2017

An Evaluation of the Effects of Treatment Integrity Errors on Skill Acquisition During Discrete-Trial Instruction

Jennifer M. Owsiany

An Evaluation of the Effects of Treatment Integrity Errors on Skill Acquisition During Discrete-Trial Instruction

Jennifer M. Owsiany, B.A.

Thesis submitted to the
Eberly College of Arts and Sciences
at West Virginia University

in partial fulfillment of the requirements for the degree of

Master of Science in
Psychology

Regina A. Carroll, Ph.D., BCBA-D, Chair
Michael Perone, Ph.D.
Amy Herschell, Ph.D.

Department of Psychology

Morgantown, West Virginia
2017

Keywords: autism spectrum disorder, discrete-trial instruction, skill acquisition, treatment integrity

Abstract

An Evaluation of the Effects of Treatment Integrity Errors on Skill Acquisition During Discrete-Trial Instruction

Jennifer M. Owsiany

Previous research has demonstrated that delivering controlling prompts and reinforcers with low-levels of treatment integrity may interfere with the acquisition of skills during discrete-trial instruction. However, implementing certain components of discrete-trial instruction with varying levels of low integrity may influence skill acquisition differently. In Experiment 1, we evaluated the effects of delivering a prompt following participants' incorrect responses with 25%, 50%, 75%, and 100% integrity on skill acquisition with two children with autism spectrum disorder. In Experiment 2, we evaluated the effects of delivering a reinforcer following participants' correct responses with 25%, 50%, 75%, and 100% integrity on skill acquisition with the same two children from Experiment 1. Participants acquired the target skills taught in the high-integrity condition and in the low-integrity conditions in both experiments. The results of the current study suggest that implementing components of discrete-trial instruction with low levels of integrity may not interfere with skill acquisition for these two participants. These results are partially inconsistent with previous research. Future research should continue to investigate the effects of implementing various components of discrete-trial instruction with low treatment integrity with children with autism spectrum disorder.

Acknowledgements

I would like to thank my advisor and mentor, Dr. Regina Carroll for her guidance and support while I was completing this project. It would not have been possible without her assistance. I would also like to thank Dr. Carroll for the vast amount of knowledge that she imparted on me during my time at West Virginia University, as both an undergraduate and graduate student. I would also like to thank Dr. Mike Perone and Dr. Amy Herschell for serving on the committee for this project and for providing me with helpful feedback. Finally, I would like to thank my fellow lab mate, Natalie Jones, and former undergraduate research assistant, Ashley Felde, for assisting with data collection for this project.

Table of Contents

List of Tables

Tables

List of Abbreviations

ASD	autism spectrum disorder
DTI	discrete-trial instruction
EVT-2	Expressive Vocabulary Test – 2nd Edition
PPVT-4	Peabody Picture Vocabulary Test – 4th Edition
VB-MAPP	Verbal Behavior Milestones Assessment and Placement Program

An Evaluation of the Effects of Integrity Errors on Skill Acquisition During Discrete-Trial Instruction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by deficits in social communication and language, restricted interests, and repetitive behaviors (American Psychiatric Association, 2013). ASD is becoming increasingly prevalent in the United States. Approximately one in every 68 children in the United States is diagnosed with ASD each year, with males tending to be diagnosed more often (1 in 42) than females (1 in 189; Center for Disease Control and Prevention, 2015).

There are several treatment models that can be used to treat ASD, however, early intensive behavioral intervention is the most widely researched and requested treatment model (Green et al., 2006). The main components of early intervention include (a) highly structured, one-on-one teaching strategies (e.g., discrete trial instruction; DTI), (b) supervised treatment from therapists trained in applied behavior analytic procedures, (c) treatment that is provided for a long duration and high intensity (i.e., between 20 and 40 hours per week for up to four years), and (d) treatment which occurs in the home, school, or a combination of the two (Lovaas, 1987; Reichow & Wolery, 2009; Reichow, Barton, Boyd, & Hume, 2012). Reichow (2012) conducted an overview of five meta-analyses that have been conducted on the effectiveness of early intensive behavioral intervention for children with ASD and found that children with ASD who received early intensive behavioral intervention showed gains in IQ, adaptive behaviors, or both, when compared with other forms of treatment (e.g., less intensive, workshop-based therapies). The current body of literature shows strong support for the use of early intensive behavioral intervention as the optimal treatment method for children with ASD.

