match-to-sample
ANOVA	Analysis of variance
ASD	Autism spectrum disorder
AuC	Area under the curve
DTI	Discrete-trial instruction
EIBI	Early intensive behavioral intervention
ITI	Intertrial interval
Min	Minutes
s	Seconds
S+	Stimulus correct in array
S-	Stimulus incorrect in array

ACKNOWLEDGMENTS

I would like to thank Dr. Tiffany Kodak, my thoughtful advisor, for her mentoring throughout this project and my graduate career. I consider myself fortunate to be her student. I am sincerely grateful to have her as a teacher, mentor, and collaborator. I would be but a shadow of the behavior analyst, scientist, and scholar that I am without her generous provision of time and feedback. I can only hope to repay her by mentoring students much in the same way.

I offer my appreciation to my dissertation committee: Dr. Shawn Cahill, Dr. Christine Larson, Dr. Jeffrey Tiger, and Dr. Jason Vladescu for their time, attention, and contributions. I am especially thankful to Dr. Cahill for his assistance with statistical analyses and wisdom in group research and to Dr. Tiger for questions and suggestions that sparked much discussion.

Thank you to my colleague, collaborator, and friend Mike Harman for hours of help with statistical analyses and graphical displays. I am indebted to Mike for his patient teaching.

This project would not have been possible without the help of Dr. Jason Hirst. Thank you for adapting the computer program for my research and working within my timeline. Thank you also to Pat Reilly for his early contributions to the program's design and logistical wizardry.

I am grateful to my family and friends for their support and efforts to remain close. Notes of appreciation are due to Tracy Hess and Brittany Juban for their immeasurable support. Thank you to my parents Wes and Cindy and aunt Robbin for encouraging me to accomplish my goals. I am filled with gratitude for the support of my husband Jonathan which included cooking most of our meals, reminding me to take breaks, and bolstering my confidence when I wavered.

Last, but certainly not least, this project was supported by a Graduate Research Grant from the Organization for Autism Research and a Project Award in Behavior Analysis from John and Lynn Schiek.

INTRODUCTION

Establishing Quality Standards for Applied Behavior Analytic Skill-Acquisition Interventions: A Translational Model with Undergraduate Students

Autism spectrum disorder (ASD) is estimated to affect one in 68 individuals (Christenson et al., 2016). In 2015, national costs associated with ASD totaled \$268 billion with \$60,000 estimated per individual to provide interventions based on applied behavior analysis (ABA; Leigh & Du, 2015). Interventions based on ABA are the most commonly used evidence-based treatments to address behavioral excesses and deficits in individuals diagnosed with ASD (Wong et al., 2014). A recent review of available interventions for individuals with ASD identified ABA as an “established” intervention (National Autism Center, 2015) with proven effectiveness in ameliorating a variety of behavioral and academic needs (e.g., problem behavior, skill deficits, language development).

The effectiveness of ABA skill-acquisition interventions continues to gain empirical support (e.g., Eikeseth, Klintwall, Jahr, & Karlsson, 2012; Reichow & Wolery, 2009). Effective treatment involves not only an empirically-supported intervention based on principles of ABA but also accurate and consistent implementation (i.e., treatment integrity; Gresham, 1989) that is likely to produce meaningful outcomes for the client; however, providers may not always implement interventions with a high degree of treatment integrity in practice (Carroll, Kodak, & Fisher, 2013; Kodak, Cariveau, LeBlanc, Mahon, & Carroll, 2017). Poor treatment integrity could hinder or prevent effective interventions and limit internal validity (e.g., Carroll et al., 2013; Detrich, 2014; Holcombe, Wolery, & Snyder, 1994; Noell, Gresham, & Gansle, 2002). Interventions implemented with inadequate treatment integrity could interfere with a consumer’s (e.g., client) right to effective treatment (Van Houten et al., 1988), which Board Certified

Behavior Analysts are ethically required to provide (Standard 2.09a; Behavior Analyst Certification Board, 2014).

Perfect treatment integrity of behavioral skill-acquisition interventions across all conditions would be ideal, but it is likely unrealistic to expect a behavior analyst, teacher, or therapist to implement an ABA skill-acquisition procedure with perfect integrity across all contexts. Threats to treatment integrity include, but are not limited to, competing demands on time and resources, insufficient opportunities for initial and continued training, and diverse learner needs (Kodak, et al., 2017; Perepletchikova, Hilt, Chereji, & Kazdin, 2009). Descriptive behavior-analytic studies on the implementation of ABA-based skill-acquisition interventions by educators teaching children with ASD reported that errors of omission and errors of commission occur (Carroll et al., 2013; Kodak et al., 2017). Errors of omission involve failing to implement a component(s) of intervention. Errors of commission involve adding an extra component(s) to intervention. Withholding a small snack after a correct response is an example of an error of omission and providing a small snack after an incorrect response is an example of an error of commission. Errors are not necessarily mutually exclusive, and more than one type of error can occur simultaneously like when a therapist does not provide a prompt (i.e., omission of prompt) and provides a small snack following an incorrect response (i.e., commission of reinforcer).

In behavior-analytic descriptive studies on treatment integrity during skill-acquisition interventions, the type and degree of error varied based on the component of instruction, and educators implemented some components with less than 50% integrity (Carroll et al., 2013; Kodak et al., 2017). Errors in reinforcer delivery were observed; that is, teachers did not provide reinforcers following correct responses (i.e., error of omission) and provided reinforcers following an incorrect response (i.e., error of commission). Carroll and colleagues (2013)

reported that teachers provided a contingent tangible after correct responses on only 21% of opportunities; that is, they observed errors of omission on 79% of trials. Kodak and colleagues (2017) observed that teachers failed to withhold preferred items after incorrect or no responses on 43% of trials with unmastered tasks; that is, they observed errors of commission on 43% of trials. In an analysis of treatment integrity during discrete-trial instruction (DTI) implemented by novice therapists, Cook et al. (2015) reported errors in reinforcement as the most frequent.

The field of behavior analysis has begun to study how impaired treatment integrity may contribute to ineffective and/or inefficient interventions (Fryling, Wallace, & Yassine, 2012). For example, Carroll et al. (2013; Study 2) evaluated the effects of low-integrity instruction in which errors in reinforcement, vocal instruction, and controlling prompts were committed on 67% of trials (33% integrity) with six participants diagnosed with ASD who were learning tacts or play skills. Compared to high-integrity instruction, low-integrity instruction proved inefficient for one participant and inefficacious for the remaining five participants. Although errors occur in applied contexts, and some studies on the effects of impaired treatment integrity were conducted in these settings (e.g., Carroll et al., 2013; Holcombe et al., 1994; Noell et al., 2002), another option is to conduct these studies within a human operant laboratory (e.g., St. Peter Pipkin, Vollmer, & Sloman, 2010) and adopt a translational approach to evaluate the effects of integrity errors on learning (Mace & Critchfield, 2010).

Translational research involves investigating processes that are of applied interest in more highly-controlled laboratory settings and may include nonclinical populations and arbitrary tasks (Lerman, 2003; McIlvane et al., 2011). This arrangement allows researchers to study underlying behavioral principles of the phenomenon before applying similar independent variables to a clinically-relevant population in an environment where ABA academic and

behavioral intervention services are provided (e.g., St. Peter Pipkin et al., 2010). In translational research, an experimenter may explore multiple conditions, variables, or procedural variations with a population for whom integrity errors are unlikely to affect future learning (e.g., undergraduate students) before applying the independent variable to a population of interest (e.g., children with ASD). In that regard, using a translational approach to examine the effects of impaired treatment integrity is appealing because one may learn more about basic behavioral processes underlying treatment integrity errors that may slow or prevent skill acquisition without exposing a population-in-need to potentially harmful conditions. The effects of treatment integrity errors in translational settings can inform subsequent research studies and applied work with clinically-relevant populations. Thus, adopting a translational model could help reduce the possibility of potentially harmful and lengthier evaluations conducted with children with ASD. Once these phenomena have additional human operant data, the field can begin assessing external validity.

A series of translational studies by Hirst, DiGennaro Reed, and Reed (2013) and Hirst and DiGennaro Reed (2015) explored the effects of treatment integrity errors (specifically errors of omission and commission of reinforcement, called “inaccurate feedback” by the authors) on the acquisition of arbitrary match-to-sample (AMTS) tasks with undergraduate students and typically developing preschool children. Hirst and DiGennaro Reed manipulated the integrity with which a computer program provided feedback on the accuracy of responses, a presumed reinforcer, to 64 undergraduate students. When implemented with perfect integrity (i.e., 100% trials with integrity), the computer consistently presented “Correct” after all correct selections and presented “Incorrect” after all incorrect selections. Those exposed to perfect-integrity instruction met the mastery criterion (i.e., 15 consecutive correct responses) in the AMTS task.

They compared the learning of participants with perfect-integrity instruction to that of participants assigned to one of three levels of impaired integrity (i.e., 25%, 50%, and 75%). All participants, regardless of original integrity condition, were exposed to subsequent perfect-integrity instruction (i.e., A-B design). The focus of Hirst and DiGennaro Reed was whether participants who were previously exposed to errors would learn the task to mastery with subsequent perfect-integrity instruction. In other words, the authors analyzed their findings to determine whether prior exposure to integrity errors negatively affected later learning and performance. The authors reported that 88%, 75%, and 83% of participants in the 25%, 50%, and 75% integrity conditions, respectively, met the mastery criterion following the change to perfect-integrity instruction. They reported a relationship between higher integrity and more correct responses with statistically significant differences in median correct responses between low integrity (i.e., 25% and 50%) and high integrity (i.e., 75% and 100%). In addition, delays to acquisition following the phase change were evident for some participants in the 25% and 50% conditions with more trials required to reach the mastery criterion or failure to master the task.

Hirst and DiGennaro Reed (2015) emphasized acquisition after the phase change to perfect integrity which is a worthy avenue of research. This is akin to environments where integrity is improved given additional training and feedback (e.g., Cook et al., 2015; DiGennaro, Martens, & Kleinmann, 2007); however, these resources may be difficult to provide in environments with limited time, money, and/or personnel. Thus, in addition to evaluating if participants learn during subsequent perfect-integrity instruction, research may focus on whether it is possible to learn when integrity does not improve. This is akin to contexts where improvements in integrity may be less likely due to lack of resources, and learners may be exposed to impaired-integrity instruction for some time. With close inspection of the cumulative

records and summary data presented by Hirst and DiGennaro Reed, it is possible to determine how many participants assigned to impaired integrity (i.e., 25%, 50%, and 75% integrity) met the mastery criterion prior to the change to perfect integrity thereby evaluating whether instruction with impaired integrity was efficacious. None of the participants ($n = 32$) in the 25% and 50% integrity groups met the mastery criterion with impaired integrity. Only one-quarter of the participants in the 75% integrity group ($n = 16$) met the mastery criterion prior to the phase change. Therefore, one could conclude that instruction with 75% of trials implemented with integrity was inefficacious for most participants.

Many of the treatment integrity evaluations conducted in translational and applied studies have included impaired-integrity values between 33% and 67% (e.g., Carroll et al., 2013; Jenkins, Hirst, & DiGennaro Reed, 2015). That is, previous studies have evaluated instruction when 33% to 67% of trials were implemented with integrity. These studies reported detrimental effects of learning at all levels. However, the effect of higher levels of impaired-integrity instruction has received less attention in the extant literature. Exceptions are the studies by Hirst et al. (2013) and Hirst and DiGennaro Reed (2015) which included 75% integrity. In addition to these studies, Bergmann, Kodak, and LeBlanc (2017) also investigated smaller decrements to integrity. Bergmann et al. evaluated the effects of isolated errors of omission of reinforcement (i.e., not providing praise and token after correct response) and errors of commission of reinforcement (i.e., providing praise and token after incorrect response) on the acquisition of an AMTS task with two typically developing children. The authors found that impaired-integrity values between 70% and 82% either hindered or prevented learning. In Experiment 1, 18% of trials with errors of omission or commission (i.e., 82% integrity) resulted in double the sessions required to produce learning compared to perfect-integrity instruction for one participant, and

17% of trials with errors of omission (i.e., 83% integrity) required double the sessions to produce learning for the other participant. In Experiment 2, errors of omission or commission occurring on 21% to 22% of trials (i.e., 79% and 78% integrity) slowed acquisition to the point that one participant required seven times the number of sessions compared to perfect-integrity instruction, and errors of omission on 20% of trials (i.e., 80% integrity) and errors of commission on 30% of trials (i.e., 70% integrity) prevented the other participant from acquiring the AMTS task altogether. Bergmann et al. programmed errors of omission and commission of reinforcement in isolation. That is, depending on the condition, the participants were exposed to either errors of omission or errors of commission. The ability to make an error on a given trial was dependent on whether the participant responded correctly or incorrectly. This limited the parametric analysis because the authors were unable to make errors equally across participants and conditions. The design in Hirst and DiGennaro Reed addressed these issues by making combined errors of omission and commission allowing for obtained errors to match programmed levels of error in their parametric analysis.

Taken together, the results of Hirst and DiGennaro Reed (2015) and Bergmann et al. (2017) suggest that the efficacy and/or efficiency of instruction can be affected when errors occur on roughly 25% of trials (i.e., 75% integrity). Understanding the implications of integrity at or above 75% is especially pertinent because 80% integrity may be considered “acceptable” in some educational and intervention contexts (Cook et al., 2015). Acceptable integrity should not result in detrimental effects on learning. Given demands in instructional settings (e.g., lack of resources like time, insufficient training, novel learner behaviors), acceptable integrity likely needs to include some room for error; nevertheless, the field does not currently have a minimum level of acceptable integrity. A minimum could help determine whether interventions are

implemented with sufficient integrity likely to produce clinically-meaningful outcomes for clients. That is, if data were collected and revealed that intervention was implemented with integrity below a standard, poor clinical outcomes may be linked to integrity rather than an inefficacious intervention. This would suggest different courses of action (e.g., train staff on components of intervention rather than discontinue intervention) to improve outcomes for the learner. Additional research is warranted to investigate how, and to what degree, different types of integrity errors (e.g., reinforcement, prompting, instruction) affect efficacy and efficiency of skill-acquisition interventions.

The purpose of the current study was to replicate and extend Hirst and DiGennaro Reed (2015). We made several changes to the computer program written in Microsoft® Visual Basic .Net developed by Hirst and DiGennaro Reed (2015) to more closely approximate how AMTS is taught in ABA skill-acquisition programs (e.g., Green, 2001; Grow, & LeBlanc, 2013). We modified the computer program in the following ways: (a) providing praise statements (e.g., Way to go!) rather than “correct” or “incorrect”, (b) contriving a potential token economy for a putative reinforcer (i.e., points and gift card), (c) including a brief intertrial interval (ITI), and (d) removing distracters from response options. To focus on efficacy of instruction with impaired integrity, we deviated from the method of staggering the introduction of perfect integrity within and across conditions (Hirst & DiGennaro Reed, 2015) by including 300 trials in the integrity comparison prior to introducing 200 trials with perfect integrity. We also sought to strengthen experimental control by using random assignment rather than group sessions with all participants assigned to one condition.

In addition to the procedural modifications described above, we extended Hirst and DiGennaro Reed (2015) by conducting a parametric analysis of additional values of impaired

integrity that were not included in previous evaluations. Specifically, we included additional values of impaired integrity above 75% and below 100%. The goal of this study was to identify the point at which instruction remained efficacious for most participants (i.e., 80% of participants met the mastery criterion) yet allowed for some errors to be made. As in Bergmann et al. (2017) and Hirst and DiGennaro Reed (2015), we focused on integrity of reinforcement. Errors of omission and commission of reinforcement were observed in descriptive studies (Carroll et al., 2013; Cook et al., 2015; Kodak et al., 2017), and the process of reinforcement is essential to increase the frequency of new responses. To obtain specific proportions of errors, we evaluated the effects of errors of omission and commission of reinforcement (i.e., combined errors) because percent of error was not dependent on the participants' behavior (e.g., Bergmann et al., 2017). We adopted a translational model to evaluate the effects of treatment integrity errors on efficacy and efficiency of instruction. That is, all analyses were conducted with undergraduate students learning an arbitrary task in a highly-controlled laboratory setting. A translational approach was used to manipulate treatment integrity with participants for whom potential detrimental effects on learning were less of a concern rather than evaluating the effects of the independent variable on the learning of a clinically-relevant population like children diagnosed with ASD who are receiving ABA services.

METHOD

Participants and Setting

In this study, 168 undergraduate students (120 women; see Table 1 for demographics) aged 18 to 66 years ($M = 21.6$) and enrolled in psychology courses at a public university in the Midwest served as participants. They were provided 1.5 hours of extra credit in compensation. In addition to extra credit, roughly one out of every two participants received a \$10.00 Target® gift

card based on earning the highest number of points in his/her randomly-assigned condition. We did not exclude participants based on any of their answers on a university-wide prescreen survey, but eligible participants reported no visual impairment affecting their ability to read on a computer. For data analysis, we planned to exclude participants with an average response latency of less than 0.5 s and those who failed to complete a minimum of 500 trials within one hour (Hirst & DiGennaro Reed, 2015). These criteria were designed to control for the estimated minimum time participants needed to attend to the visual and textual stimuli on the screen and to equate the number of trials to which participants were exposed, respectively. However, no participants in our sample met either criterion. All sessions were conducted in a university computer lab (6.7 m by 10.4 m) with 30 Dell® touch-screen desktop computers.

Materials

Black cardboard dividers (157 cm by 56 cm) were placed around computers to obstruct views of other computer screens. Printed materials (i.e., consent packet, debriefing form) were distributed to participants. A computer program written in Microsoft® Visual Basic .Net presented all components of the AMTS task. Stimuli were presented on a dark gray background with light gray response buttons that could be clicked with the computer mouse, and the program was designed to occupy the entire computer screen. The AMTS task included five black-and-white Japanese hiragana, a phonetic lettering system that was modified to prevent future difficulties with learning the language, and five nonsense words (see Table 2; Hirst et al., 2013; Hirst & DiGennaro Reed, 2015). Each visual stimulus was paired with a textual stimulus. The correct textual stimulus in each trial (i.e., S+) was determined by the visual stimulus presented as the sample. The four other textual stimuli were incorrect on that trial (i.e., S-); however, each stimulus was both an S+ and an S- throughout the task (i.e., conditional discrimination).

At the onset of a trial, the computer program presented one black-and-white visual stimulus (2.5 cm by 2.5 cm) on the left side of the dark gray background (Appendix A). Five response options displayed in dark gray text atop 6.4 cm by 1.3 cm light gray rectangular boxes were aligned to the right of the figure and appeared simultaneously. Green text boxes (6.4 cm by 1.3 cm) with three rotating praise statements (e.g., You did it!) appeared under the visual stimulus following correct responses and incorrect responses with a programmed error of commission and remained for 1.5 s. The visual and textual stimuli were removed for the 0.5-s ITI until the onset of the next trial. No text box appeared following incorrect responses nor correct responses with programmed errors of omission; instead, the visual and textual stimuli were removed for the entire 2-s ITI until the onset of the next trial. A white text box located in the top-center of the computer screen displayed points earned and was visible to the participant throughout the study. The experimenter recorded points earned on a clipboard at the end of the session to determine to whom gift cards were distributed.

Dependent Variables and Response Measurement

The main dependent variable was the *efficacy* of instruction determined by whether a participant reached the mastery criterion. The *mastery criterion* was defined as 15 consecutive correct responses. *Correct responses* were defined as selecting the textual stimulus that corresponded with the visual stimulus presented as the sample on that trial; in other words, a correct response was clicking on the S+ rather than any of the S- comparison stimuli in the array. *Incorrect responses* were defined as selecting any textual stimulus other than the S+ on a trial. In addition to efficacy, we also compared the *efficiency* of instruction. Efficiency was determined by comparing the number of trials required to reach the mastery criterion across conditions and participants. Efficiency was only compared for participants for whom instruction was

efficacious. Stimuli were ordered using random rotation without replacement whereby a trial of each sample stimulus was presented once before repeating a sample stimulus. Therefore, to meet the mastery criterion, participants needed to select the S+ in the presence of each sample three consecutive times.

Interobserver Agreement and Treatment Integrity

The Microsoft® Visual Basic .Net computer program administered all aspects of the experiment and collected data for all dependent measures. The computer program was thoroughly tested to ensure accurate data collection and consequence presentation as specified in the procedure below. Treatment integrity was evaluated by comparing the obtained percentage of errors to the programmed percentage of errors. That is, we determined whether the computer program implemented the correct number of errors per the integrity condition. There were no differences between obtained and programmed integrity; in other words, the computer program's obtained percentage of integrity matched the programmed integrity level 100% of the time.

Experimental Design

A randomized-control between-groups design was used to compare the effects of errors of omission and commission on the acquisition of an AMTS task in the parametric analysis. In addition, data from participants assigned to impaired-integrity conditions provided an opportunity for within-subject analyses, because acquisition under impaired- and perfect-integrity conditions could be compared (i.e., A-B design).

Participants were randomly assigned (Appendix B) to an integrity condition. To randomly assign participants, the experimenter copied the list of students who signed up for a research session through an online research database. Then, the experimenter used a list of random numbers that were obtained from an online random number generator

(www.randomizer.org) to order the participants' names and assign them participant numbers and computers. Next, the experimenter generated a list of the conditions in a random order using the same random number generator. She then assigned these conditions to each participant number. Finally, the experimenter setup the program on each computer per these randomized conditions.

Procedure

Each participant was given a consent packet (Appendix C) at their computer station, and the experimenter reviewed the consent packet with the group. The participants had the opportunity to ask questions, sign the consent form, and withdraw their participation at any time. If participants chose to withdraw from the study, they received extra credit compensation in accordance with the duration of participation. For example, if a participant withdrew after 30 min, the experimenter assigned 0.5-hour extra credit to their account on the research database. No participants withdrew from the study.

After consenting, the participants selected "Begin" on their computer screens. They answered a few short demographic questions regarding their age, ethnicity, and disability status (Hirst & DiGennaro Reed, 2015). Next, participants read written instructions (Appendix A). Participants were informed that they would earn points for correct responses, these points were displayed on a point counter near the top of their screens, and those with the highest values of points would qualify for a \$10.00 Target® gift card at the end of the session. This statement was used to potentially contrive value for the points. Participants were not informed that the computer program would make errors in the provision of these points. Deception was used to mirror integrity errors made during instruction with children; for example, a teacher would be unlikely to tell a student that s/he engaged in an error while teaching. The use of deception was

approved by the Human Subjects Institutional Review Board. Finally, participants clicked a button that indicated they understood the instructions and were ready to start the experiment.

The participants completed a total of 500 trials in one hour. When participants completed the 500 trials, the computer program stopped, and they were told to wait for further instructions from the experimenter. Following the completion of the AMTS task by all participants, the experimenter distributed gift cards to the participants who earned the highest number of points within their randomly assigned condition, read the debriefing form aloud (Appendix D), provided an opportunity for participants to ask questions, and excused the participants.

Perfect-integrity control condition. Twenty-four participants were randomly assigned to the control condition. Each trial began with the presentation of a visual sample stimulus on the left and five textual response options directly to its right. The visual stimulus varied from trial-to-trial, according to random replacement without rotation, and the location of the S+ in the response array varied according to an algorithm written in the program's code. The computer program presented differential consequences based on whether the participant's response was correct or incorrect. Following all correct responses, the computer provided a point and a praise statement (e.g., Way to go!) in a green text box directly below the visual stimulus for 1.5 s with a brief 0.5-s ITI of a gray screen without visual or textual stimuli. Following all incorrect responses, the computer did not provide a point nor written praise statement, and it proceeded directly to a 2-s ITI with a gray screen. Participants assigned to perfect integrity experienced 500 perfectly-implemented trials.

Impaired-integrity conditions. The remaining 144 participants ($n = 24$) were assigned to one of six impaired-integrity conditions (see Table 3 for conditions). In the integrity comparison phase, these participants were exposed to a proportion of trials (Table 3) as

described in the perfect-integrity condition above. However, on the remaining trials (Table 3), the computer program made either an error of omission or commission depending on the participant's response. If the participant responded correctly on a programmed error trial, the computer program made an error of omission in which it moved directly to a 2-s ITI without providing praise nor a point. Conversely, if the participant responded incorrectly on a programmed error trial, the computer made an error of commission in which it provided a point and a praise statement for 1.5 s before moving onto a 0.5-s ITI.

Integrity errors were programmed to occur during a certain proportion of trials in the first 300 trials of the AMTS task (i.e., integrity comparison); this value is based on the maximum trial duration used by Hirst and DiGennaro Reed (2015). The computer divided the 300-trial integrity comparison into 20-trial sessions to control the percentage of errors and distribute errors evenly across sample stimuli. Programmed error trials were designed to occur across all five stimuli equally and were distributed across trials. Subsequently, all participants were exposed to 200 trials implemented with perfect integrity. That is, after 300 trials with a proportion of errors of omission or commission (e.g., 50% integrity/50% errors), the participants experienced 200 perfectly-implemented trials (i.e., 100% integrity; Table 3).

Data Analyses

To evaluate efficacy, we calculated the proportion of participants in each condition that reached the mastery criterion. These data were computed by dividing the number of participants who met the mastery criterion during the integrity comparison by the total number of participants randomly assigned to the condition ($n = 24$) and multiplying the quotient by 100 to obtain a percentage. We used a chi-square (χ^2) test of independence to analyze whether integrity condition was related to mastery of the AMTS task in the integrity comparison phase. That is, as a measure

of internal validity, we examined if integrity condition and achieving or failing the mastery criterion, in the integrity comparison phase, were independent of one another. We also computed the total number of participants, by condition, who achieved the mastery criterion when exposed to perfect-integrity instruction following the phase change as well as the number of participants who failed to master the AMTS task.

If participants met the mastery criterion at any point in the experiment, we compared the number of trials required to master as a measure of efficiency. Conditions averaging fewer trials to mastery were considered more efficient. We conducted a nonparametric analysis of variance (ANOVA; Kruskal-Wallis) on the number of trials to criterion by condition. A nonparametric statistic was used because we compared efficiency for only those participants for which instruction was also efficacious; therefore, we had unequal cell sizes and non-homogeneous variance. The ANOVA revealed whether there were statistically significant differences in trials to criterion under impaired-integrity and perfect-integrity conditions. When statistically significant differences were found, we conducted multiple comparison posttests (Mann-Whitney) with a Bonferroni correction procedure to reduce the probability of Type I error (i.e., false positives; adj. alpha, $p = .002$). These posttests were used to identify between which conditions (e.g., 75% integrity, 100% integrity) differences were statistically significant.

Visual inspection of cumulative records of correct responding during the integrity comparison and subsequent perfect integrity was completed, and patterns in responding were identified. To supplement visual analysis of cumulative records, we computed the area under the curve (AuC) before and after the phase change (Hirst et al., 2013; Hirst & DiGennaro Reed, 2015). We used the trapezoidal method (Myerson, Green, & Warusawitharana, 2001) to estimate the area under each participant's acquisition curve (i.e., cumulative correct responses) in the

integrity comparison and under perfect integrity. The trapezoidal method is used in behavioral economics to estimate differences between groups for data displayed in graphical formats and does not require curve fitting (Hirst & DiGennaro Reed, 2015). To calculate AuC, the area of each trapezoid was computed and then the areas were summed: $AuC = \sum(x_2 - x_1)[(y_1 + y_2)/2]$. We reset the floor to zero for each participant on trial 301 after the phase change to perfect-integrity instruction. To standardize this measure and allow a comparison to the results reported by Hirst and DiGennaro Reed (2015), proportional AuC was calculated by dividing the AuC for each participant in each condition prior to and following the phase change by the AuC of cumulative correct responses of perfect, hypothetical acquisition (slope = 1). A nonparametric ANOVA (Kruskal-Wallis) was conducted to compare proportional AuC across conditions. A nonparametric statistic was used because of a non-normal distribution which violated one of the assumptions of parametric analyses. We conducted post-hoc comparisons with corrections as described above.

Errors in reinforcement delivery could increase the likelihood that conditional discriminations come under faulty sources of stimulus control. For example, in the presence of a particular sample stimulus, selecting an S- was reinforced during an error of commission. Thereafter, participants may continue to select the same S- in the presence of that sample. Alternatively, selecting an S+ could contact extinction during an error of omission. One might also see the development of stimulus biases wherein participants continue to select a particular comparison stimulus regardless of the sample. To evaluate whether faulty stimulus control occurred and impacted mastery, we conducted conditional probability analyses. We identified the S- selected in the presence of each sample stimulus to evaluate whether incorrect responses were allocated to any particular stimulus during the integrity comparison. That is, we divided the

proportion of incorrect responses that participants allocated to each comparison stimulus in the presence of each sample across trials in the integrity comparison (i.e., 60 trials of each sample in 300 trials) to look for values above chance (i.e., 0.25 in array of four S-).

RESULTS

We evaluated the efficacy of each integrity condition included in the parametric analysis. Figure 1 displays the percentage of participants who mastered during the integrity comparison, those who mastered following the phase change to perfect integrity, and participants who never mastered the task. We found that as integrity decreased so did the proportion of participants who mastered the conditional discriminations during the integrity comparison. The efficacy comparison was supported by a chi-square test of independence. We compared the observed frequencies of participants (Table 4) who met the mastery criterion in the integrity comparison and the observed frequencies of those who failed to meet the mastery criterion to expected frequencies (i.e., chance levels based on the parent distribution). These observed frequencies differed significantly from expected frequencies ($\chi^2(6) = 96.13; p < .001$) suggesting these differences were unlikely due to a variable other than integrity condition. Thus, these analyses support internal validity that integrity affected efficacy of instruction. In the parametric analysis, the lowest impaired-integrity level that resulted in at least 80% of participants mastering the task was 85% integrity.

To compare relative efficiency of instruction at different levels of integrity, we computed the number of trials to mastery for each participant by condition (Figure 2; Table 4). We examined whether differences in efficiency, or trials to criterion, were statistically significant using a Kruskal-Wallis nonparametric analysis. We only compared efficiency for participants for whom instruction was efficacious; that is, we included trial data for participants who mastered

the task within 500 trials. The omnibus statistic was significant ($\chi^2(6) = 79.32; p < .001$) and suggested differences in efficiency. We conducted Mann-Whitney post-hoc comparisons with a Bonferroni Correction procedure and adjusted alpha ($p = 0.002$) to identify between which groups there were statistically significant differences. These posttests found significant differences ($p < .001$) between the (a) 50% condition and conditions with at least 80% integrity, (b) 75% condition and conditions with at least 80% integrity, and (c) 80% integrity and conditions with at least 90% integrity. Of note, statistically significant differences were not found between 85% integrity and all conditions with at least 80% integrity. The lack of effect suggested that differences in relative efficiency were not observed once integrity reached 85%.

To look for trends in acquisition curves, we calculated the median cumulative correct responses for participants in each condition (Figure 3). Participants in the 50% and 75% conditions engaged in fewer correct responses across 500 trials compared to higher levels of integrity. Acquisition of participants in the 80% condition was slower than conditions with at least 85% integrity. The distance between data paths and similar slope of conditions with at least 85% integrity showed that acquisition was similar for participants across these conditions. Therefore, visual inspection of median cumulative records suggested that acquisition under 85% integrity was akin to acquisition under higher integrity. In addition to group data, each participant's cumulative correct responses were graphed, and we inspected cumulative records for slope, changes in slope, and mastery. We sorted cumulative records by condition according to slope and whether participants mastered the conditional discriminations before the phase change, after the phase change, or never. We categorized each participant's cumulative record of correct responding based on these attributes. Figures 4 to 9 are representative cumulative records for each condition (see Appendix E for all records); summaries of participant acquisition are below.

Cumulative records of correct responses for the 24 participants in the 100% integrity condition are in Figure 4 (upper panel). These participants acquisition curves fit two patterns with most participants' correct responding increasing in a manner like perfect hypothetical acquisition (Figure 4; left upper). All participants in the 100% control condition met the mastery criterion in an average of 99 trials (range, 35-257; Table 4).

Cumulative records of correct responding for the 24 participants randomly assigned to the 95% integrity are in Figure 4 (lower panel). Participants' acquisition curves fit into two categories with most participants acquiring the conditional discriminations in a manner similar to 100% integrity instruction (Figure 4; left lower). Twenty-three (96%) participants met the mastery criterion during the integrity comparison. One participant did not master the AMTS task after 500 trials (Figure 4; right lower). Participants in the 95% condition had a mean of 117 trials to mastery (range, 38-282; Table 4).

Cumulative records for the 24 participants assigned to the 90% integrity condition are in Figure 5. These participants' acquisition curves fit into three patterns. Of the 24 participants randomly assigned to this condition, 21 (88%) met the mastery criterion in the integrity comparison (Figure 5; upper left). Of the three participants who did not master the AMTS task in the integrity comparison, one mastered in perfect integrity (Figure 5; upper right) and two never mastered the task (Figure 5; lower). Participants in this condition averaged 119 trials to mastery (range, 41-315; Table 4).

Cumulative records for the 24 participants in the 85% integrity condition are shown in Figure 6. These participants' acquisition curves fit into three patterns. Twenty (83%) participants mastered the task in the integrity comparison (Figure 6; upper left). Two of the remaining four participants met the mastery criterion in perfect integrity (Figure 6; upper right), and two

participants did not meet the mastery criterion after 500 trials (Figure 6; lower). The mean number of trials to mastery for participants randomly assigned to this condition was 150 (range, 46-373; Table 4).

Cumulative records for the 24 participants in the 80% integrity condition are in Figure 7. Their data fit into three patterns of responding. Fifteen (63%) participants met the mastery criterion in the integrity comparison (Figure 7; upper left). Of the nine participants who did not meet mastery in the integrity comparison, eight mastered the AMTS task in perfect integrity (Figure 7; upper right) and one never mastered (Figure 7; lower). The average number of trials to mastery for this condition was 227 (range, 22-420; Table 4).

Cumulative records of correct responses for the 24 participants in the 75% integrity condition are in Figure 8. These participants' data were consistent with three patterns. Five of the 24 participants (20.8%) assigned to this condition met the mastery criterion in the integrity comparison (Figure 8; upper left). Of the remaining 19 participants, 14 participants met the mastery criterion in perfect integrity (Figure 8; upper right) and five did not master the task (Figure 8; lower). The mean number of trials to mastery for participants in this condition was 359 (range, 153-469; Table 4).

Cumulative records of correct responses for the 24 participants assigned to the 50% integrity are in Figure 9. These graphs showed shallow-sloped learning curves compared to other conditions. None of the participants in the 50% integrity condition met the mastery criterion in the integrity comparison; 11 participants met the mastery criterion following the phase change perfect integrity (Figure 9; left). Thirteen participants assigned to this condition never mastered the AMTS task (Figure 9; right). The average number of trials to mastery for participants in this condition was 387 (range, 342-486; Table 4).

Visual analysis of cumulative records was supplemented by calculating the proportional AuC for each participant before (i.e., trial 1-300) and after (i.e., trial 301-500) the phase change (Figure 10). These values quantified the amount of learning that occurred before and after the phase change expressed as a proportion of learning one would expect if each trial was a correct response. The proportional AuC values for each condition for both phases were analyzed using a Kruskal-Wallis ANOVA which revealed statistically significant differences between groups in both phases of the study (integrity comparison: $\chi^2(6) = 96.99, p < .001$; perfect integrity: $\chi^2(6) = 78.52, p < .001$). We conducted exhaustive post-hoc analyses with multiple comparison posttests (Mann-Whitney) with a Bonferroni correction procedure (adj. alpha, $p = .002$). These posttests revealed statistically significant differences ($p < .001$) in the integrity comparison between (a) 50% and conditions with at least 75% integrity, (b) 75% and conditions with at least 80% integrity, and (c) 80% and conditions with at least 95% integrity. In perfect integrity, differences in proportional AuC were statistically significant ($p < .001$) between (a) 50% and conditions with at least 80% integrity and (b) 75% and conditions with at least 80% integrity. Differences between 80% and 90% ($p = .002$), in the integrity comparison, and 80% and 100% ($p = .004$), in perfect integrity, were approaching significance with the correction procedure. Of note, no statistically significant differences were found between 85% and conditions with at least 90% integrity in either phase.

Conditional probabilities are shown in Figure 11. We calculated the probability that a participant would select each incorrect stimulus (S-) in the array in the presence of each sample stimulus throughout the integrity comparison (i.e., 60 trials of each sample). These calculations were a measure of discrimination strength and may provide evidence for acquisition of incorrect discriminations or faulty stimulus control. Across conditions, participants generally responded to

an S- at or below chance level (i.e., 0.25 given four S- in array) when they engaged in incorrect responses in the presence of the sample stimuli. Few participants engaged in incorrect selection responses above chance level in the conditions with at least 85% integrity. However, more participants engaged in incorrect responses to an S- in the presence of a sample that exceeded chance level responding as indicated by more data points above chance level responding, in conditions with 80% or less integrity. These data suggest that acquisition of faulty stimulus control like stimulus biases and/or incorrect discriminations were more likely in lower integrity conditions.

DISCUSSION

We evaluated whether efficacy and efficiency of instruction would be compromised with different levels of impaired integrity in a parametric analysis which focused on integrity between 75% and 95% using a translational approach with undergraduate student participants. Programmed integrity errors in the form of omission of reinforcement and commission of reinforcement affected the efficacy and efficiency of instruction. The results of the study suggested a relationship between integrity of the procedure and acquisition of the AMTS task. That is, as integrity increased, more participants met the mastery criterion which was our operational definition of efficacy. Our data also showed a relationship between the efficiency of instruction, defined as the number of trials to mastery, and the integrity of instruction. Higher levels of integrity needed fewer trials to reach mastery. Overall, instruction implemented with higher levels of integrity was more likely to be efficacious and efficient; this outcome replicates previous parametric analyses of programmed treatment integrity errors (e.g., Bergmann et al., 2017; Carroll et al., 2013; Hirst & DiGennaro Reed, 2015; Noell et al., 2002).

We sought to identify the lowest level of impaired integrity at which most participants (i.e., at least 80%) met the mastery criterion, and the 85% integrity condition produced these outcomes. In addition, there were no statistically significant differences between trials to criterion for participants in the 85% integrity condition and trials to mastery for participants assigned to higher levels of integrity. Cumulative records for participants assigned to the 85% integrity condition showed little discrepancy between participants' acquisition in 85% to 100% integrity conditions. Statistical analyses of proportional AuC did not find statistically significant differences between 85% and conditions with greater integrity in the integrity comparison nor subsequent perfect integrity. In other words, participants' rate of acquisition in the 85% integrity condition was not discrepant from participants' acquisition in conditions with 90% or greater integrity. These data suggest that receiving instruction with 85% integrity was as efficacious and efficient as instruction with perfect integrity. Thus, it may be that learning under conditions with 85% of trials implemented correctly is akin to learning under conditions with 100% of trials implemented correctly.

Compared to the participants in the perfect-integrity control condition, the participants in the 85% condition required, on average, 50 additional trials to master these five conditional discriminations. This difference was not statistically significant suggesting 85% integrity is as efficient as perfect-integrity instruction. Nevertheless, differences that are not statistically significant may be clinically significant. Considering that many individuals receiving behavior analytic skill-acquisition interventions need to learn far more than five conditional discriminations, this 50-trial difference could become clinically meaningful if extrapolated to include more targets and skills. Eighty-five percent integrity may be a level of impaired integrity at which efficacy and efficiency of intervention are unlikely to be compromised; however,

additional analyses need to be conducted before a minimum or “acceptable” level can be identified.

Our data replicated the outcomes of Hirst and DiGennaro Reed (2015) in that all the participants assigned to the 100% integrity condition were able to reach the mastery criterion, about a quarter of participants assigned to the 75% integrity condition met mastery prior to the phase change, and none of the participants assigned to the 50% integrity condition met mastery. However, we did not find a statistically significant difference after the phase change to subsequent perfect integrity between 50% and 75% integrity, whereas Hirst and DiGennaro Reed reported a statistically significant difference between these groups. However, it is unclear whether the posttests they employed included a correction procedure and adjusted alpha to reduce Type I error.

These data add to the extant literature on the effects of treatment integrity errors on skill acquisition in several ways. This study employed a randomized-control group design to evaluate the effects of treatment integrity errors on efficacy and efficiency of instruction. We used a randomized-control group design because we were interested in identifying overall trends to add to single-subject data with more idiosyncratic findings across participants (e.g., Bergmann et al., 2017; Carroll et al., 2013). With these group data, we identified 85% integrity as a potential level of impaired integrity that may not result in delayed acquisition or inefficacious intervention for at least 80% of learners. The only other published study, to our knowledge, that examined the effects of treatment integrity with 20% or fewer trials with errors (Bergmann et al., 2017) reported that instruction was less efficient but still efficacious when errors in reinforcer delivery occurred on approximately 18% of trials (i.e., 82% integrity) with two typically developing participants. The efficacy data in the current study support the findings of Bergmann et al. in that

impaired-integrity values above 80% are likely to still be efficacious. Together, these data lend support for an “acceptable” level of integrity above 80% (Cook et al., 2015). These data should be interpreted with caution, however, given that both studies only evaluated errors in reinforcer delivery and utilized translational research methods including non-clinical populations, arbitrary tasks, and trial-and-error instruction.