Discrete-Trial Instruction

DTI is an empirically supported instructional procedure and is one of the main components of early intensive behavioral intervention (e.g., Reichow & Wolery, 2009). There are five main components of DTI, which are (a) a clear instruction (e.g., a therapist says, "Read the word" when presenting a word card); (b) a controlling prompt (e.g., therapist models the correct response for the child); (c) the child's response (e.g., a correct or incorrect answer following the therapist's instruction); (d) a consequence following the child's response (e.g., the therapist may provide praise and deliver a preferred item following a correct response); and (e) a brief pause between trials (e.g., a 1- to- 2-s pause between each trial; Smith, 2001).

Therapists can use DTI to teach a variety of skills, such as language skills (e.g., Geiger et al., 2012) and play skills (e.g., Grow, Kodak, & Clements, 2016). While therapists typically implement DTI in a one-on-one, teacher-to-student format, it can also be implemented in small-group formats (e.g., Taubman, Brierly, Wishner, McEachin, & Leaf, 2001). Children with ASD have shown improvements in daily living, academic, and communication skills in early intervention settings where the primary teaching procedure is DTI (e.g., Weiss, 1999; Howard, Sparkman, Cohen, Green, & Stanislaw, 2005; Rivard, Terroux, & Mercier, 2014). Previous research suggests that while DTI is an effective teaching procedure for children with ASD, the effectiveness of DTI may depend in part on the extent to which therapists implement the components of DTI accurately (Symes, Remington, Brown, & Hastings, 2006).

Treatment Integrity

The extent to which a therapist implements an instructional procedure, like DTI, accurately is referred to as treatment integrity (Peterson, Homer, & Wonderlich, 1982). Previous research has evaluated the influence of treatment integrity errors on responding with behavioral

interventions such as interventions related to decreasing challenging behavior (e.g., Dib & Sturmey, 2007; Shillingsburg, Bowen, & Shapiro, 2014) and increasing appropriate behavior (e.g., Geiger et al. 2012; Grow et al., 2016). Implementing instructional procedures with high treatment integrity may help to ensure optimal learning outcomes for children with ASD (Symes et al., 2006). However, if a therapist does not implement instructional procedures accurately (i.e., with low treatment integrity), it may be difficult for them to draw accurate conclusions about the effectiveness of those procedures. For example, it may be difficult for a therapist to determine if a child learned as a direct result of a teaching procedure. Implementing instructional procedures with low treatment integrity may prohibit therapists from making necessary changes to instructional procedures (Vollmer, Sloman, & St. Peter Pipkin, 2008).

Previous research suggests that it is important to implement instructional procedures with high integrity, however professionals that implement these procedures may not always do so with high integrity (e.g., Carroll, Kodak, & Fisher, 2013). Carroll and colleagues conducted a series of studies to identify the most common treatment integrity errors and subsequently assessed the influences of those integrity errors on the acquisition of skills during DTI with children with ASD. Specifically, in Study 1, they observed teachers and paraprofessionals providing one-on-one or small-group instruction to children with ASD to determine the most commonly made treatment-integrity errors. During each observation data were collected on (a) making sure the child was ready, (b) providing a clear and concise instruction, (c) delivering a prompt following a child's incorrect response, (d) delivering a reinforcer following a child's correct response, and (e) responding to a child's problem behavior. The results from Study 1 showed that the three most commonly made integrity errors by teachers and paraprofessionals

were providing an unclear instruction, withholding reinforcement following a correct response, and withholding a prompt following an incorrect response.