The 85% integrity condition led to at least 80% of participants meeting the mastery criterion in the integrity comparison. We used 80% to represent the “majority” of participants to make decisions regarding efficacy of instruction with different levels of integrity. This value was selected based on standards for efficacy often used in general education settings (e.g., Detrich, 2014) whereby instruction is deemed efficacious if 80% of a classroom learns a skill via instruction and integrity is assumed acceptable. Given the individualized nature of ABA intervention, it is unclear what might be considered a sufficient proportion of individuals who benefit from instruction to judge efficacy. It could be that behavior analysts require effective behavior change for 100% of individuals receiving applied behavior analytic interventions; therefore, 80% may not be sufficient to judge efficacy. In the current study, the only condition which resulted in mastery for all participants was 100% integrity. Applying this standard suggests that 100% integrity instruction would be the only acceptable level of integrity. As discussed previously, 100% integrity, may not be realistic across all situations and settings (e.g., Carroll et al., 2013; Kodak et al., 2017). More research and discussion are needed to identify a standard for efficacy that considers potential treatment integrity errors that are likely to occur, at least to some degree, in naturalistic contexts which can then be applied in future studies.

The primary focus of this study was whether impaired-integrity instruction could be efficacious and efficient. We found that instruction implemented with at least 85% integrity

could be efficacious and efficient for most participants. Our data also permitted analysis of potential carry-over effects of previous impaired-integrity instruction when integrity was improved. We compared the proportion of participants who never met mastery and proportional AuC across conditions. The proportion of participants unable to acquire the conditional discriminations during or after the integrity comparison increased as integrity decreased. Specifically, 54% and 21% of participants in the 50% and 75% integrity conditions, respectively, never mastered whereas a maximum of 8% of participants never mastered in the conditions with at least 80% integrity. The proportional AuC values for participants in the 50% and 75% integrity conditions were not statistically significant from one another; however, these differences were statistically significant when proportional AuC for 50% and 75% integrity was compared with conditions with at least 80% integrity. These data suggest that differences in acquisition remained for some participants assigned to 50% and 75% integrity, and there may have been carry-over effects hindering acquisition under subsequent perfect-integrity conditions. In other words, decrements to acquisition remained despite improved integrity conditions for many participants. Conditional probabilities calculated during the integrity comparison point to one potential explanation for these carry-over effects; more participants assigned to 50% and 75% integrity may have learned discriminations under faulty sources of stimulus control compared to participants assigned to higher integrity.

These potential carry-over effects replicated Hirst and DiGennaro Reed (2015) and Jenkins et al. (2015) wherein some participants failed to acquire the target skills under perfect-integrity instruction following previous exposure to impaired-integrity instruction. These findings run counter to the results of Bergmann et al. (2017) and Carroll et al. (2013) wherein participants' acquisition under subsequent high-integrity instruction showed little or no delay or

impairment. Bergmann et al. and Carroll et al. (Study 3) manipulated treatment integrity errors in isolation (e.g., only omission of reinforcement) whereas the current study, Hirst and DiGennaro Reed, and Jenkins et al. manipulated combined errors in reinforcer delivery. Thus, further research is needed to elucidate the effects of isolated and combined errors on acquisition, and descriptive research could seek to report the occurrence of isolated and combined errors in various instructional arrangements. Future research could examine whether different types, degrees, or duration of exposure to integrity errors lead to continued impairments in acquisition and how to modify instruction to improve outcomes.

To replicate Hirst and DiGennaro Reed (2015), we exposed all participants to a total of 500 trials in the study. We kept the number of trials in the study constant while the integrity varied across conditions. The integrity-comparison phase was 300 trials and the perfect-integrity phase was 200 trials. Therefore, participants assigned to impaired-integrity conditions experienced fewer trials overall with perfect integrity compared to the perfect-integrity control group. For example, participants assigned to 50% integrity experienced 150 perfect-integrity trials in the first 300 trials of instruction (i.e., 150 is 50% of 300). After the phase change, they received 200 trials of instruction conducted with perfect integrity. Thus, in total, participants in the 50% integrity condition received 350 trials (150 in the integrity comparison phase and 200 in the perfect integrity phase) implemented with perfect integrity, which is 150 fewer perfectly-implemented trials than participants in the 100% integrity condition. By keeping the total number of trials constant across participants, participants in impaired-integrity conditions experienced fewer than 500 trials with perfect integrity. It could be that all participants, regardless of the impaired-integrity condition in the integrity comparison, would have mastered the AMTS task had they been exposed to 500 perfect-integrity trials. For example, we do not know whether

participants randomly assigned to the 50% integrity condition would have eventually met mastery if exposed to an additional 150 trials of perfect integrity (i.e., 150 implemented with poor integrity, 500 implemented with perfect integrity). However, it should be noted that all the participants in the 100% integrity condition met the mastery criterion in fewer than 350 trials. Nevertheless, it could be that the proportion of error is less important for efficacy than the number of trials implemented with integrity. To make that comparison in future research, one would need to compare acquisition under conditions where integrity differs, but the total number of trials implemented with integrity is the same. For example, 50% integrity could have 150 trials with integrity and 150 trials with errors in the integrity comparison and 350 correctly-implemented trials in the perfect-integrity phase; therefore, these participants would experience the same number of trials implemented with integrity as those assigned to 100% integrity.

The current study included several limitations that should be addressed in future research. Our primary design was a randomized-control design to examine between-subject differences at the group level; however, we also employed a single-subject design. We used an A-B design, the weakest single-subject design that still allows for some level of experimental control; however, internal validity of the efficacy and efficiency under different levels of impaired integrity could be strengthened by establishing a baseline. An extension could be to expose participants to multiple levels of impaired integrity and compare his/her acquisition under impaired integrity to acquisition with perfect integrity using an adapted alternating treatments design (Sindelar, Rosenberg, & Wilson, 1985).

Another limitation involves our conditional probability data. Although we were able to convert data to compute conditional probability to examine potential stimulus biases by condition, we were unable to use our existing data to examine the potential acquisition of

additional error patterns like position biases. A participant's behavior would be indicative of a position bias if he/she routinely responded to a particular position in the array (e.g., first stimulus, last stimulus) regardless of which sample was present and which comparison stimulus was in that position during that trial. Future studies employing similar methodology could configure the computer program to create an output that includes the position of the S+ and the position selected in addition to the name of the S+ and name of the stimulus selected.

Error trials were programmed to occur an equal number of times with each sample stimulus and throughout the integrity comparison. An algorithm was used to randomize these errors across the phase. However, it could be that certain kinds of errors (i.e., commission or omission) may be differentially detrimental depending on whether the learner is in the beginning, middle, or end of acquisition. Perhaps errors occurring earlier on in instruction are more likely to lead to delayed acquisition or deficient instruction. The current program placed errors across the entire phase and used algorithms based on Hirst and DiGennaro Reed (2015), but versions of the program (there were four versions of each condition) varied somewhat in relation to placement of errors within blocks of 20 trials. The proportion of errors of omission and errors of commission could also fluctuate across the acquisition curve. In earlier stages of acquisition, one may be more likely to experience errors of commission because incorrect responses are more prevalent. As conditional discriminations are acquired, one is more likely to experience errors of omission because incorrect selection responses decrease unless behavior comes under the control of faulty sources of stimulus control. Future studies could specifically investigate whether certain types of error (i.e., omission or commission) for specific instructional components (e.g., reinforcement, prompting) are likely to occur at different stages of instruction and may have differential effects on acquisition.

The current study extended Hirst and DiGennaro Reed (2015) by incorporating additional components in the computer program to better approximate best practice recommendations for teaching conditional discriminations to individuals with disabilities (Grow & LeBlanc, 2013). These modifications included removing distracters for conditional discriminations (i.e., each comparison stimulus functioned as both an S+ and an S-), programming a brief ITI, and using praise and points to approximate conditioned reinforcers provided in DTI. However, we did not modify the computer program to include other recommended components of instruction that future researchers could consider incorporating to increase potential ecological validity. These components could include: (a) adding a trial-initiation response (Saunders & Williams, 1998) such as requiring a click to a box to start the next trial, (b) requiring an orienting/observing response (Dinsmoor, 1985) such as clicking on the sample before the comparison array will appear, or (c) programming a differential observing response (Dinsmoor, 1985) such as clicking in a unique location on each sample (e.g., right upper corner for Bifdo, left lower corner for Punfi) or clicking according to a specific schedule that is unique to each sample (e.g., fixed-ratio 4 clicks for Bifdo, differential reinforcement of low rates 30 s for Punfi; Saunders & Spradlin, 1989). At present, it is unknown how these modifications might interact with integrity errors.

Limited external validity is a limitation of the current study. We chose to employ a translational approach to research to evaluate the efficacy and efficiency of instruction in our parametric analysis of programmed treatment integrity errors. This decision was made for several reasons: (a) to replicate Study 1 in Hirst and DiGennaro Reed (2015), (b) to recruit a large enough population to examine differences between conditions, and (c) to avoid exposing a population-in-need to unnecessary errors that could have unintended and potentially long-lasting effects on acquisition. Several descriptive studies (Brand, Elliffe, & DiGennaro Reed, 2017;

Carroll et al., 2013; Cook et al., 2015; Kodak et al., 2017) reported that treatment integrity errors occur during instruction of children with ASD; however, we cannot say whether the errors and degree of errors evaluated in this study are likely to affect acquisition of children with ASD in a similar way. Once additional studies have replicated our findings and suggest a minimum level of integrity when combined errors occur during instruction, future researchers should consider employing a parametric analysis of treatment integrity errors with children with ASD who are receiving behavior analytic skill-acquisition interventions. Researchers should carefully consider the potential risks and benefits of conducting programmed treatment integrity error evaluations with populations in need, however. It could be recommended to include potential safeguards like discontinuation criteria, arbitrary stimuli that are unlikely to affect future acquisition, and teaching unlearned stimuli with high-integrity instruction prior to terminating participation.

We adopted a translational approach to research a phenomenon of applied interest within a highly-controlled human operant setting and with a non-clinical population. This analysis was informed by previous descriptive research (e.g., Carroll et al., 2013). For example, we included errors in reinforcement delivery which were observed to occur in the special education classroom. However, the rates at which these errors were observed in natural settings may be higher than the rates that we programmed for direct evaluation in our experiment. We calibrated our conditions to explore values that may approximate what could be considered “acceptable” integrity and values included in Bergmann et al. (2017) that affected efficacy and/or efficiency of instruction. Additional research is needed to discern which errors are occurring in instructional settings and to what degree; researchers could also collect data on students’ learning to investigate whether concomitant changes in learning occur as a function of integrity. Research manipulating treatment integrity errors should be informed by these descriptive data as an

understanding of the functional relations between different degrees of integrity errors on many instructional components and skill acquisition is warranted. In addition, future studies could examine potential antecedent and consequence manipulations that could improve treatment integrity of skill-acquisition interventions (e.g., Cook et al., 2015; DiGennaro Reed et al., 2007) in a variety of environments. These analyses could extend beyond DTI to naturalistic teaching procedures as well (Donnelly & Karsten, 2017; Pence & St. Peter, 2015).

We conducted a parametric analysis of programmed treatment integrity errors with undergraduate students. Specifically, we used a randomized-control design to examine the effects of errors in reinforcement, errors of omission and errors of commission, when implemented with higher levels of impaired integrity than most previous evaluations of treatment integrity errors. Our parametric analysis focused on integrity between 75% and 95%, and our data revealed that instruction implemented with at least 85% integrity was efficacious and efficient for most participants. More research needs to be conducted before we can establish a standard for “acceptable” integrity but applied behavior analytic skill-acquisition interventions may remain efficacious and efficient for most when implemented with at least 85% integrity.

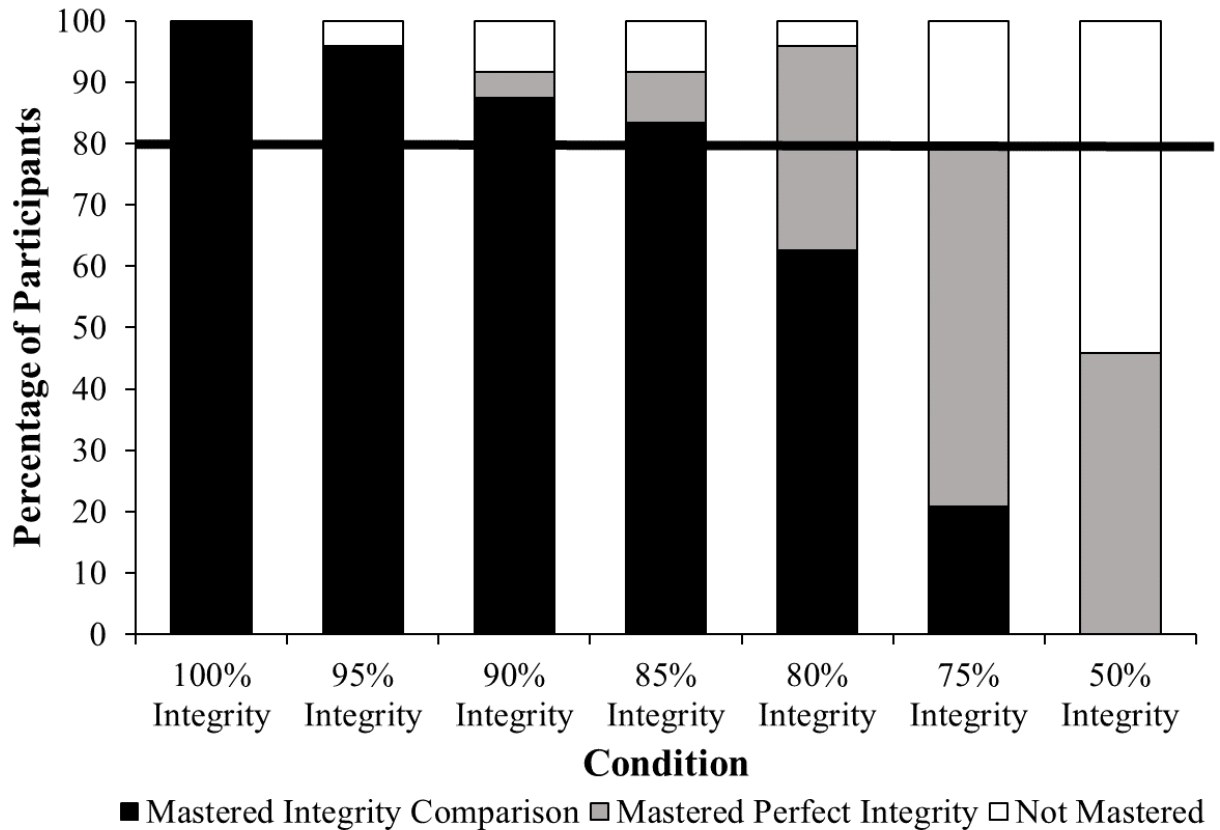


Figure 1. Percentage of participants who met the mastery criterion during the integrity comparison, met the mastery criterion with perfect integrity, and those who never met the mastery criterion. The horizontal line at $y=80$ represents the 80% cutoff for efficacious outcomes.

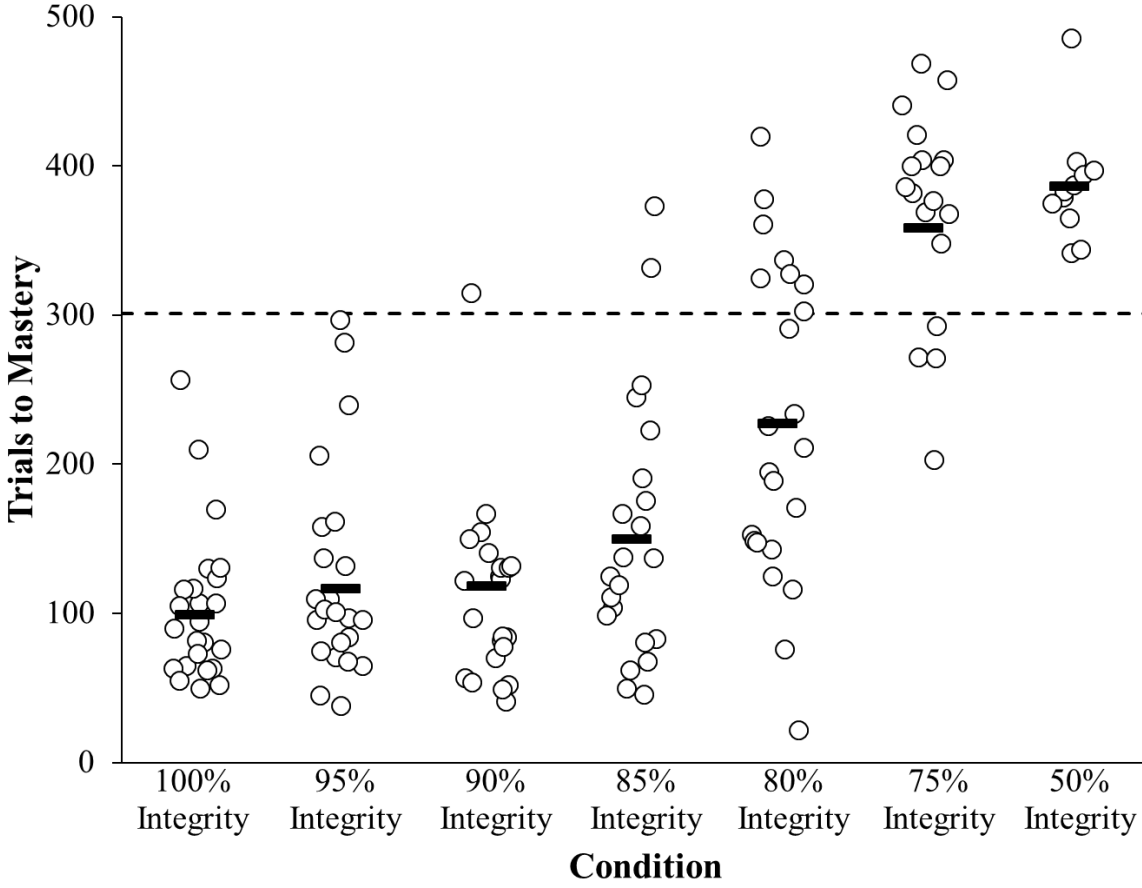


Figure 2. The number of trials to meet the mastery criterion by condition. Each data point is one participant; the black horizontal lines are the mean number of trials to mastery for each condition. The dashed horizontal line at $y=301$ represents the phase change from the integrity comparison to perfect integrity. Data points above the dashed line mean participants met mastery after the phase change to perfect integrity.