Treatment integrity errors with a controlling prompt. A treatment integrity error with a controlling prompt may consist of a therapist withholding a prompt that would help the child respond correctly following an error. Previous research has evaluated the influence of prompting errors on the acquisition or improvement of skills and has found that prompting errors interfere with learning (e.g., Holcombe, Wolery, & Snyder, 1994; Noell, Gresham, & Gansle, 2002; Carroll et al., 2013). For example, Holcombe and colleagues compared the acquisition of new skills across a high-integrity condition and a low-integrity condition for six pre-school aged children with intellectual disabilities. During the high-integrity condition, the therapist delivered a prompt (i.e., the therapist modeled the correct response following the child's incorrect response) on all trials in a session. During the low-integrity condition, the therapist withheld the controlling prompt on 50% of the trials in a session. That is, following an incorrect response, the therapist ended the trial. During both conditions, the therapist provided praise following a correct response and a tangible item at end of the session. Five of the six participants acquired the skill taught during the high-integrity condition, and four of the six participants acquired the skill taught in the low-integrity condition. Although four of the six participants acquired the skill taught in the low-integrity condition, learning was more efficient in the high-integrity condition (i.e., three of these four participants acquired the skill in the high-integrity condition following fewer training sessions compared to the low-integrity condition). These results suggest that delivering a prompt with low treatment integrity may interfere with the acquisition of skills. However, Holcombe and colleagues only assessed the influence of prompting errors that occurred on 50% of the trials on skill acquisition. It is possible that programmed prompting

errors during fewer trials (i.e., 33% of trials) would be less likely to influence the acquisition of new skills.

In an extension of Holcombe et al. (1994), Noell et al. (2002) evaluated the improvement of math skills in a high-integrity condition and two low-integrity conditions with six typically developing elementary-school children using a computer program. During the high-integrity condition, if the children did not respond within 2 s of the presentation of the problem, the experimenters programmed a prompt to occur. The experimenters programmed prompts to keep occurring until either the child entered in a response, or the correct solution was given. During the low-integrity conditions, the experimenters programmed prompts to occur on either 67% of trials or 33% of trials. In all conditions, the experimenters programmed feedback (i.e., signals that the child's answer was correct or incorrect) following every correct and incorrect response. All six participants showed an increase in the number of math problems answered correctly during the high-integrity condition and four participants showed decreases in the number of math problems answered correctly in both of the low-integrity conditions. These results suggest that programmed prompting errors may decrease performance on established skills. Although these results suggest that programmed prompting errors may interfere with the improvement of an established skill, direct comparisons cannot be made about the effects of programmed prompting errors on the acquisition of new skills.

Similar studies to Holcombe et al. (1994) and Noell et al. (2002) have shown that prompting errors may decrease the effectiveness of instructional procedures (e.g., Grow et al., 2009; Carroll et al., 2013). However, the delivery of a controlling prompt is only one of the five main components of DTI. It is possible that implementing other components of DTI with low

treatment integrity, such as the delivery of a reinforcer following a correct response, may influence the acquisition of skills.

Treatment integrity errors with delivery of reinforcers. A treatment integrity error with the delivery of reinforcers may consist of a therapist withholding reinforcement following correct responses, or therapists delivering reinforcement following incorrect responses. Previous research has evaluated the influence of programmed reinforcement errors on the acquisition of new skills during DTI (e.g., DiGennaro Reed et al., 2011; Jenkins et al., 2015; Carroll et al., 2016). For example, Jenkins et al. evaluated the influence of delivering a reinforcer following an incorrect response when four young children with ASD were taught to identify Japanese characters. During the high-integrity condition, the therapist only delivered a reinforcer following correct responses. During the two low-integrity conditions, the therapist delivered a reinforcer following either 50% of incorrect responses or 100% of incorrect responses.