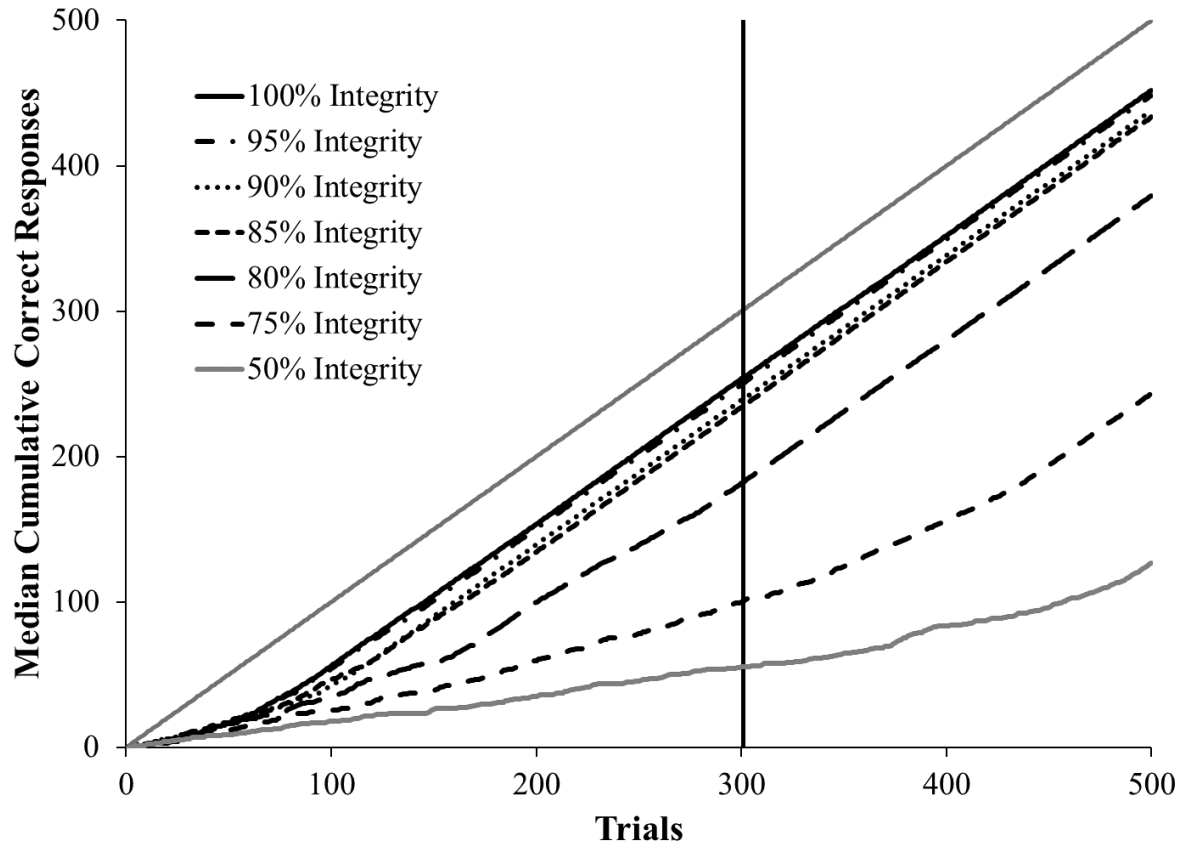


Figure 3. The median cumulative correct responses by condition. The gray diagonal line represents hypothetical perfect acquisition (slope=1). The vertical black line represents the phase change from the integrity comparison to perfect integrity.

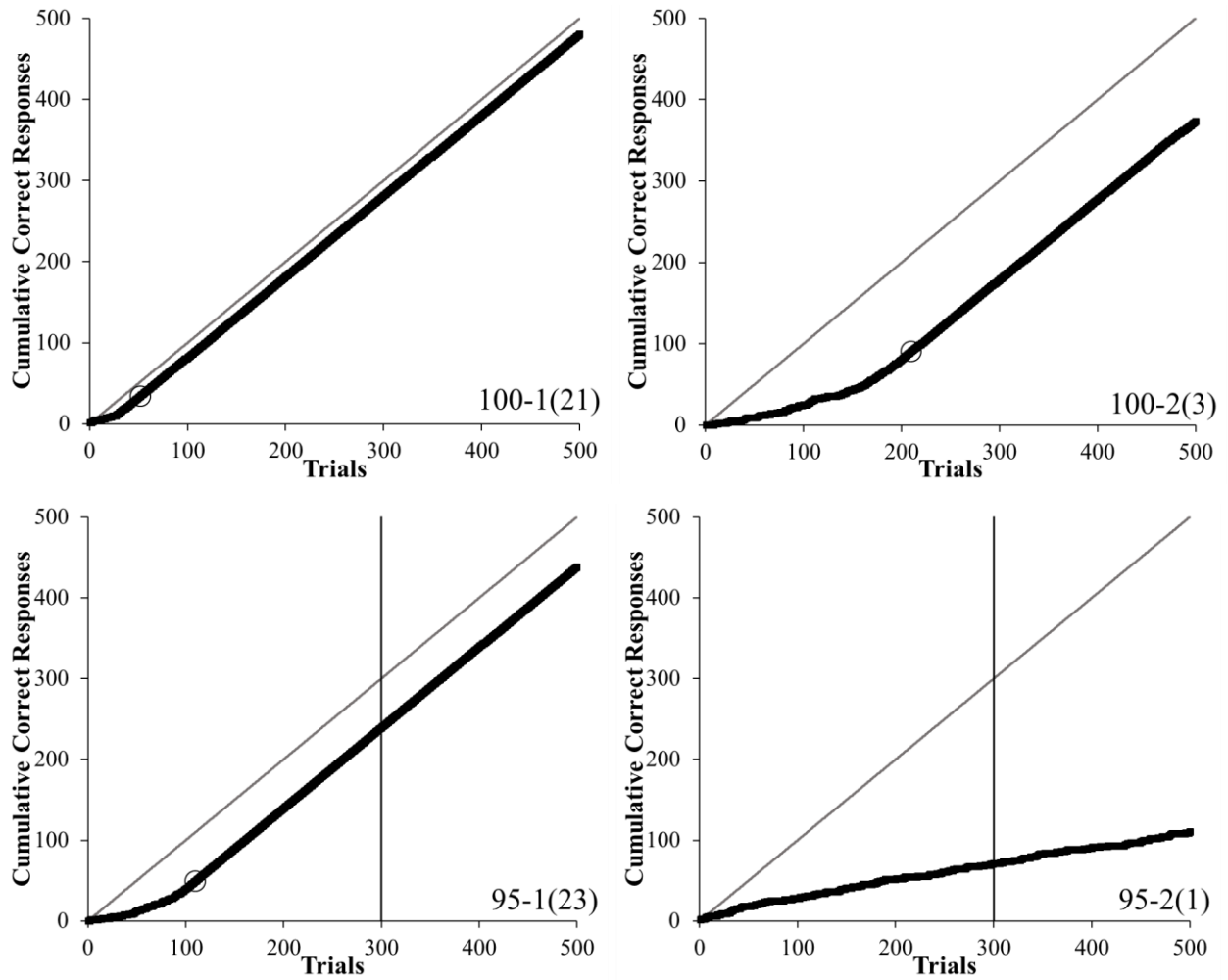


Figure 4. Representative cumulative records for participants in the 100% condition (upper panel) and 95% condition (lower panel). The number in parenthesis denotes the number of participants each graph represents. The dark gray line represents hypothetical perfect responding with a slope of 1. The black data path depicts the participant's cumulative correct responses. An open circle denotes mastery.

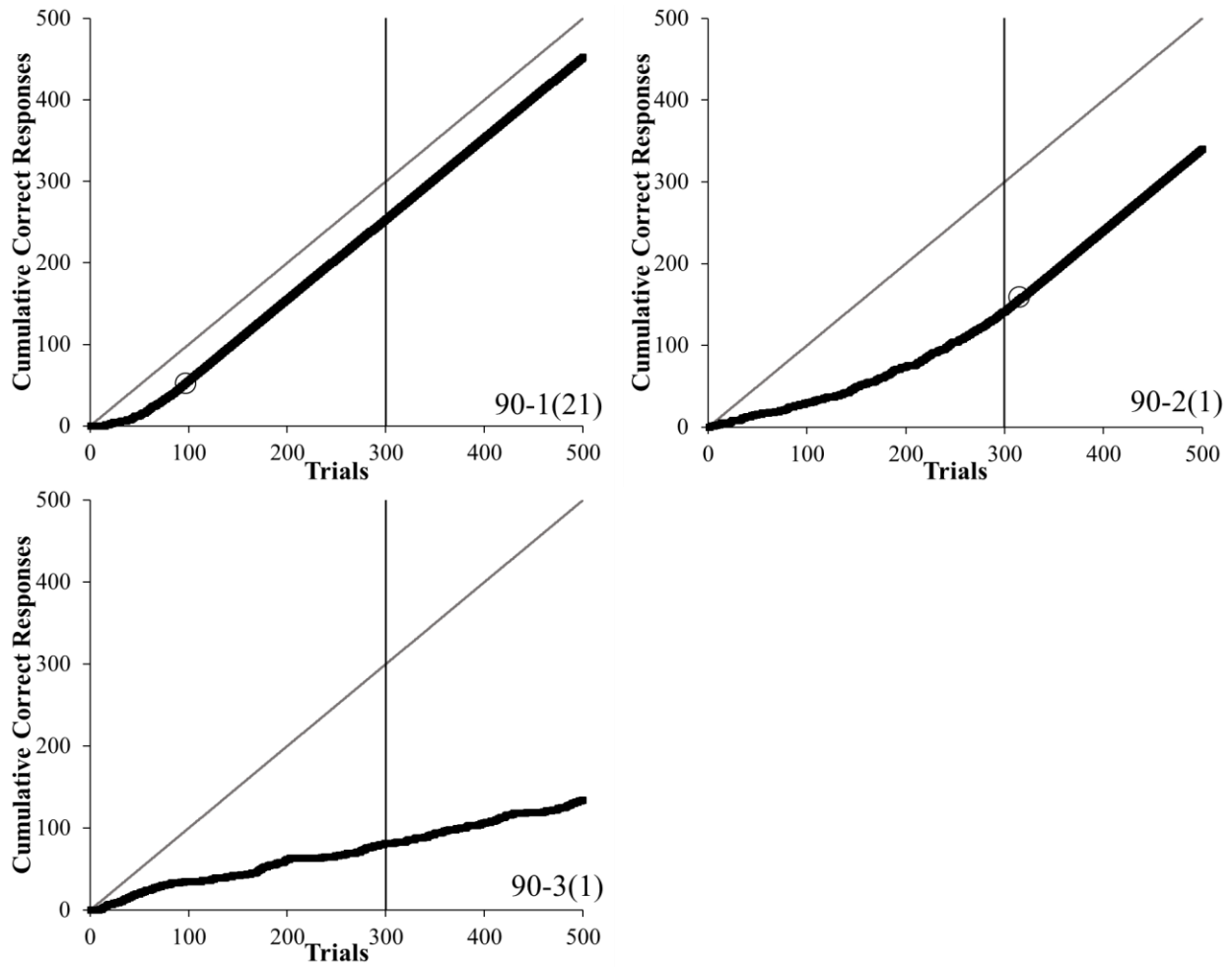


Figure 5. Representative cumulative records for participants in the 90% condition. The number in parenthesis denotes the number of participants each graph represents. The dark gray line represents hypothetical perfect responding with a slope of 1. The black data path depicts the participant's cumulative correct responses. An open circle denotes mastery.

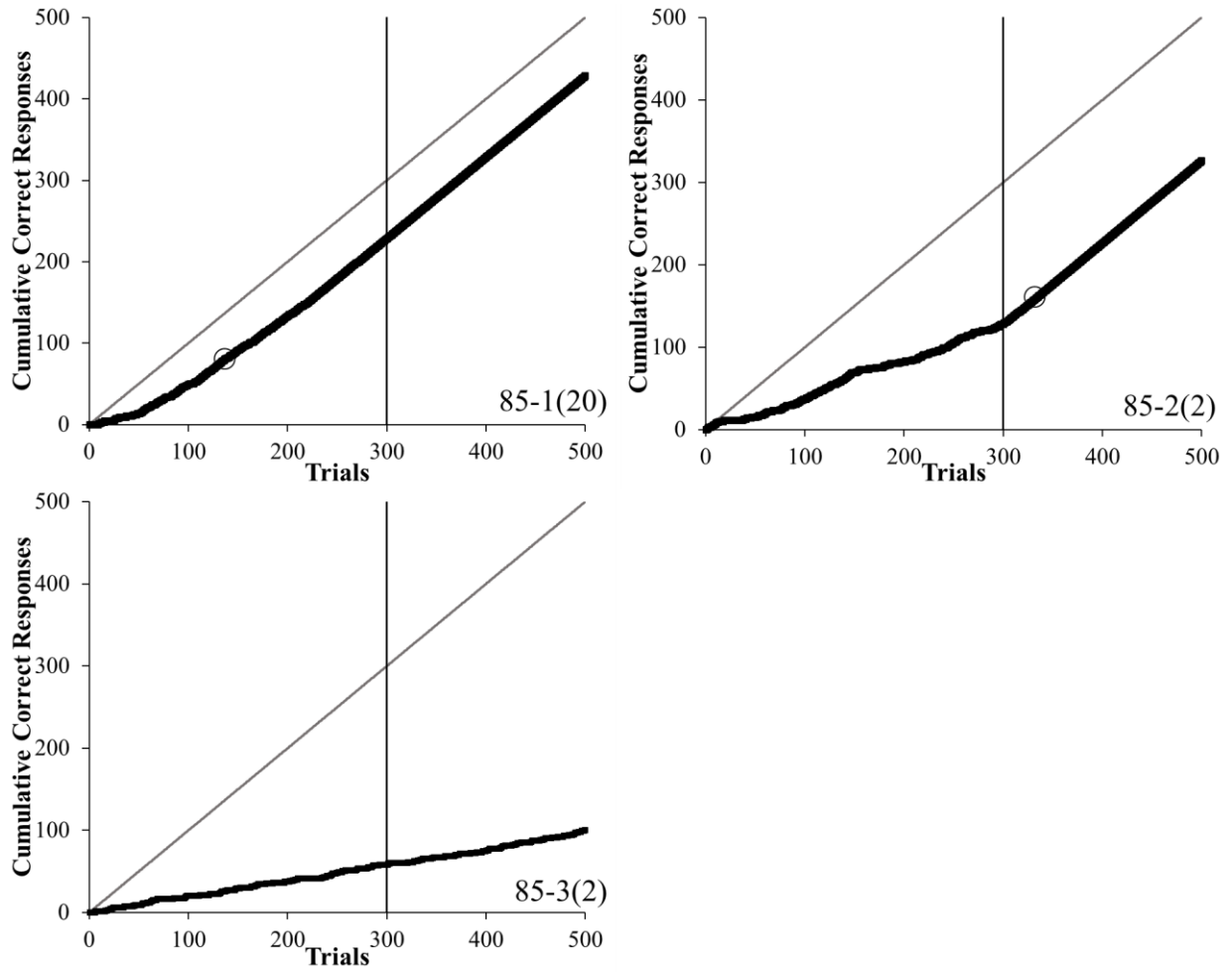


Figure 6. Representative cumulative records for participants in the 85% condition. The number in parenthesis denotes the number of participants each graph represents. The dark gray line represents hypothetical perfect responding with a slope of 1. The black data path depicts the participant's cumulative correct responses. An open circle denotes mastery.

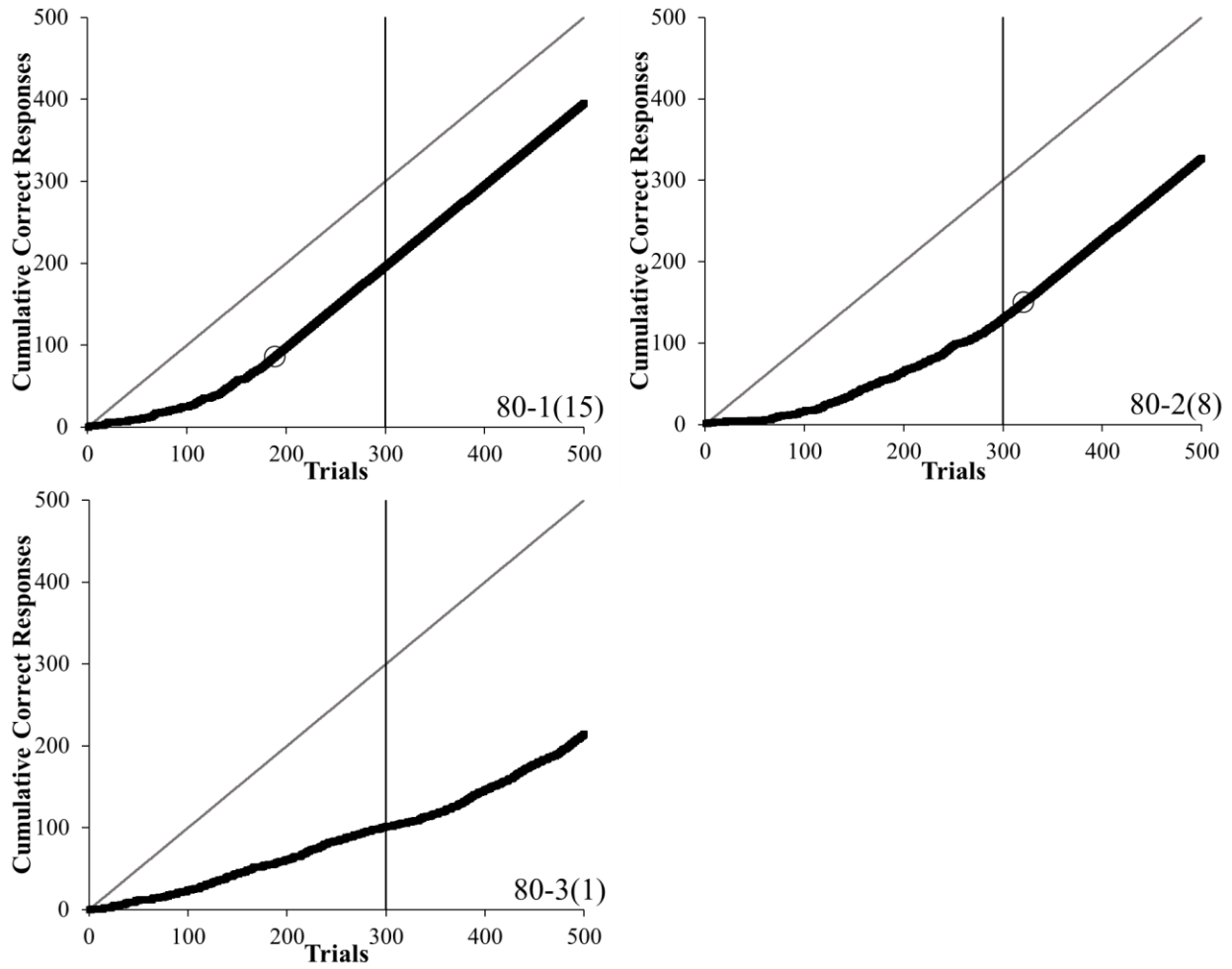


Figure 7. Representative cumulative records for participants in the 80% condition. The number in parenthesis denotes the number of participants each graph represents. The dark gray line represents hypothetical perfect responding with a slope of 1. The black data path depicts the participant's cumulative correct responses. An open circle denotes mastery.

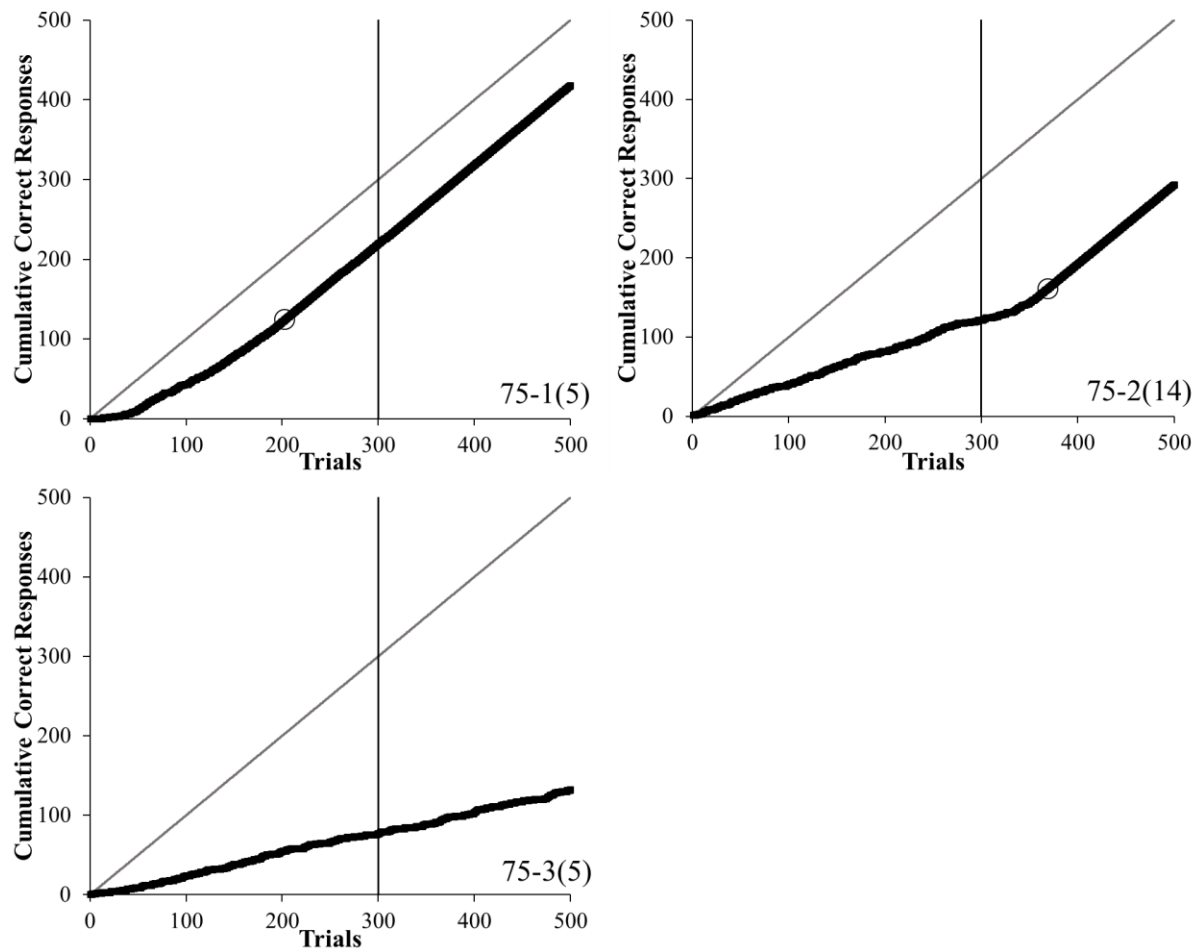


Figure 8. Representative cumulative records for participants in the 75% condition. The number in parenthesis denotes the number of participants each graph represents. The dark gray line represents hypothetical perfect responding with a slope of 1. The black data path depicts the participant's cumulative correct responses. An open circle denotes mastery.