All four participants in Jenkins et al. (2015) acquired the skill taught in the high-integrity condition. Three of the four participants did not acquire the skill taught in the low-integrity conditions. The therapist re-taught the skills previously exposed to low integrity to the three participants who did not acquire the skill taught in the low-integrity conditions. This enabled Jenkins and colleagues to evaluate the effects of previous exposure to programmed reinforcement errors on skill acquisition. Two of the three participants took longer to acquire the skills previously exposed to low treatment integrity after they were re-taught with high integrity. These results suggest that previous exposure to low integrity may delay skill acquisition. While Jenkins and colleagues evaluated the effects of programmed reinforcement errors on skill acquisition on 100% and 50% of trials, it is possible that more (i.e., 75%) or less (i.e., 25%) errors will have different effects on skill acquisition. Jenkins and colleagues also delivered

reinforcers following incorrect responses. It is unclear whether these same effects would be obtained had they withheld the delivery of reinforcers following correct responses. Although these results provide further insight into the potential influences of teaching with low treatment integrity on skill acquisition, the effects of programmed reinforcement errors on the acquisition of new functional skills remains unclear. It is also possible that a therapist can make a treatment integrity error with more than one component of DTI on a single trial in a session, which warrants further investigation.

Combined treatment integrity errors during DTI. It is possible for therapists to simultaneously make errors with different components of DTI. For example, a therapist might withhold a controlling prompt on one trial following an incorrect response and withhold a reinforcer following a correct response on another trial. Carroll et al. (2013) conducted a series of studies to identify common treatment integrity errors and to evaluate the influences of those errors on skill acquisition with children with ASD.

Based on the results from Study 1, Carroll et al. (2013) conducted Study 2 with six children with ASD. In this study, they compared the acquisition of new skills during a high-integrity condition and a low-integrity condition. During the high-integrity condition, the therapist did not make any integrity errors. That is, they provided a prompt following a child's incorrect response, delivered a reinforcer following a correct response, and presented a clear and concise instruction on every trial. During the low-integrity condition, the therapist withheld a prompt following an incorrect response, withheld a reinforcer following a correct response, and presented an additional instruction that differed from the protocol on 67% (8 of 12) of trials. On the remaining 33% (4 of 12) of trials in the low-integrity condition, the therapist did not make any integrity errors. All six participants acquired the skills taught in the high-integrity condition. Only one of

the six participants acquired the skills taught in the low-integrity condition, and for this participant, it took twice the number of training sessions compared to the high-integrity condition. The other five participants did not acquire the skills taught in the low-integrity condition. Carroll and colleagues subsequently taught the skills initially taught in the low-integrity condition with high-integrity procedures. Following this training, the other five participants acquired these skills. The results of Study 2 suggest that implementing various components of DTI with low treatment integrity simultaneously may impair skill acquisition; but the individual influence of low treatment integrity for each component of DTI on skill acquisition remains unclear.

In Study 3, Carroll et al. (2013) evaluated the effects of the three most common integrity errors on skill acquisition individually with three children diagnosed with ASD. The high-integrity condition was implemented with procedures identical to those in Study 2. Each of the three integrity errors was programmed to occur individually on 67% (8 of 12) of trials in a session. That is, instead of one low-integrity condition, there were three low-integrity conditions, one for each type of integrity error. The three low-integrity conditions in Study 3 were low-integrity prompt (i.e., the therapist withheld a prompt on 67% of trials following an incorrect response), low-integrity reinforcement (i.e., the therapist withheld a reinforcer on 67% of trials following a correct response), and low-integrity instruction (i.e., the therapist presented an additional instruction that differed from the protocol). Two of the three participants acquired the skills taught in the high-integrity condition and in all low-integrity conditions. All of the participants acquired skills in the least amount of time in the high-integrity condition. One participant took the longest time to acquire the skills taught in the low-integrity prompt condition, while another participant took the longest time to acquire the skills taught in the low-

integrity reinforcement condition. The third participant did not acquire the skills taught in the low-integrity prompt and low-integrity instruction conditions until they were taught with high-integrity procedures. These results suggest that implementing different components of DTI, such as the delivery of prompts and reinforcers, with low-treatment integrity may differentially influence skill acquisition across individual learners.

Carroll et al. (2013) programmed prompting and reinforcement errors to occur on only 67% of trials in a session. As previously stated, it is possible that programming more or less errors in a session may influence skill acquisition differently. It may be necessary to conduct an analysis of a range of programmed prompting and reinforcement errors, given the results of Carroll and colleagues' third study, where implementing different components of DTI at low integrity (i.e., 67%) influence each child's skill acquisition differently.

Purpose

The existing literature (e.g., Holcombe et al., 1994; Noell, et al., 2002; Carroll, et al., 2013; Jenkins et al., 2015; Carroll et al., 2016) suggests that implementing components of DTI (e.g., delivering a controlling prompt and delivering reinforcers) with low treatment integrity impairs skill acquisition for learners with ASD. The current study replicates previous research by evaluating the influences of prompting errors and reinforcement errors on skill acquisition with children with ASD during DTI. However, previous research has only evaluated prompting errors that occur during 50% or 67% of trials, it is unclear whether more or less programmed errors would influence skill acquisition differently. Thus, the purpose of Experiment 1 was to evaluate the effects of varying levels (i.e., 25%, 50%, 75%, and 100%) of prompting errors on the acquisition of new skills with children with ASD during DTI. Previous research has evaluated programmed reinforcement errors on 0%, 50%, and 100% of trials in a session; thus, it remains

unclear whether different amounts of programmed errors with the delivery of reinforcers following correct responses (e.g., 25%) would influence skill acquisition differently. In addition, previous research that has evaluated programmed reinforcement errors (e.g., Jenkins et al., 2015) delivered reinforcers following incorrect responses. Thus, it is unclear how skill acquisition would be influenced if a therapist withheld a reinforcer following a correct response. The purpose of Experiment 2 was to evaluate the effects of varying levels (i.e., 25%, 50%, 75%, and 100%) of programmed reinforcement errors, where the therapist withholds a reinforcer following a correct response, on skill acquisition with children with ASD during DTI.

General Method

Participants, Settings, and Materials

Three children with diagnoses of ASD participated in Experiment 1. Two participants from Experiment 1 also participated in Experiment 2. Chad was a 7-year 11-month old male who communicated using four- to five-word phrases. Prior to the start of the study, he had 3-years 6-months of experience with DTI. Ulysses was a 3-year 5-month old male who communicated using two- to three-word phrases. Prior to the start of the study, Ulysses had 6 months of experience with DTI. Hannah was a 5-year 9-month old female who did not communicate vocally. Prior to the start of the study, Hannah had 2-years 9-months of experience with DTI. Before the study began, we conducted language assessments with each participant, including the Peabody Picture Vocabulary Test-4 (PPVT-4; Dunn & Dunn, 2007), the Expressive Vocabulary Test-2 (EVT-2; Williams, 2002); and the Verbal Behavior Milestones Assessment Placement Program (VB-MAPP; Sundberg, 2008).

The PPVT-4 is a standardized language assessment that measures an individual's ability to identify common objects and their characteristics, actions, and emotions (Dunn & Dunn,

2007). During the PPVT-4, the test administrator does not require individuals to respond vocally. That is, individuals typically respond by touching or pointing to pictures. The EVT-2 is a standardized language assessment that measures an individual's ability to vocally label common objects and their characteristics, actions, and emotions (Williams, 2002). The VB-MAPP is a language assessment that therapists use to assess an individual's language skills and to monitor language acquisition (Sundberg, 2008). The VB-MAPP consists of four assessments, however, we only used the Milestones Assessment for the current study. We used the Milestones Assessment to evaluate 170 language skills. Based on the results of the Milestones Assessment, we categorized each participant into one of three developmental age ranges (i.e., 0-30 months, 18-30 months, and 30-48 months).

We used the results of each participant's language assessment (see Table 1) to identify potential target skills (i.e., the specific skills we taught the participants). Based on the results of the language assessments, we taught Chad to identify categories of items and Ulysses to identify the functions of items and categories of items during Experiment 1 (see Table 2). Hannah moved to a different state while we were conducting Experiment 1. Thus, we were unable to complete the study with Hannah. During Experiment 2, we taught Chad to answer questions about items and activities and Ulysses to identify features of items (see Table 3).

The therapist conducted sessions either in a private room of an early intensive behavioral intervention clinic (Chad) or in a university-based research laboratory (Ulysses). Both rooms were equipped with a table, chairs, and all materials necessary to conduct sessions (e.g., a camera and tripod, data sheets and timers, preferred items, and picture cards). When we taught Ulysses to identify functions of items during Experiment 1, we printed colored pictures of common items on white 102 mm by 172 mm cards.