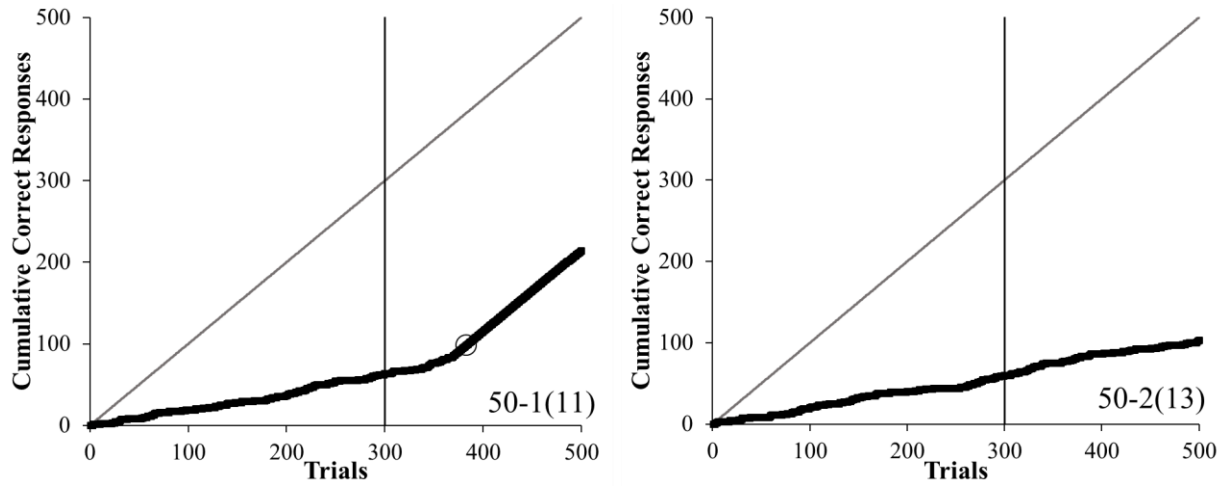


Figure 9. Representative cumulative records for participants in the 50% condition. The number in parenthesis denotes the number of participants each graph represents. The dark gray line represents hypothetical perfect responding with a slope of 1. The black data path depicts the participant's cumulative correct responses. An open circle denotes mastery.

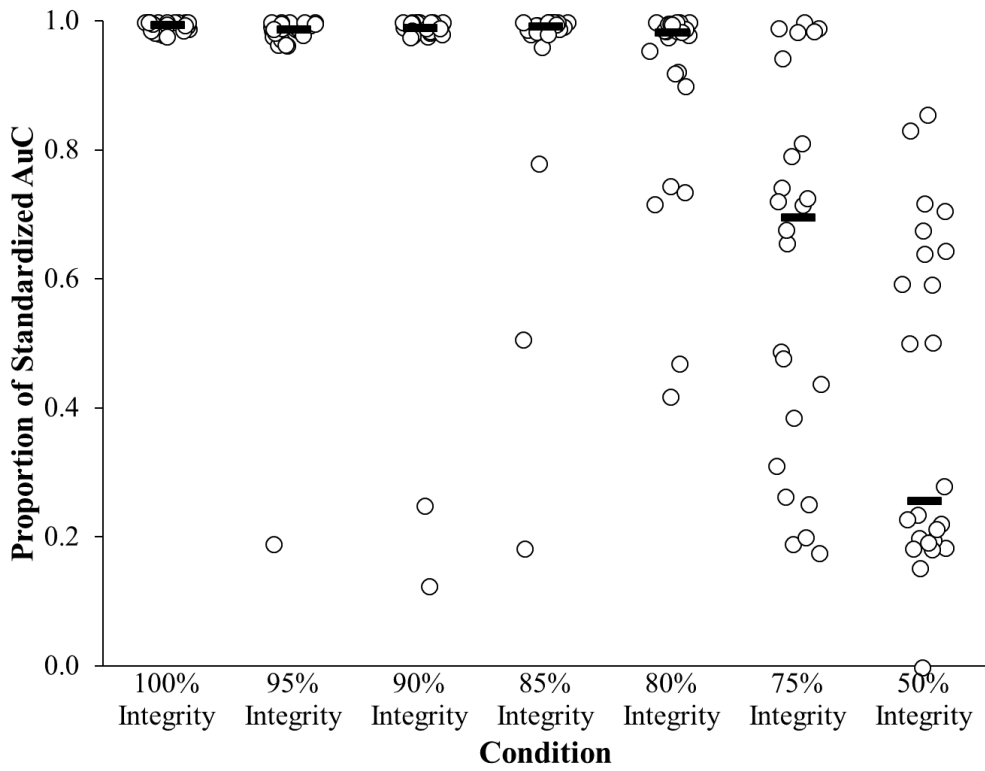
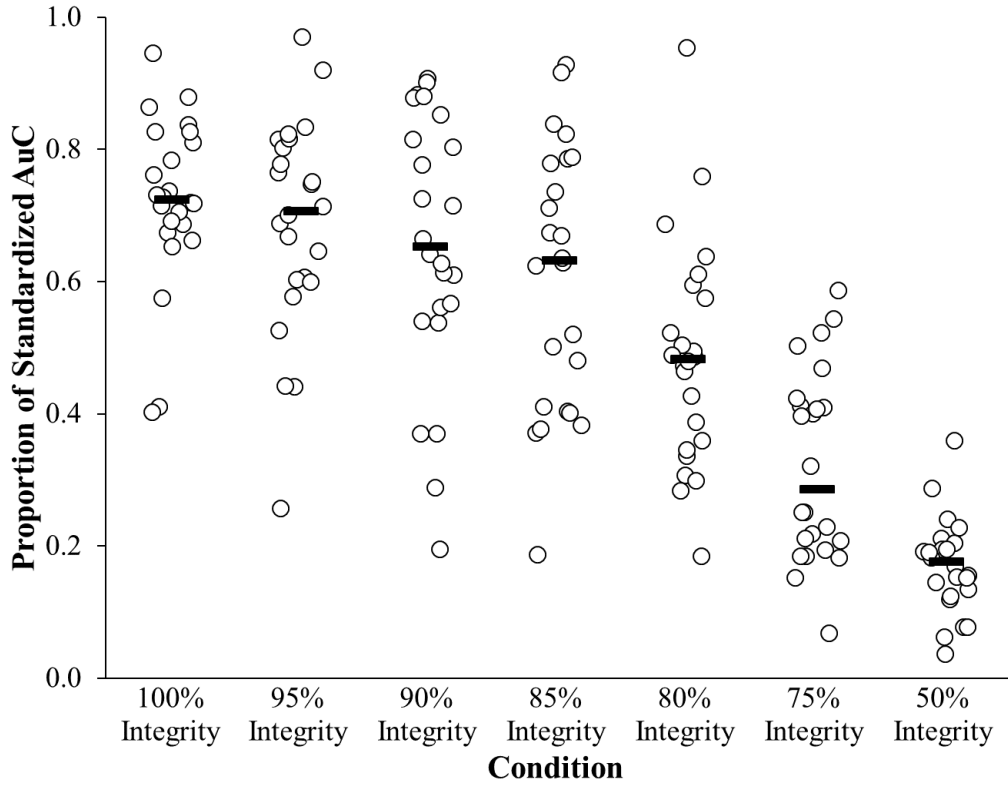


Figure 10. %AuC by condition for the integrity comparison (top panel) and following the phase change to perfect integrity (bottom panel). Each data point is a participant. The horizontal black lines are the median %AuC for each condition.

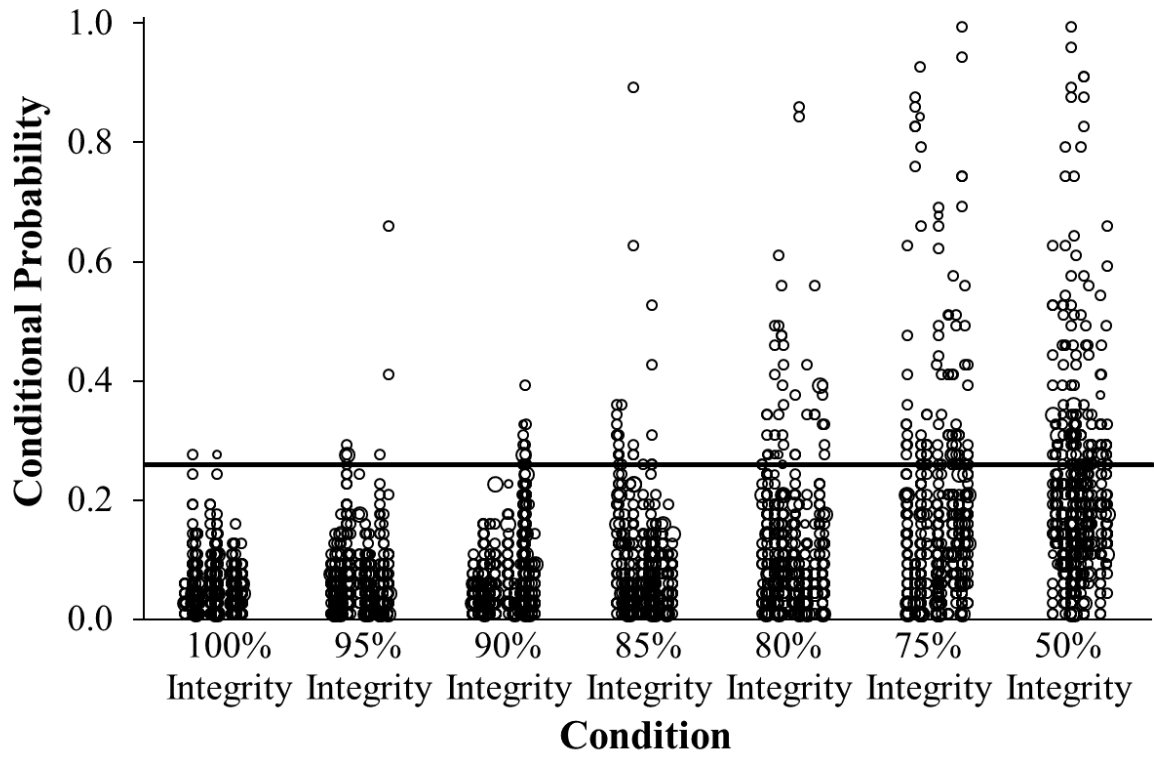


Figure 11. The conditional probability of incorrect selection responses in the presence of each sample stimulus for all participants by condition. The horizontal line at $y=0.25$ represents chance level responding based on four S- in each array.

Table 1

Self-reported Demographic Data as a Percentage of the Sample

Characteristic	Percent (N=168)
Gender	
Male	28
Female	71
No Answer	1
Race/Ethnicity	
Caucasian/White	64
African American	7
Asian	12
Hispanic/Latino	9
Native American	1
Multiple	7
No Answer	1
Disability	
None	92
ADHD	4
Intellectual/DD	0
Physical	3
No Answer	1

Table 2

Textual and Visual Stimuli Presented in AMTS Task



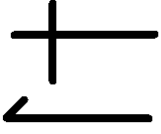
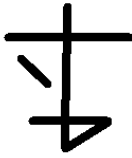

Textual Stimuli	Bifdo	Punfi	Raopol	Smuzy	Zitaaf
Visual Stimuli					

Table 3

Trials Implemented with Integrity and Errors across Conditions

Integrity Condition (%)	Integrity Comparison		Perfect Integrity	
	Integrity	Error	Integrity	Error
100	300	0	200	0
95	285	15	200	0
90	270	30	200	0
85	255	45	200	0
80	240	60	200	0
75	225	75	200	0
50	150	150	200	0

Table 4

Mastery by Condition and Trials to Mastery

Integrity Condition (%)	Frequency of Mastery ^a		Trials to Mastery	
	Integrity Comparison	Perfect Integrity	Mean	Range
100	24	--	99 (52.1)	35-257
95	23	0	117 (60.1)	38-282
90	21	1	119 (71.5)	41-315
85	20	2	150 (88.3)	46-373
80	15	8	227 (105.8)	22-420
75	5	14	359 (84.4)	153-469
50	0	11	387 (38.5)	342-486

Note. Standard deviations in parentheses.

^a $n = 24$

References

- Behavior Analyst Certification Board, Inc. (2014). Professional and ethical compliance code for behavior analysts. Retrieved from <http://bacb.com/wp-content/uploads/2016/03/160321-compliance-code-english.pdf>
- Bergmann, S., Kodak, T., & LeBlanc, B. (2017). Effects of programmed errors of omission and commission during auditory-visual conditional discrimination training with typically developing children. *Psychological Record*, 67(1), 109-119. doi:10.1007/s40732-016-0211-2
- Brand, D., Elliffe, D., & DiGennaro Reed, F. D. (2017). Using sequential analysis to assess component integrity of discrete-trial teaching programs. *European Journal of Behavior Analysis*, 1-18. doi:10.1080/15021149.2017.1404392
- Carroll, R. A., Kodak, T., & Fisher, W. W. (2013). An evaluation of programmed treatment-integrity errors during discrete-trial instruction. *Journal of Applied Behavior Analysis*, 46(2), 379-394. doi:10.1002/jaba.49
- Cook, J. E., Subramaniam, S., Brunson, L. Y., Larson, N. A., Poe, S. G., & St. Peter, C. C. (2015). Global measures of treatment integrity may mask important errors in discrete-trial training. *Behavior Analysis in Practice*, 1-11. doi:10.1007/s40617-014-0039-7.
- Detrich, R. (2014). Treatment integrity: Fundamental to education reform. *Journal of Cognitive Education and Psychology*, 13(2), 258-271. doi:10.1891/1945-8859.13.2.258
- DiGennaro, F. D., Martens, B. K., & Kleinmann, A. E. (2007). A comparison of performance feedback procedures on teachers' treatment implementation integrity and students' inappropriate behavior in special education classrooms. *Journal of Applied Behavior Analysis*, 40(3), 447-461. doi: 10.1901/jaba.2007.40-447

- Dinsmoor, J. A. (1985). The role of observing and attention in establishing stimulus control. *Journal of the Experimental Analysis of Behavior*, 43(3), 365-381.
- Donnelly, M. G. & Karsten, A. M. (2017). Effects of programmed teaching errors on acquisition and durability of self-care skill. *Journal of Applied Behavior Analysis*, 50, 511-528. doi: 10.1002/jaba.390
- Eikeseth, S., Klintwall, L., Jahr, E., & Karlsson, P. (2012). Outcome for children with autism receiving early and intensive behavioral intervention in mainstream preschool and kindergarten settings. *Research in Autism Spectrum Disorders*, 6, 829-835. doi: 10.1016/j.rasd.2011.09.002
- Fryling, M. J., Wallace, M. D., & Yassine, J. N. (2012). Impact of treatment integrity on intervention effectiveness. *Journal of Applied Behavior Analysis*, 45(2), 449-453. doi:10.1901/jaba.2012.45-449
- Green, G. (2001). Behavior analytic instruction for learners with autism: Advances in stimulus control technology. *Focus on Autism and Other Developmental Disabilities*, 16(2), 72-85.
- Gresham, F. M. (1989). Assessment of treatment integrity in school consultation and prereferral intervention. *School Psychology Review*, 18(1), 37-50.
- Grow, L. & LeBlanc, L. (2013). Teaching receptive language skills: Recommendations for instructors. *Behavior Analysis in Practice*, 6(1), 56-75.
- Hirst, J. M., & DiGennaro Reed, F. D. (2015). An examination of the effects of feedback accuracy on academic task acquisition in analogue settings. *Psychological Record*, 65, 49-65. doi:10.1007/s40732-014-0087-y
- Holcombe, A., Wolery, M., & Snyder, E. (1994). Effects of two levels of procedural fidelity with constant time delay on children's learning. *Journal of Behavioral Education*, 4(1), 49-73.

- Lerman, D. C. (2003). From the laboratory to community application: Translational research in behavior analysis. *Journal of Applied Behavior Analysis*, 36(4), 415-419.
doi:10.1901/jaba.2003.36-415
- Jenkins, S. R., Hirst, J. M., DiGennaro Reed, F. D. (2015). The effects of discrete-trial training commission errors on learner outcomes: An extension. *Journal of Behavioral Education*, 24, 196-209. doi:10.1007/s10864-014-9215-7
- Kodak, T., Cariveau, T., LeBlanc, B., Mahon, J., & Carroll, R. A. (2017). Selection and implementation of skill acquisition programs by special education teachers and staff for students with autism spectrum disorder. *Behavior Modification*, 1-26, doi: 10.1177/0145445517692081
- Mace, F. C., & Critchfield, T. S. (2010). Translational research in behavior analysis: Historical traditions and imperative for the future. *Journal of Experimental Analysis of Behavior*, 93(3), 293-312. doi: 10.1901/jeab.2010.93-293
- McIlvane, W. J., Dube, W. V., Serna, R. W., Lionello-DeNolf, K. M., Barros, R. S., & Galvao, O. F. (2011). Some current dimensions of translational behavior analysis: From laboratory research to intervention for persons with autism spectrum disorder. In J. A. Mulick & E.A. Mayville (Eds.), *Behavioral Foundations of Effective Autism Treatment* (pp. 155-181). Cornwall-on-Hudson, NY: Sloan Publishing.
- Myerson, J., Green, L., & Warusawitharana, M. (2001). Area under the curve as a measure of discounting. *Journal of the Experimental Analysis of Behavior*, 76(2), 235-243.
- National Autism Center. (2015). Findings and conclusions: National standards project, phase 2. Randolph, MA; Author.

- Noell, G. H., Gresham, F. M., & Gansle, K. A. (2002). Does treatment integrity matter? A preliminary investigation of instructional implementation and mathematics performance. *Journal of Behavioral Education*, 11(1), 51-67. doi:10.1023/A:1014385321849
- Pence, S. T., & St Peter, C. C. (2015). Evaluation of treatment integrity errors on mand acquisition. *Journal of Applied Behavior Analysis*, 48(3), 575–589. doi:10.1002/jaba.238.
- Perepletchikova, F., Hilt, L. M., Chereji, E., & Kazdin, A. E. (2009). Barriers to implementing treatment integrity procedures: Survey of treatment outcome researchers. *Journal of Consulting and Clinical Psychology*, 77(2), 212-218. doi: 10.1037/a0015232
- Reichow, B. & Wolery, M. (2009). Comprehensive synthesis of early intensive behavioral interventions for young children with autism based on the UCLA Young Autism Project model. *Journal of Autism and Developmental Disorders*, 39, 23-41. doi: 10.1007/s10803-008-0596-0
- Saunders, K. J. & Spradlin, J. E. (1989). Conditional discrimination in mentally retarded adults: The effect of training the component simple discriminations. *Journal of the Experimental Analysis of Behavior*, 52(1), 1-12.
- Saunders, K. J. & Williams, D. C. (1998). Stimulus control procedures. In K. Lattal & M. Perone (Eds.), *Handbook of research methods in human operant behavior* (pp. 193-228). New York, NY: Plenum Press.
- Sindelar, P. T., Rosenberg, M. S., & Wilson, R. J. (1985). An adapted alternating treatments design for instructional research. *Education and Treatment of Children*, 8(1), 67–76.
- St. Peter Pipkin, C., Vollmer, T. R., & Sloman, K. N. (2010). Effects of treatment integrity failures during differential reinforcement of alternative behavior: A translational model. *Journal of Applied Behavior Analysis*, 43(1), 47-70. doi:10.1901/jaba.2010.43-47

Van Houten, R., Axelrod, S., Bailey, J., Favell, J., Foxx, R., Iwata, B., & Lovaas, O. (1988). The Right to Effective Behavioral Treatment. *The Behavior Analyst*, 11(2), 111-114.