Experiment 1: A Parametric Analysis of Prompting Errors During Discrete-Trial Instruction

Previous research (e.g., Holcombe et al., 1994; Carroll et al., 2013) has evaluated the effects of prompting errors with 50-67% integrity on skill acquisition. It is possible that programming prompting errors on more or less trials in a session will influence skill acquisition differently. The purpose of Experiment 1 was to evaluate skill acquisition when a therapist taught target skills with high integrity and varying levels of low integrity (i.e., 25%, 50%, and 75% Integrity) with the delivery of a prompt during DTI.

Dependent Measures and Data Collection

During each session, we collected data on the following participant responses (a) correct responses, defined as the participant providing a pre-determined correct vocal response within the allotted prompt delay (e.g., the child says, “drive” when shown a picture of a car and asked, “What do you do with it?”); (b) prompted responses, defined as the participant providing the correct vocal response following a vocal model of the correct response; (c) incorrect responses, defined as the participant providing a vocal response that is different than the pre-determined correct response; and (d) no responses, defined as the participant not responding within the prompt delay. We also collected data on problem behavior for each participant. We defined problem behavior specifically for each participant. For example, aggression (Ulysses) was defined as hitting, kicking, scratching, biting, pinching or head-butting another person or throwing objects within one foot of another person. We converted each dependent measure to a percentage of trials by dividing the number of trials with a participant response by the total number of trials in a session, and multiplying by 100. We measured the total number of sessions and total training time required for a participant to reach a pre-determined mastery criterion.

During low-integrity sessions, we programmed prompting errors to occur during 25% to 75% of trials in a session. However, the therapist could only implement a prompting integrity error on trials that the participant engaged in an incorrect or no response. Thus, the actual percentage of trials with treatment-integrity errors was sometimes lower than the programmed percentage of prompting errors in a given session. For each low-integrity session, we calculated the actual obtained percentage of prompting errors by dividing the number of trials that the participant engaged in an incorrect or no response by the number of trials when a therapist delivered a prompt, and multiplying by 100.

Interobserver Agreement and Treatment Integrity

A secondary observer independently collected data on participant responses for an average of 63% (range, 46% to 89%) of sessions across conditions for each participant. We compared the primary and secondary observers' data on a trial-by-trial basis. We scored an agreement if both observers scored the same participant response (e.g., both observers scored the participant's response as correct). We scored a disagreement if each observer scored a different participant response (e.g., the primary observer scored the participant's response as correct and the secondary observer scored the participant's response as incorrect). We calculated interobserver agreement by taking the number of trials with an agreement in a session, dividing it by the number of agreements plus disagreements, and multiplying by 100. The average interobserver agreement for Chad was 99% (range, 92% to 100%) and the average interobserver agreement for Ulysses was 98% (range, 80% to 100%).

The secondary observer also collected data on the therapist's correct implementation of the teaching protocol for an average of 63% (range, 46% to 89%) of sessions across all conditions for each participant. During all conditions, the therapist's correct implementation of a

trial included (a) securing attention, defined as waiting to present the instruction until the participant looked at (i.e., made eye contact with) each picture card or was sitting upright and oriented toward the table; and (b) presenting the instruction exactly as outlined in the teaching protocol. During baseline, maintenance, and control trials, correct implementation of a trial included (a) withholding a prompt for incorrect or no responses, defined as not delivering a vocal model following an incorrect or no response; (b) withholding reinforcement for correct responses, defined as not delivering descriptive praise and access to a preferred item following a correct response; and (c) providing opportunities to earn reinforcement for mastered tasks, defined as presenting a mastered task six times during a session and delivering descriptive praise and brief access (i.e., 25 s) to a preferred item following a correct response.

During high-integrity trials, correct implementation of a trial included (a) delivering a prompt following an incorrect or no response, and (b) delivering praise and brief access to a preferred item following a correct response. During low-integrity trials, correct implementation of a trial included (a) withholding a prompt following an incorrect or no response and (b) delivering praise and brief access to a preferred item following a correct response. We scored a trial as either correct (i.e., implemented with 100% integrity) or incorrect (i.e., implemented with less than 100% integrity). We calculated treatment integrity for each session by taking the number of trials implemented correctly, dividing it by the total number of trials in a session, and multiplying by 100. Average treatment integrity for Chad's sessions was 99% (range, 92% to 100%) and average treatment integrity for Ulysses sessions was 99% (range, 92% to 100%).

Preference Assessment

Before the first session of each day, the therapist conducted a brief multiple stimulus without replacement preference assessment (Higbee, Carr, & Harrison, 2000) to identify

preferred edible and tangible items to use during teaching. During the preference assessment, the therapist set out either five edible items or five tangible items in front of the participant. The therapist then secured the participant's attention (i.e., ensured the participant looked at each item), labeled each item, and presented the instruction, "Pick one." Following a selection, the therapist allowed the participant brief access (i.e., 15 s) to the item. After the participant interacted with the item they selected, the therapist removed that item. For Chad, the therapist continued to present the remaining items until he selected three items. The therapist used the three items Chad chose as preferred items during sessions. For Ulysses, the therapist used the first tangible item and the top three edible items he selected as preferred items during sessions. We conducted additional preference assessments following every two to three sessions or following signs that the current items were no longer preferred (e.g., the participant requested other items or stopped manipulating the items provided).

Experimental Design

We used an adapted alternating treatments design (Sindelair, Rosenberg, & Wilson, 1985) to compare skill acquisition across high-integrity prompt, low-integrity prompt, and control conditions for two children with ASD. Specifically we compared the effects of a therapist delivering a prompt with high treatment integrity (i.e., 100% Integrity), varying levels of low treatment integrity (i.e., 25%, 50%, and 75% Integrity), and not directly teaching the target skills (i.e., control condition) on skill acquisition.

General Procedures

Pretest and target assignment. We conducted a pretest to ensure that the participants did not know the target skills before the start of the experiment. At the start of a trial, the therapist secured the participant's attention and presented a potential target skill (e.g., the

therapist held up a picture of a car and asked, "What do you do with it?"). If the participant responded correctly, (i.e., said "Drive") the therapist did not provide praise or deliver a preferred item and ended the trial. If the participant responded incorrectly, the therapist did not deliver a prompt. However, the therapist provided praise and 25-s access to a preferred item following approximately every two trials for appropriate session behavior (e.g., quiet hands, defined as sitting up with their hands resting on the table in front of them).

For each participant we identified 15 target skills that were not associated with correct responding during the pretest. We took several steps to equate the difficulty of the target skills across conditions. For all target skills for which we required a vocal response, we conducted an echoic assessment to identify target skills that the participant echoed incorrectly or inconsistently. During the echoic assessment, the therapist presented a vocal model of one of the target skills (e.g., "drive"). If the participant correctly echoed the therapist's model, then the therapist provided immediate descriptive praise (e.g., "Awesome, drive!") and brief access to a preferred item. The therapist presented each target skill at least three times during the echoic assessment. We excluded any target skills that the participant had difficulty echoing, echoed inconsistently, or that sounded too similar to another target skill. We also assigned target skills with similar numbers of syllables to each condition. For all target skills that required picture cards (Ulysses only), we ensured that the objects in the pictures did not look similar within and across conditions. We assigned three target skills to each condition.

Teaching. We presented target skills an equal number of times during a 12-trial session. The therapist conducted two sessions for the high- and low-integrity conditions and up to one control session per day. The therapist conducted sessions with Chad and Ulysses two days a

week. Thus, the maximum number of sessions that the therapist conducted a week equaled 16 for each participant.

The therapist used a constant prompt-delay procedure (Gast, Ault, Wolery, Doyle, & Belanger, 1988) to teach target skills. At the start of teaching, the therapist conducted two 0-s prompt-delay sessions with target skills in the high-integrity and low-integrity conditions. Following two 0-s prompt-delay sessions, the therapist increased the delay between the presentation of the instruction and the delivery of the prompt to a 2-s prompt delay (Chad and Ulysses Sets 1 and 2) or a 5-s prompt delay (Ulysses Set 3).