Wong, C., Odom, S. L., Hume, K., Cox, A. W., Fettig, A., Kucharczyk, S., ... Schultz, T. R. (2014). *Evidence-based practices for children, youth, and young adults with autism spectrum disorder*. Chapel Hill: The University of North Carolina, Frank Porter Graham Child Development Institute, Autism Evidence-Based Practice Review Group.

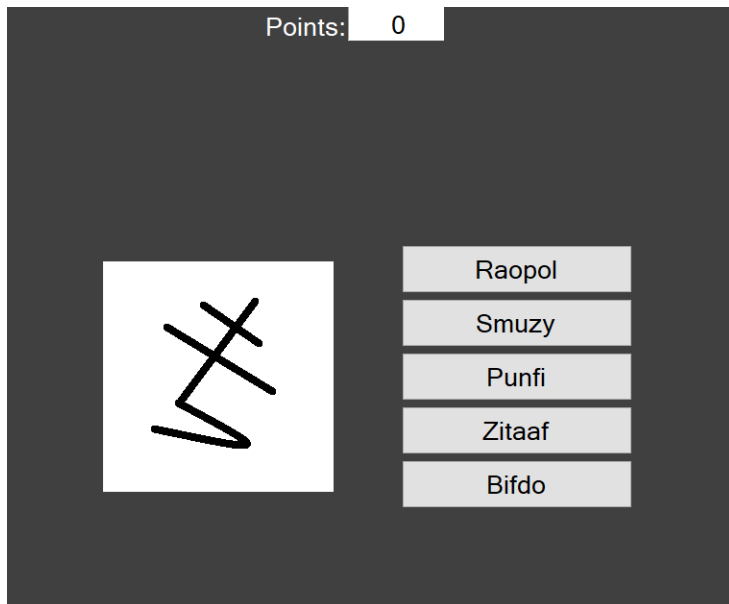
Appendix A:

Participant Instructions and Screenshots of Microsoft® Visual Basic .Net

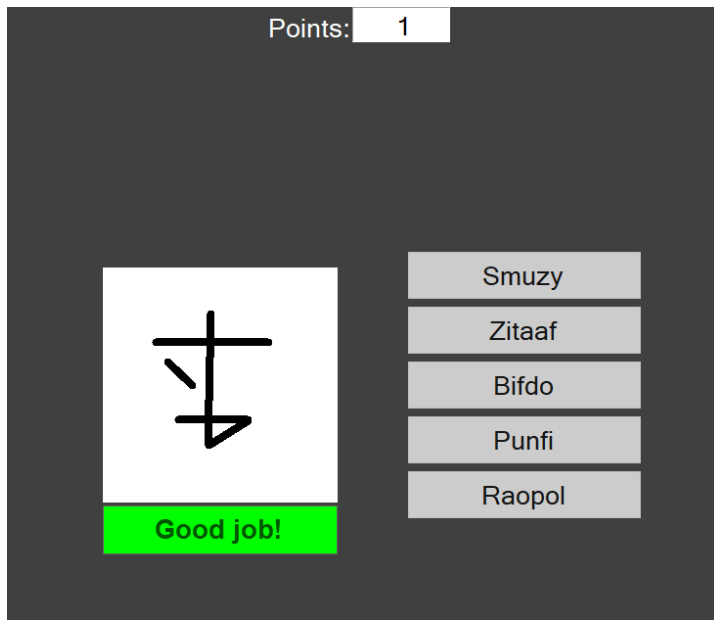
Instructions shown on screen following consent and before the first trial

You will learn to match symbols with words. In this study, you will receive feedback on your performance. When you get an answer right, you will read a positive statement and get a point. These points will be tracked and participants with the highest scores will earn a \$10.00 Target® gift card. When you get an answer wrong, you will not get a positive statement nor a point. This study will last one hour. Your participation is completely voluntary. If you understand these instructions, you may click the button below to begin. If you have questions, please raise your hand and the experimenter will be available to assist you.

Trial arrangement with visual stimulus on left and textual stimuli response options on right



Screen display following a correct response

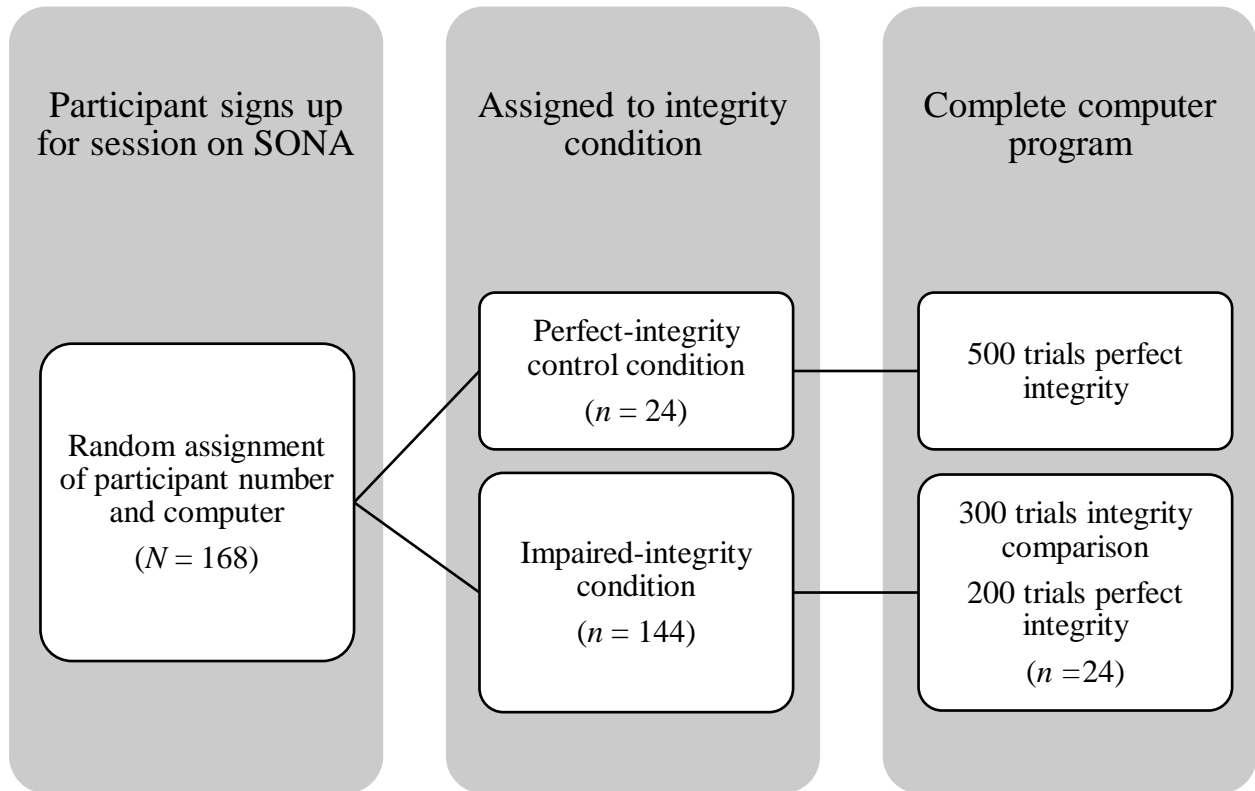


Screen display during ITI



Appendix B:

Random Assignment of Participants to Conditions



Appendix C:

Informed Consent Packet

UNIVERSITY OF WISCONSIN – MILWAUKEE CONSENT TO PARTICIPATE IN RESEARCH

1. General Information

Study title: Establishing Quality Standards for Behavior Analytic Intervention: A Translational Model with Undergraduate Students and Children Diagnosed with Autism Spectrum Disorder

Person in Charge of Study (Principal Investigator): Tiffany Kodak, Ph.D., Assistant Professor in the Department of Psychology at the University of Wisconsin-Milwaukee

2. Study Description

You are being asked to participate in a research study. Your participation is completely voluntary. You do not have to participate if you do not want to.

Study description:

The purpose of this study is to examine how people learn new skills. The goal of the study is to examine how quickly and accurately you learn new information when presented on a computer screen. The study will be conducted in a university computer laboratory at the University of Wisconsin-Milwaukee. We will recruit up to 208 undergraduate students to participate in this study. Participation requires attending one 1.5-hour appointment.

3. Study Procedures

What will I be asked to do if I participate in the study?

If you agree to participate, you will be asked to come to the university computer laboratory at the University of Wisconsin-Milwaukee for one 1.5-hour appointment. During the appointment, you will be shown pictures on a computer screen and asked to match the pictures with words. You will earn points for each correct match. We will track your total number of points. Points will be used to determine the highest point earners in each condition, there will be multiple conditions (i.e., four) per appointment. All activities will take place on a computer. Gift cards will be provided to the participants with the highest points in their condition prior to the end of the study session.

4. Risks and Minimizing Risks

What risks will I face by participating in this study?

There is little-to-no risk associated with participating in this research study. You may feel fatigued or tired during the tasks; if your eyes become strained, please look away from the computer screen briefly. If you need a break to use the restroom, please inform the researcher. If you experience frustration, stress, or distress during or following this study, please consider seeking clinical resources at Norris Health Center 414-229-4716 (mentalhealth.uwm.edu). Counseling services for students are available on the fifth floor of the Northwest Quadrant Building and can be accessed by using the RED elevators to go to reception. Hours M-Th 8:00-4:45, F 9:00-4:45.

5. Benefits

Will I receive any benefit from my participation in this study?

There is no direct benefit to you from your participation in this study. Your participation could help inform future work with children with autism spectrum disorder and developmental disabilities.

6. Study Costs and Compensation

Will I be charged anything for participating in this study?

You will not be responsible for any of the costs from taking part in this research study.

Are subjects paid or given anything for being in the study?

You will receive extra credit for participating in this study. You will receive 1.5 hours recorded on SONA. The participants with the highest scores will also receive a \$10 Target gift card. One out of every four participants will earn a gift card for their high score in each condition. Approximately four participants per session, one participant per condition, will earn a gift card prior to the end of the study session.

7. Confidentiality

What happens to the information collected?

All information collected during the course of this study will be kept confidential to the extent permitted by law. We may decide to present what we find to others, or publish our results in scientific journals or at scientific conferences. We are not collecting information that identifies you nor will any information be shared without your permission. Only the PI and research assistant will have access to the information. However, the Institutional Review Board at UW-

Milwaukee or appropriate federal agencies like the Office for Human Research Protections may review your related records.

Your information will be stored and coded based on a participant number. Your name will not be attached to the participant number. The information collected in this study will be stored in encrypted and password-protected storage devices or on paper that will be stored in a locked cabinet. The data will be stored in locked filing cabinets for seven years after completion of this study. Following seven years, all information will be destroyed.

8. Alternatives

Are there alternatives to participating in the study?

There are alternatives to participating in this study to receive class credit. You should consult the SONA research participation website for information about other ways to receive extra credit points for participating in research studies as well as a non-study option for extra credit.

9. Voluntary Participation and Withdrawal

What happens if I decide not to be in this study?

Your participation in this study is entirely voluntary. You may choose not to take part in this study. If you decide to take part, you can change your mind later and withdraw from the study. Your decision to withdraw from the study will not change any present or future relationships with the University of Wisconsin Milwaukee. Refusal to participate in the study will not affect your grade or class standing. You will earn extra credit for the amount of time spent in the study. You will not be eligible to earn the \$10 Target gift card if you do not complete the study.

10. Questions

Who do I contact for questions about this study?

For more information about the study or the study procedures or treatments, or to withdraw from the study, contact:

Tiffany Kodak, Ph.D.
Department of Psychology
2441 E. Hartford Ave., Garland 238E
Milwaukee, WI 53211
414-229-7383
kodak@uwm.edu

Who do I contact for questions about my rights or complaints towards my treatment as a research subject?

The Institutional Review Board may ask your name, but all complaints are kept in confidence.

Institutional Review Board

Human Research Protection Program
Department of University Safety and Assurances
University of Wisconsin – Milwaukee
P.O. Box 413
Milwaukee, WI 53201
(414) 229-3173

11. Signatures

Research Subject's Consent to Participate in Research:

To voluntarily agree to take part in this study, you must sign on the line below. If you choose to take part in this study, you may withdraw at any time. You are not giving up any of your legal rights by signing this form. Your signature below indicates that you have read or had read to you this entire consent form, including the risks and benefits, and have had all of your questions answered, and that you are 18 years of age or older.

Printed Name of Subject/ Legally Authorized Representative

Signature of Subject/Legally Authorized Representative

Date

Principal Investigator (or Designee)

I have given this research subject information on the study that is accurate and sufficient for the subject to fully understand the nature, risks and benefits of the study.

Printed Name of Person Obtaining Consent

Study Role

Signature of Person Obtaining Consent

Date

Appendix D:

Debriefing Information

Thank you for participating in this study. The purpose of the study was to examine the effects of making some errors during instruction on learning. We did not tell you about this purpose of the study during informed consent. This was done to reduce the likelihood that you noticed the errors the computer made and changed how you responded during the task as a result. During some of the trials (range 0%-75%), the computer committed an error during instruction. Sometimes, the computer did not provide praise (e.g., "Well done!") and a point when you selected the correct words. Sometimes, the computer provided praise and a point when you selected the wrong words. During the first 300 trials, the computer made errors on _____% of your trials. During the last 200 trials, the computer made errors on 0% of your trials.

There were multiple conditions in this study. Each condition had a different rate of computer-programmed errors. That means that participants assigned to different conditions were exposed to different amounts of errors and these errors could have impacted the speed of learning.

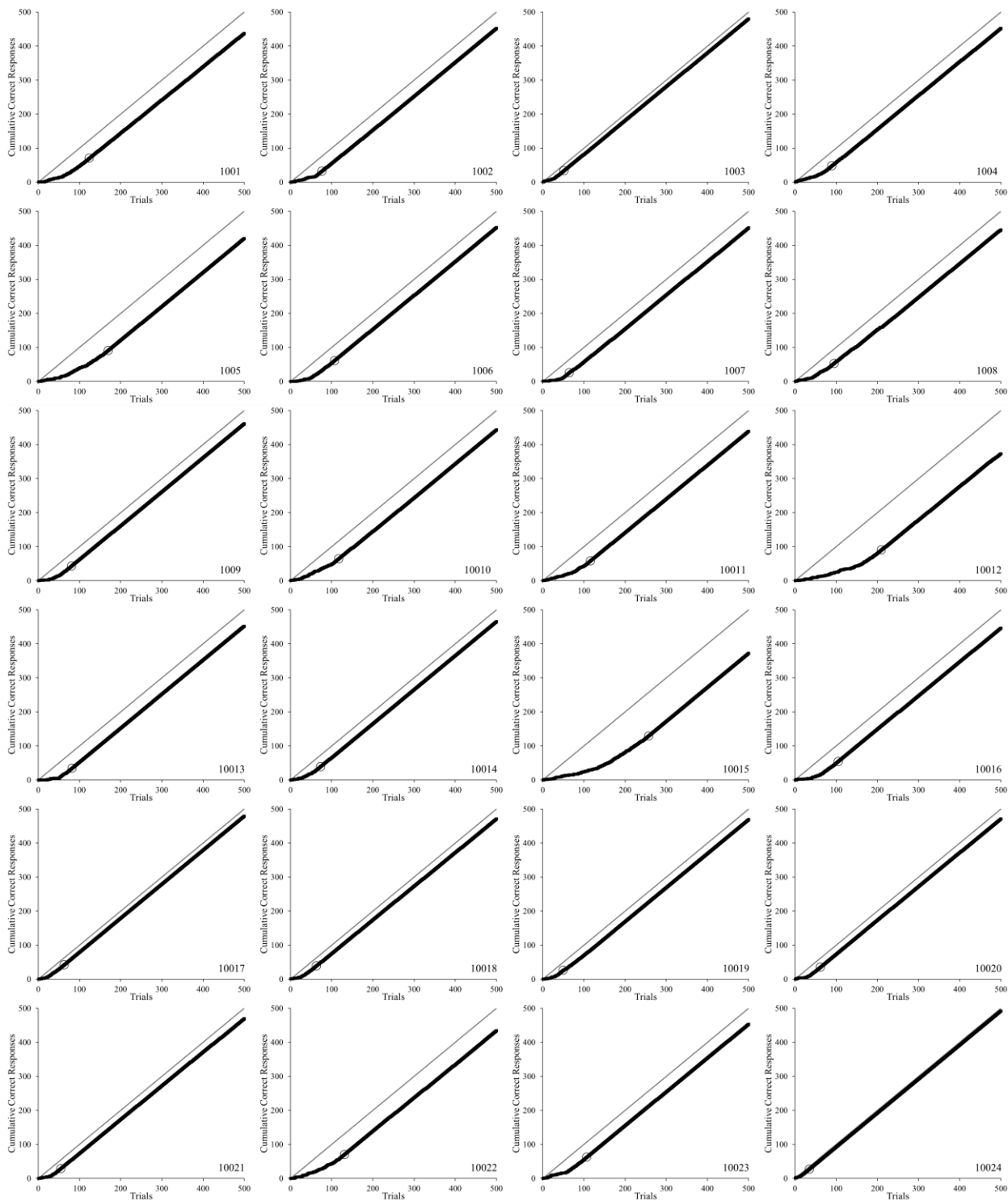
These kinds of errors have been observed to occur during educational activities with children. We used these procedures to see if these kinds of errors would prevent you from learning or slow down your learning. This information will help us learn more about how errors affect learning and whether learning can withstand some errors. This will help us in future research with children learning skills. Please inform your experimenter if you would like to know more about the number of errors made or if you would like to see a key of correct answers.

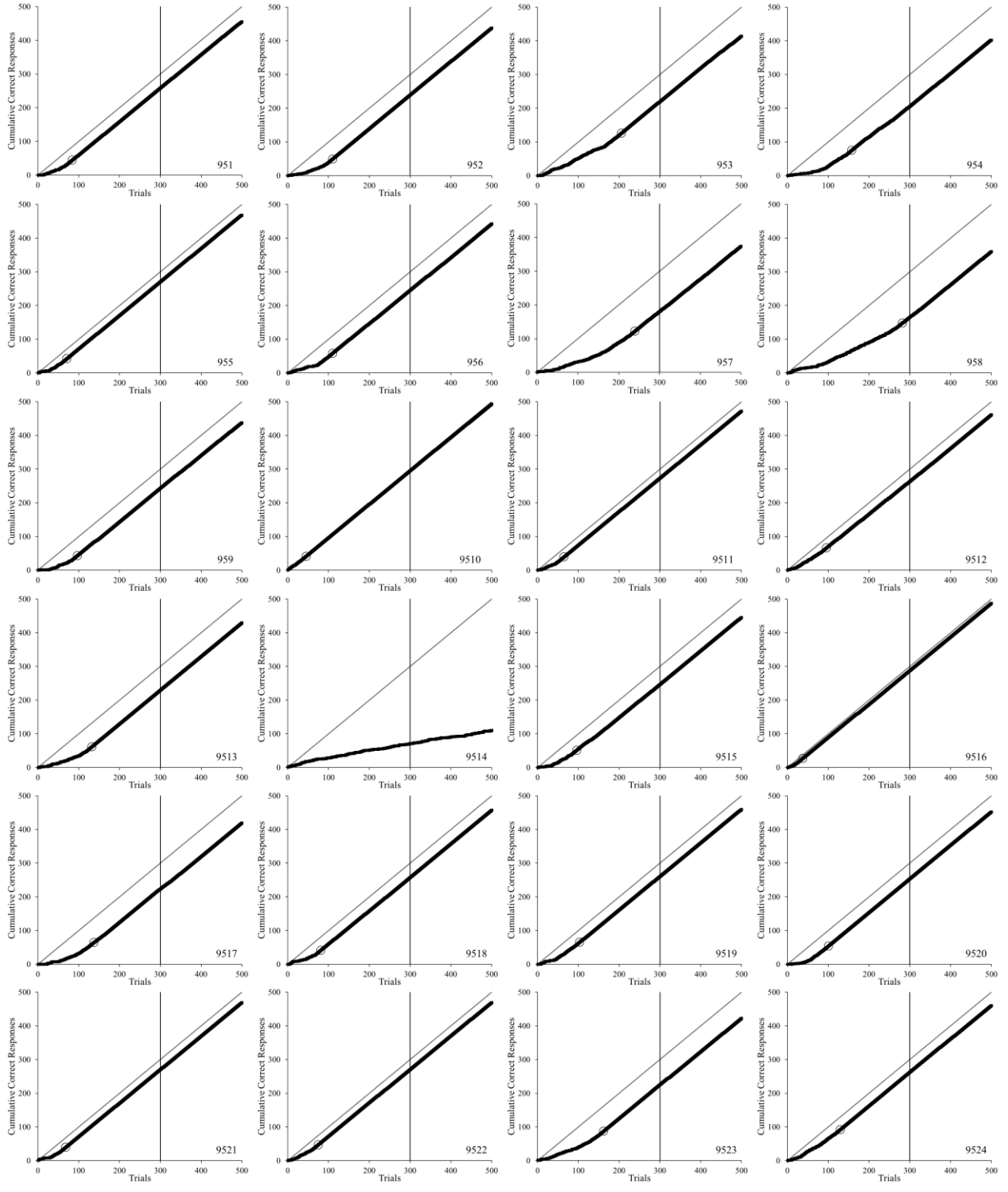
If you experience frustration, stress, or distress during or following this study, please consider seeking clinical resources at Norris Health Center 414-229-4716 (mentalhealth.uwm.edu). Counseling services for students are available on the fifth floor of the Northwest Quadrant Building and can be accessed by using the RED elevators to go to reception. Hours M-Th 8:00-4:45, F 9:00-4:45.

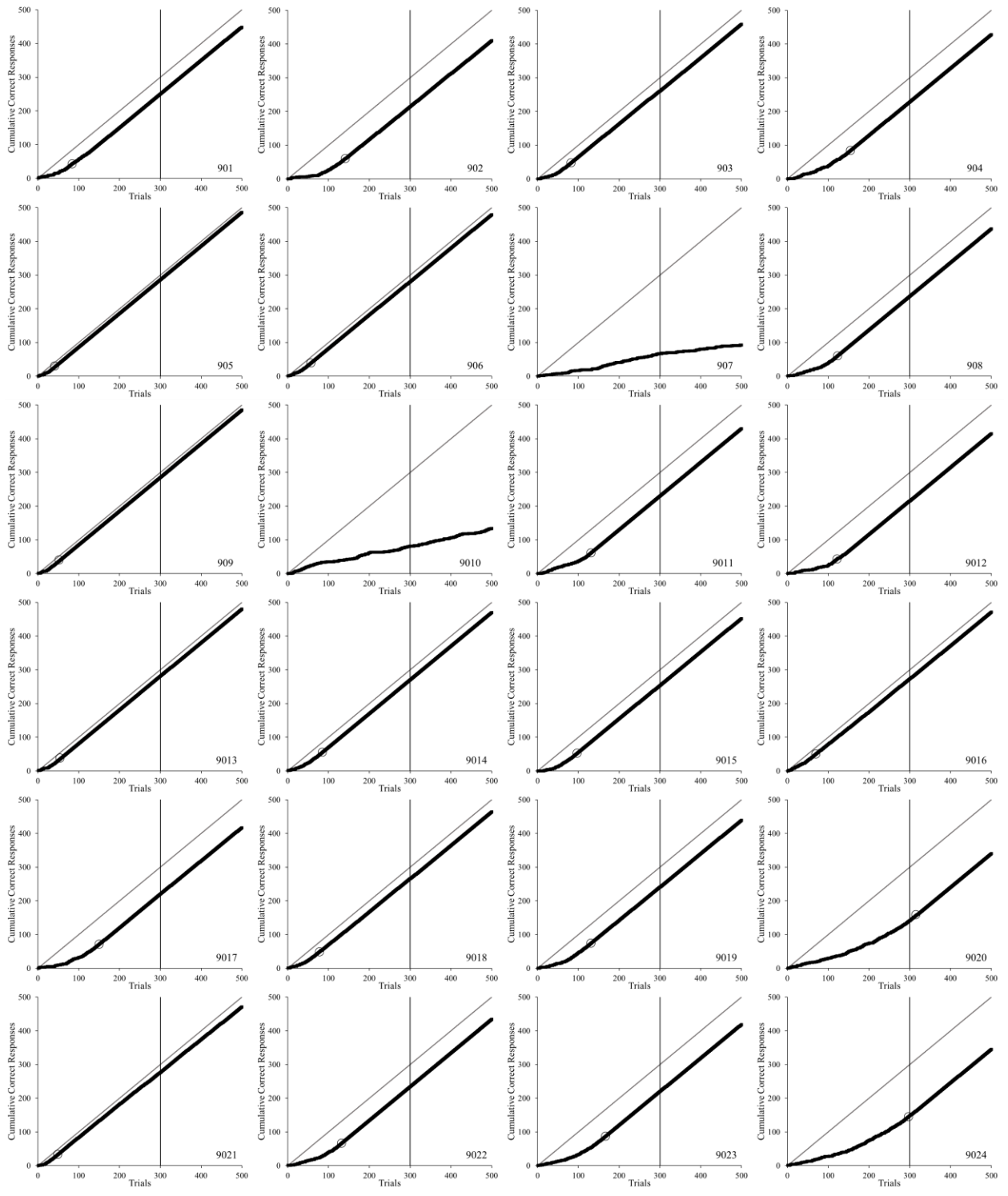
If you have additional questions about this study, feel free to talk to the experimenter or contact Dr. Tiffany Kodak at kodak@uwm.edu

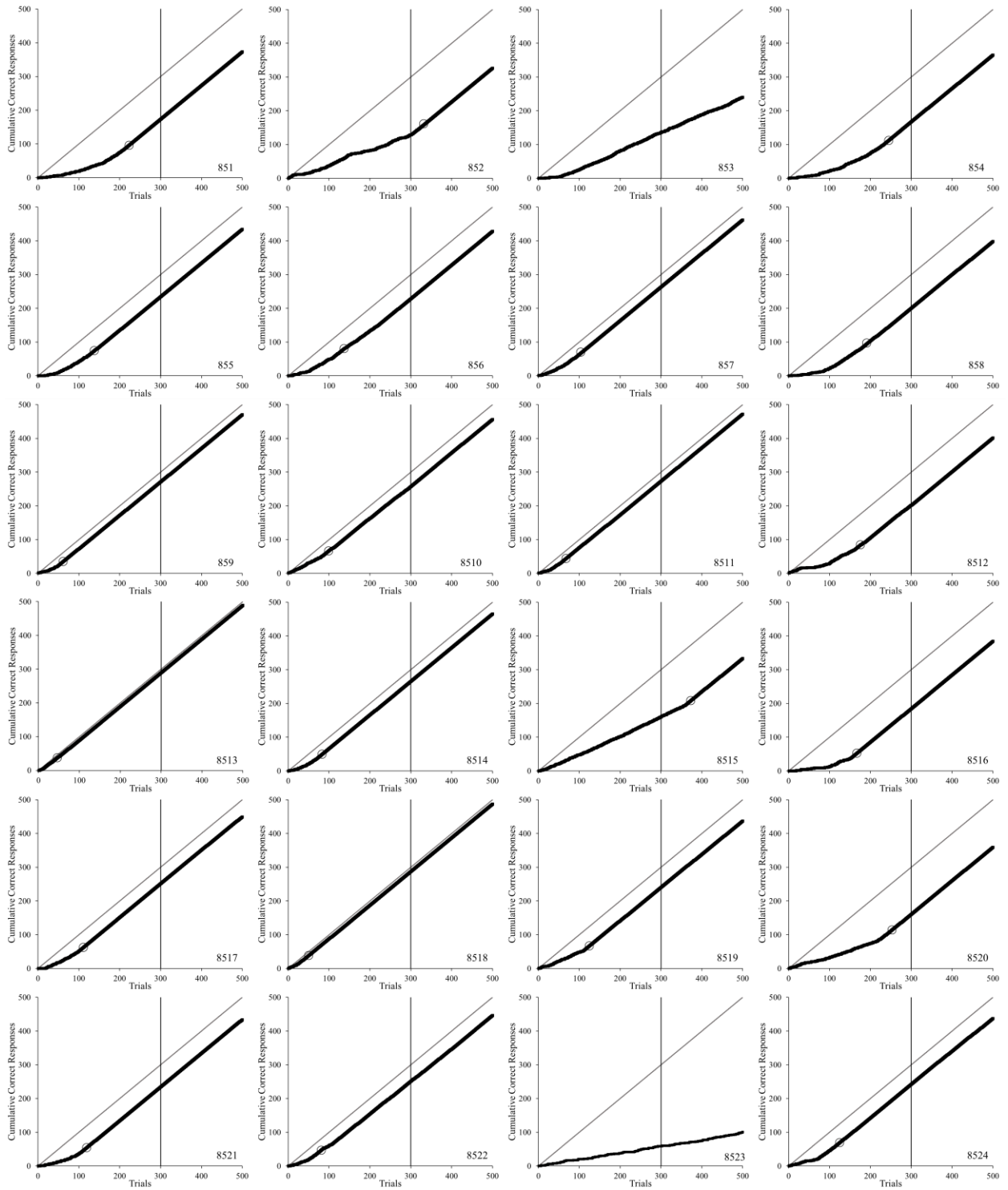
Appendix E:

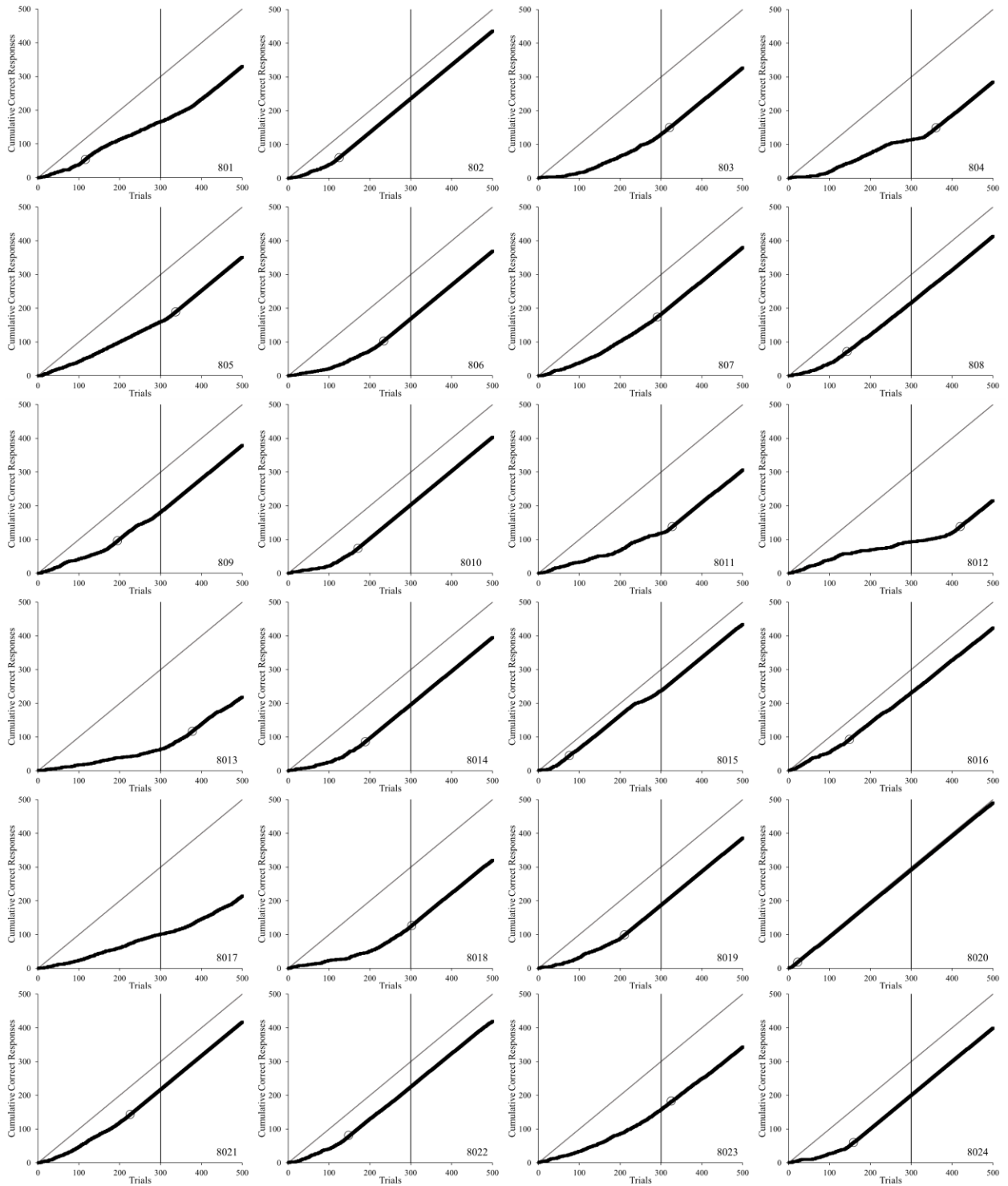
Cumulative Records

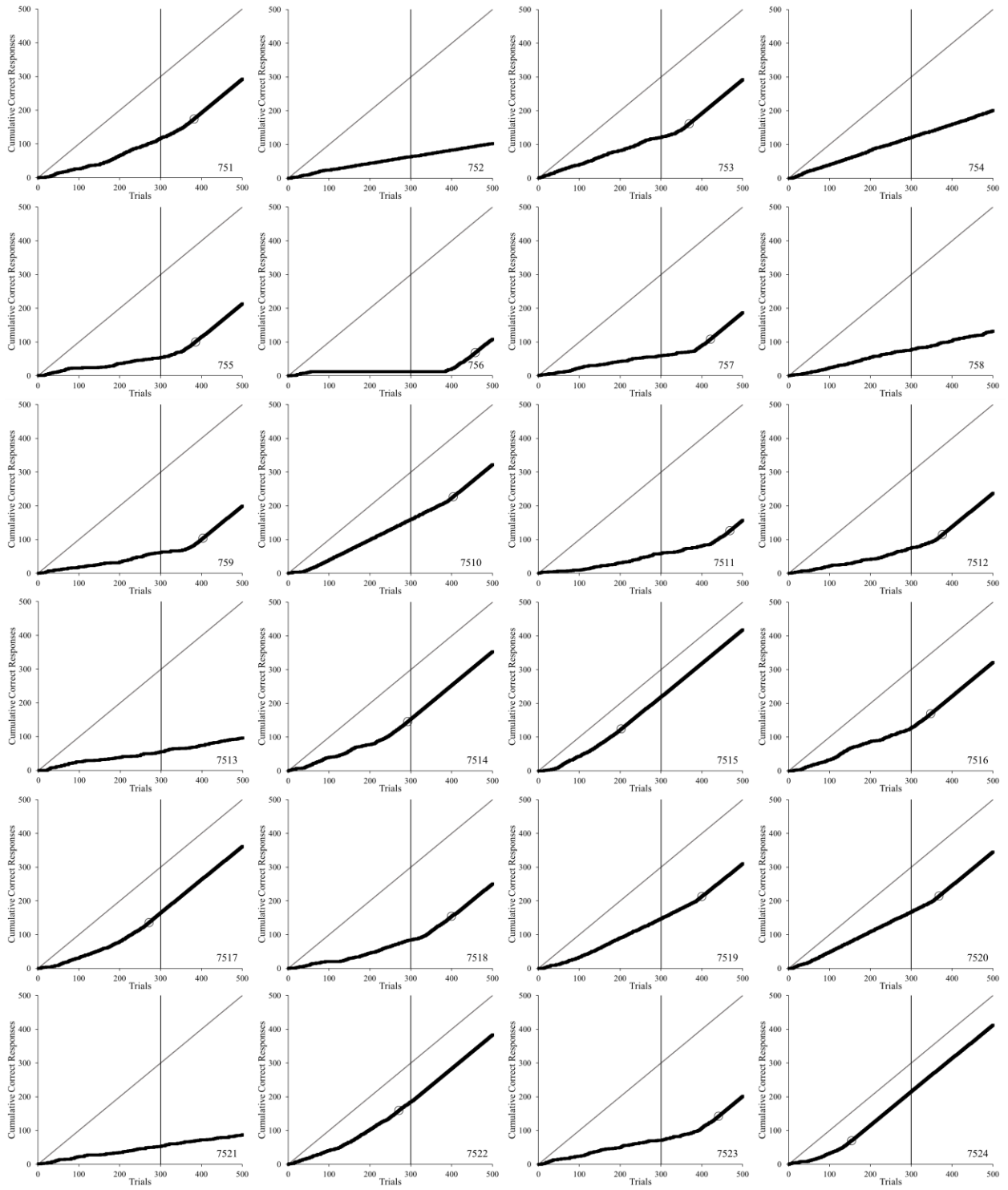


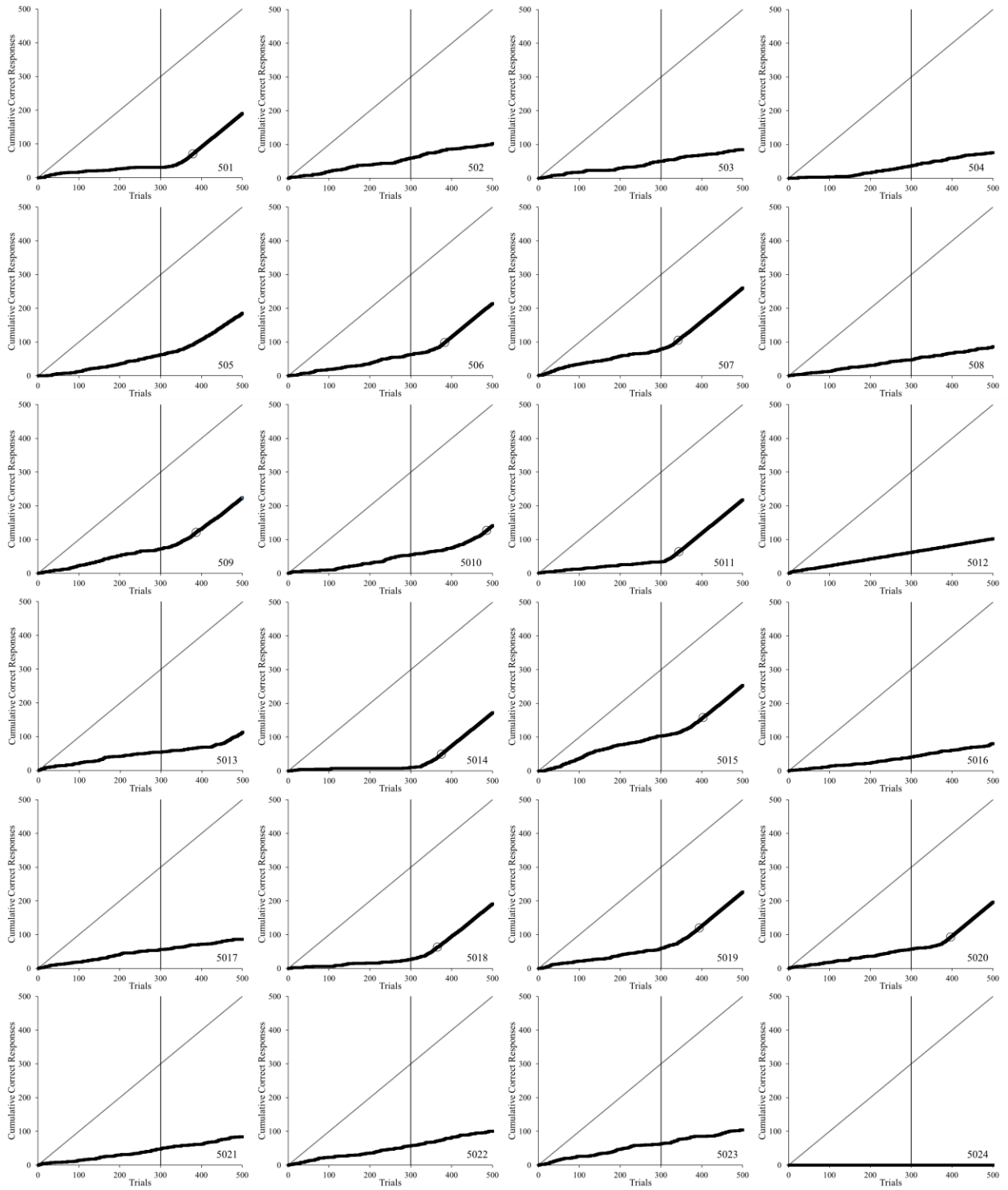












CURRICULUM VITAE

Samantha Bergmann

Place of birth: Saint Louis Park, MN

Education

B.A., University of Minnesota, May 2011
Major: Psychology and Spanish Studies, *summa cum laude*

M.A., Western Michigan University, May 2013
Major: Behavior Analysis

Ph.D., University of Wisconsin-Milwaukee, May 2018
Major: Psychology, Behavior Analysis
Minor: Developmental Psychopathology

Dissertation Title: Establishing Quality Standards for Applied Behavior Analytic Skill-Acquisition Interventions: A Translational Model with Undergraduate Students

CERTIFICATION AND LICENSE

Board Certified Behavior Analyst (1-13-14033), Behavior Analyst Certification Board
Licensed Behavior Analyst (84-140), Wisconsin Department of Safety and Professional Services

AWARDS AND HONORS

- 2017 Distinguished Dissertator Fellowship, University of Wisconsin-Milwaukee
- 2017 John and Lynn Schiek Project Award in Behavior Analysis, University of Wisconsin-Milwaukee
- 2017 John and Lynn Schiek Travel Award, University of Wisconsin-Milwaukee
- 2017 Graduate Student Travel Award, University of Wisconsin-Milwaukee
- 2016 Distinguished Graduate Student Fellowship, University of Wisconsin-Milwaukee
- 2016 Department of Psychology Summer Research Fellowship Award, University of Wisconsin-Milwaukee
- 2016 Summer School Institute Scholarship, Morningside Teachers' Academy
- 2016 John and Lynn Schiek Project Award in Behavior Analysis, University of Wisconsin-Milwaukee
- 2016 John and Lynn Schiek Travel Award, University of Wisconsin-Milwaukee
- 2016 Graduate Student Travel Award, University of Wisconsin-Milwaukee
- 2015 Forrest J. Files Student Research Award, Mid-American Association for Behavior Analysis
- 2015 John and Lynn Schiek Project Award in Behavior Analysis, University of Wisconsin-Milwaukee
- 2015 John and Lynn Schiek Travel Award, University of Wisconsin-Milwaukee

- 2015 Graduate Student Travel Award, University of Wisconsin-Milwaukee
- 2015 Best Presentation, Third Place, Association of Graduate Students in Psychology, University of Wisconsin-Milwaukee,
- 2014 John and Lynn Schiek Project Award in Behavior Analysis, University of Wisconsin-Milwaukee
- 2014 Dynamic Measurement Group Award, University of Oregon
- 2014 DIBELS Summer Training Institute Graduate Student Scholarship, Dynamic Measurement Group
- 2014 School Psychology Travel Grant, University of Oregon
- 2013 Dean's List, Western Michigan University
- 2012 Dean's List, Western Michigan University
- 2011 Dean's List, Western Michigan University

PUBLICATIONS

Peer-Reviewed Journal Articles

- 2017 **Bergmann, S.,** Kodak, T., LeBlanc, B. (2017). Effects of programmed errors of omission and commission during auditory-visual conditional discrimination training with typically developing children. *Psychological Record*. 67(1), 109-119. doi:10.1007/s40732-016-0211-2
- 2016 Kodak, T., Campbell, V., **Bergmann, S.,** LeBlanc, B., Kurtz-Nelson, E., Cariveau, T., Haq, S., Zemantic, P., & Mahon, J. (2016). Error-correction procedures to teach sight words and prepositions to children with autism spectrum disorder. *Journal of Applied Behavior Analysis*, 49(3), 532-547. doi:10.1002/jaba.310

Book Chapters

Kodak, T., **Bergmann, S.,** & Grow, L. (in press). Behavioral interventions for autism. In W. Fisher, C. Piazza, & H. Roane (Eds). *Handbook of Applied Behavior Analysis*, 2nd Ed. Guilford Press.

Encyclopedia Entries

- 2017 **Bergmann, S.,** & Kodak, T. (2017). Treatment Integrity in A. Wenzel (Ed.), *Encyclopedia of Abnormal and Clinical Psychology* (pp. 3652), Thousand Oaks: SAGE Publications, Inc. doi:10.4135/9781483365817.n1454

GRANTS

- 2017 Organization for Autism Research, Graduate Student Research Grant, University of Wisconsin-Milwaukee, 2017-2019, \$2,000.

CONFERENCE ACTIVITY/PARTICIPATION

Symposia

- 2018 **Bergmann, S.**, Kodak, T., Van Den Elzen, G., Cliett, T., Wood, R., & Doyle, H. (2018, May). Evaluating the Efficacy and Efficiency of Auditory Discrimination Procedures: A Translational Study with Two Typically Developing Children. Association for Behavior Analysis International Annual Convention, San Diego, California.
- 2018 Waite, M., Kodak, T., **Bergmann, S.**, & Fulton, C. (2018, March) Use of the Functional Analysis for Identifying the Function of Mouthing in Dogs. Interdisciplinary Forum for Applied Animal Behavior, New Orleans, Louisiana.
- 2017 **Bergmann, S.**, Kodak, T., Benitez, B, Van Den Elzen, G., Cliett, T., Knutson, S., Bohl, L., & Wood, R. (2017, May). Comparing the Efficacy and Efficiency of Strategies to Teach Auditory Discrimination. Association for Behavior Analysis International Annual Convention, Denver, Colorado.
- 2017 **Bergmann, S.**, Kodak, T., Van Den Elzen, G., Cliett, T., Harman, M., & Wood, R. (2017, May). Matrix Training: Considerations for Recombinative Generalization and Efficiency of Acquisition. Association for Behavior Analysis International Annual Convention, Denver, Colorado.
- 2017 Harman, M., Kodak, T., **Bergmann, S.**, LeBlanc, B., Baumann, C., Bohl, L., & Van Den Elzen, G. (2017, May). Examining the Effects of an Unrelated Auditory Stimulus on the Accuracy and Latency to Respond to a Mental Math Task. Association for Behavior Analysis International Annual Convention, Denver, Colorado.
- 2016 **Bergmann, S.**, Kodak, T., Mahon, J., LeBlanc, B., & Van Den Elzen, G. (2016, October). Matrix Training: Considerations for Recombinative Generalization and Efficiency of Acquisition. Berkshire Association for Behavior Analysis and Therapy conference, Amherst, Massachusetts.
- 2016 **Bergmann, S.**, Kodak, T., & LeBlanc, B. (2016, May). Examining the Parity Hypothesis with English-Speaking Undergraduate Students. Association for Behavior Analysis International Annual Convention, Chicago, Illinois.
- 2016 Kodak, T., **Bergmann, S.**, Fiske, K, Isenhower, R., Bamond, M., Delmolino Gatley, L., & Francis, C. (2016, May). A Replication and Extension of a Skills Assessment for Auditory-Visual Conditional Discrimination Training. Association for Behavior Analysis International Annual Convention, Chicago, Illinois.
- 2016 LeBlanc, B., Kodak, T., **Bergmann, S.**, Zettel, S., Benitez, B., Knutson, S., & Shannon-Jackson, A. (2016, May). Using Video Modeling to Teach Parents to Use the Natural Language Paradigm. Association for Behavior Analysis International Annual Convention, Chicago, Illinois.
- 2015 **Bergmann, S.**, Kodak, T., LeBlanc, B., (2015, October). Effects of Programmed Errors of Omission and Commission during Auditory-Visual Conditional Training with Typically Developing Children. Mid-American Association for Behavior Analysis, Kansas City, Missouri.
- 2015 **Moberg, S.**, Kodak, T., Campbell, V., Cariveau, T., Mahon, J., Ruppert, T., Rush, K., Kurtz-Nelson, E., (2015, May). Comparing the Efficiency of Error Correction Procedures and Assessing Children's Preference for Instruction. Association for Behavior Analysis International Annual Convention, San Antonio, Texas.
- 2015 LeBlanc, B., Kodak, T., **Moberg, S.**, Kammer, J., Haq, S., Zemantic, P., (2015, May). Comparing the Effects of Errors of Commission and Omission on Skill Acquisition. Association for Behavior Analysis International Annual Convention, San Antonio, Texas.

Poster Presentations

- 2017 **Bergmann, S.**, Kodak, T., Bamond, M., Fiske Massey, K., Isenhower, R. (2017, November). Extending and Assessment of Repertoires Related to Auditory-Visual Conditional Discriminations. Mid-American Association for Behavior Analysis, Milwaukee, Wisconsin.
- 2017 Olsen, M., Kodak, T., **Bergmann, S.**, Harman, M., Van Den Elzen, G., Costello, D., Halbur, M. (2017, November). Teaching Early Literacy Skills to Individuals with ASD Using Headsprout®. Mid-American Association for Behavior Analysis, Milwaukee, Wisconsin.
- 2017 Kodak, T., **Bergmann, S.**, Van Den Elzen, G., Knutson, S., Gorgan, E., Olsen, M., Harman, M., Benitez, B., & Halbur, M. (2017, November). Comparing Within-Stimulus Prompts to Teach Intraverbal Conditional Discriminations of Function. Mid-American Association for Behavior Analysis, Milwaukee, Wisconsin.
- 2017 Benitez, B., Kodak, T., **Bergmann, S.**, Knutson, S., Van Den Elzen, G., & Halbur, M. (2017, May). Addressing Stimulus Overselectivity during Tact Training with a Child with Autism Spectrum Disorder. Association for Behavior Analysis International Annual Convention, Denver, Colorado.
- 2017 Costello, D., **Bergmann, S.**, & Kodak, T. (2017, May). Increasing the Frequency and Duration of Eye Contact of a Child with Autism Spectrum Disorder. Association for Behavior Analysis International Annual Convention, Denver, Colorado. Mid-American Association for Behavior Analysis, Milwaukee, Wisconsin (2017, November).
- 2016 Benitez, B., Kodak, T., **Bergmann, S.**, LeBlanc, B., Harman, M., & Ayazi, M. (2016, May). Examination of Auditory and Visual Stimuli on the Accuracy and Latency to Respond in a Mental Arithmetic Task. Association for Behavior Analysis International 42nd Annual Conference, Chicago, Illinois.
- 2016 Knutson, S., **Bergmann, S.**, & Kodak, T. (2016, May). Using a Multiple Schedule to Reduce Classroom Problem Behavior. Association for Behavior Analysis International 42nd Annual Conference, Chicago, Illinois.
- 2016 Van Den Elzen, G., **Bergmann, S.**, & Kodak, T. (2016, April). Evaluating Procedures to Teach Auditory Discriminations with Children. Poster presented at the UW-System Undergraduate Research Symposium, Milwaukee, Wisconsin. 18th Annual Association for Graduate Students in Psychology Research Symposium, Milwaukee, WI (2016, April).
- 2016 Harman, M., LeBlanc, B., **Bergmann, S.**, Kodak, T., Baumann, C., Bohl, L., & Van Den Elzen, G. (2016, April). The Effects of Different Auditory Stimuli on the Completion of Math Problems. UW-System Undergraduate Research Symposium, Milwaukee, Wisconsin and 18th Annual Association for Graduate Students in Psychology Research Symposium, Milwaukee, WI (2016, April).
- 2016 LeBlanc, B., Kodak, T., **Bergmann, S.**, Zettel, S., & Benitez, B. (2015, October). Teaching Parents to use the Natural Language Paradigm with Video Modeling. Mid-American Association for Behavior Analysis annual conference, Kansas City, Missouri.
- 2015 **Bergmann, S.**, Kodak, T., LeBlanc, B., (2015, September). Effects of Programmed Errors of Omission and Commission during Auditory-Visual Conditional Training with Typically Developing Children. the Minnesota Northland Association for Behavior Analysis Conference, Minneapolis, MN.

- 2015 **Moberg, S.**, Kodak, T., LeBlanc, B., Ayazi, M., Gorgan, E., Liu, Z., (2015, March). An Experimental Analysis of Verbal Behavior: Effects of Auditory Stimuli on Accuracy and Latency to Respond on a Mental Math Task. Heartland Association for Behavior Analysis Annual Conference, Omaha, Nebraska.
- 2014 **Moberg, S.**, Kodak, T., Campbell, V., Cariveau, T., Ruppert, T., Rush, K., Kurtz-Nelson, E., (2014, May). Comparing the Effectiveness and Efficiency of Error Correction Procedures. Association for Behavior Analysis International Annual Convention, Chicago, Illinois.
- 2014 Ruppert, T., Kodak, T., Cariveau, T., Zemantic, P., **Moberg, S.**, Ledoux, M., (2014, May). An Assessment of Children's Preference for Error Correction Procedures. Association for Behavior Analysis International Annual Convention, Chicago, Illinois
- 2014 Zemantic, P., Porritt, M., **Moberg, S.**, Beattie, T., Barrett, E., Suarez Pedraza, C.S., Brandel, D., Meng, P., Hayes, M., Kodak, T., Good III, R.H., (2014, February). Examining Comprehension and Oral Reading Fluency for Students with Autism. National Association of School Psychologists Annual Convention, Washington, D.C.
- 2013 **Moberg, S.**, Kline, B., Korneder, J., & Malott, R. (2013, May). An Alternating Treatment Design with Matching-to-Sample: Error Correction and Errorless Learning with Time Delay. Association for Behavior Analysis International Annual Convention, Minneapolis, MN. Also presented at the meeting of the Association for Behavior Analysis International Autism conference, Portland, OR (2013, January).

CAMPUS OR DEPARTMENTAL TALKS

- 2018 **Bergmann, S.**, (2018, April). Establishing Quality Standards for Applied Behavior Analytic Skill-Acquisition Interventions: A Translational Model with Undergraduate Students. Behavior Analysis Colloquium Series, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin. Association of Graduate Students in Psychology Invited Research Colloquium, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin
- 2017 **Bergmann, S.**, (2017, October). Establishing Quality Standards for Applied Behavior Analytic Skill-Acquisition Interventions: A Translational Model with Undergraduate Students. Behavior Analysis Colloquium Series, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin
- 2015 **Moberg, S.**, Kodak, T., Campbell, V., Cariveau, T., Mahon, J., Ruppert, T., Rush, K., Kurtz-Nelson, E., (2015, April). Comparing the Efficiency of Error Correction Procedures and Assessing Children's Preference for Instruction. Association of Graduate Students in Psychology Invited Research Colloquium, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin
- 2014 **Moberg, S.**, Kodak, T., LeBlanc, B., & Mahon, J. (2014, November). The Assessment and Treatment of Selective Mutism: A Case Study. Behavior Analysis Colloquium Series, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin
- 2014 **Moberg, S.** (2014, February). Board Certified Behavior Analyst. In What to do with your psychology degree panel, University of Oregon, Eugene, Oregon

TEACHING EXPERIENCE

University of Wisconsin-Milwaukee

- 2017 Associate Lecturer Introduction to Psychology
- 2017 Guest Instructor Proseminar in Behavior Analysis: Verbal Behavior (Prof. Kodak)
- 2016 Associate Lecturer Introduction to Psychology
- 2015 Guest Lecturer Introduction to Psychology: Social Psychology (Prof. Smith)
- 2015 Associate Lecturer Introduction to Psychology

University of Oregon

- 2014 Guest Instructor Behavioral Assessment: Extinction, Differential Reinforcement, and Antecedent-based Strategies (Prof. Kodak)

Western Michigan University

- 2013 Graduate Student Instructor Research Assistant and Teaching Assistant Course
- 2012 Graduate Student Instructor Research Assistant and Teaching Assistant Course
- 2011 Graduate Student Instructor Introduction to Behavior Analysis

RESEARCH EXPERIENCE

- 2014 Graduate Student Researcher (2014-2018), University of Wisconsin-Milwaukee, Milwaukee, WI. Led research projects related to the assessment and treatment of auditory discriminations, instructional strategies to promote generalized learning, and early literacy skills.
- 2013 Research Assistant (2013-2014), Academic and Behavioral Intervention Clinic, University of Oregon, Eugene, OR. Studies focused on evaluating the efficiency of and preference for instructional methods, reducing repetitive play responses, and developing assessment tools.
- 2013 Data Manager (2013-2014), Early Literacy Research Team, University of Oregon, Eugene, OR. Examined oral reading fluency and comprehension with students diagnosed with an autism spectrum disorder.

SUPERVISION AND MENTORING EXPERIENCE

- 2015 Behavioral Treatment Licensed Supervisor (2015-2018), Center for Language Acquisition and Social Skills Interventions, Mequon, WI.
- 2014 Consultant (2014-2017), Douglass Developmental Disabilities Center, Rutgers University, New Brunswick, NJ. Collaborated with clinicians and researchers on topics related to conditional discrimination.
- 2014 Undergraduate Research Assistant Supervisor (2014-2017), Kodak Early Intervention Laboratory, University of Wisconsin-Milwaukee, Milwaukee, WI.
- 2012 Supervisor and Support Coordinator (2012-2013), Early Childhood Special Education Classroom, WoodsEdge Learning Center, Portage, MI.
- 2012 Undergraduate Thesis Mentor (2012-2013), Behavioral Science Program, Western Michigan University, Kalamazoo, MI. Mentored and advised a senior undergraduate student on the completion of a thesis to earn departmental honors.

- 2012 Advanced Practicum Student Supervisor (2012-2013), Early Childhood Special Education Classroom, WoodsEdge Learning Center, Portage, MI.
- 2011 Senior System Manager (2011-2013), Behavioral Research Supervisory System, Western Michigan University, Kalamazoo, MI. Supervised graduate and undergraduate students in the completion of system tasks and continuous quality improvements.

CLINICAL EXPERIENCE

- 2015 Board Certified Behavior Analyst (2015-2017), Center for Language Acquisition and Social Skills Intervention, Mequon, WI.
- 2015 Consultant, Children’s Learning Center, University of Wisconsin-Milwaukee, Milwaukee, WI. Collaborative consultation to help address teachers’ concerns regarding academic and behavioral difficulties of students.
- 2013 Clinical Case Lead (2013-2014), Academic and Behavioral Interventions Clinic, University of Oregon, Eugene, OR.
- 2012 Clinical Intern (2012-2013), Neuropsychology Associates, Kalamazoo, MI. Assisted in patient intake and follow-up for individuals with traumatic brain injuries, learning disabilities, developmental disabilities, organic-based cognitive disabilities, and psychological disorders. Conducted supervised psychotherapy sessions, monitored patient progress as assigned.
- 2011 Behavior Technician (2011-2012), Early Childhood Special Education Classroom, WoodsEdge Learning Center, Portage, MI.

ASSISTANTSHIPS

- Program Assistant**, Kodak Early Intervention Laboratory, Semester I and II, 2015-2016
- Program Assistant**, Kodak Early Intervention Laboratory, Semester I and II, 2014-2015

SERVICE TO PROFESSION

- 2018 Guest reviewer, *Journal of Applied Behavior Analysis*
- 2018 Chair. (2018, May). Assessing Procedural Variations to Evaluate Efficacy and Efficiency of Conditional Discrimination Interventions. Association for Behavior Analysis International Annual Convention, San Diego, California.
- 2017 Chair. (2017, November). Resurrection of the Animal Lab in Behavior Analysis Training Programs through the use of Invertebrates. Mid-American Association for Behavior Analysis, Milwaukee, Wisconsin.
- 2017 Chair. (2017, May). The Science of Skinner’s Analysis of Verbal Behavior: Theory and Basic Research. Association for Behavior Analysis Annual Convention, Denver, Colorado.
- 2017 Guest reviewer, *Journal of Applied Behavior Analysis*
- 2016 Guest reviewer, *Journal of Applied Behavior Analysis*
- 2016 Guest reviewer, *Learning and Motivation*
- 2016 Chair. (2016, May). Recent Advancements in Caregiver and Staff Training. Association for Behavior Analysis International Annual Convention, Chicago, Illinois

DEPARTMENTAL/UNIVERSITY SERVICE

2013 Faculty-Appointed Assistant Admissions Liaison (2013-2014), Department of School Psychology, University of Oregon

EXTRACURRICULAR UNIVERSITY SERVICE

2016 Vice President (2016-2017), Association of Students in Behavior Analysis, University of Wisconsin-Milwaukee

2015 Vice President (2015-2016), Association of Graduate Students in Psychology, University of Wisconsin-Milwaukee

2014 Secretary (2014-2015), Association of Graduate Students in Psychology, University of Wisconsin-Milwaukee

PROFESSIONAL MEMBERSHIPS/AFFILIATIONS

Association for Behavior Analysis International
Mid-America Association for Behavior Analysis