Mastery and early-termination criteria. We considered a set of target skills mastered once the participant responded correctly on at least 92% (11 of 12) of trials for two consecutive sessions. We also applied an early-termination criterion, similar to that of Carroll et al. (2013), for target skills that the therapist was still teaching following mastery of target skills in another condition. Specifically, the therapist discontinued conducting sessions in a condition that reached twice the number of sessions required for the participant to master target skills in the first condition.

Within-participant replication. We conducted within-participant replications with Ulysses with two additional sets of target skills. We used the same pretest procedures as described above to identify the additional target skills.

Procedures

Baseline. The purpose of the baseline condition was to ensure that the participants did not know the target skills before teaching. At the start of a trial, the therapist presented a picture card (e.g., a picture of a car; Ulysses only), secured the participant's attention, and presented the instruction (e.g., "What do you do with it?"). The therapist gave the participant 5 s to respond

following the instruction. If the participant responded correctly, incorrectly, or did not respond, the therapist ended the trial. The therapist did not provide any differential consequences for correct, incorrect, or no responses. During baseline sessions, the therapist presented a mastered task approximately once following every two trials. For each participant mastered tasks were skills that the participant reliably engaged in correct responses. For Chad, mastered tasks were labeling animals, shapes, and colors. For Ulysses, mastered tasks were labeling common objects, animals, shapes, and colors. If the participant responded correctly to a mastered task, the therapist provided praise and delivered a preferred item for 25 s. If the participant responded incorrectly to a mastered task, the therapist provided a model of the correct response. We incorporated mastered tasks into baseline to increase the likelihood that participants would continue to respond and engage in appropriate session behavior, despite the absence of differential consequences for responding to the target skills.

The therapist presented each target skill an equal number of times during a 12-trial session. The therapist also interspersed mastered tasks throughout the session on an additional 6 trials. Each baseline session consisted of 18 trials.

Control. The purpose of the control condition was to monitor correct responding in the absence of direct teaching. The therapist used procedures identical to baseline sessions for a trial in the control condition. The therapist ran a control session after two sessions had been conducted in the high-integrity condition and all low-integrity conditions. The therapist presented each target skill an equal number of times during a 12-trial session. The therapist interspersed mastered tasks throughout the session on an additional 6 trials. Each control session consisted of 18 trials.

High-integrity prompt (100% Integrity). The purpose of the high-integrity prompt condition was to evaluate the effects of delivering a prompt with high integrity on skill acquisition with children with ASD. The therapist began teaching in this condition with two 0-s prompt-delay sessions. Following two 0-s prompt-delay sessions, the therapist increased the delay between the presentation of the instruction and the prompt to 2 s (Chad and Ulysses Sets 1 and 2) or 5 s (Ulysses Set 3).

During a high-integrity trial, the therapist presented a picture card (e.g., a picture of a car; Ulysses only), secured the participant's attention, and presented the instruction (e.g., "What do you do with it?"). During 0-s prompt-delay trials, the therapist delivered a prompt immediately following the instruction (e.g., model, "drive"). On 2- or 5-s prompt-delay trials, the therapist waited 2 s or 5 s for the participant to respond. If the participant echoed the therapist's prompt correctly or responded correctly within the prompt delay, the therapist provided descriptive praise (e.g., "Excellent, drive!") and delivered brief access to a preferred item. If the participant echoed the therapist's prompt incorrectly or did not echo the therapist's prompt on 0-s prompt-delay trials, the therapist ended the trial. If the participant responded incorrectly or did not respond within the prompt delay during 2- or 5-s prompt-delay trials, the therapist delivered a prompt (i.e., modeled the correct response). The therapist began the next trial following a brief inter-trial interval (i.e., a 2- to 3-s pause between trials). During teaching with a 2- or 5-s prompt delay, the therapist delivered only descriptive praise for correct responses following a prompt (i.e., the therapist no longer delivered a preferred item for prompted responses) once the participant responded correctly on at least 50% (6 of 12) trials for two consecutive sessions. The purpose of this was to prevent prompt dependence (i.e., waiting for the therapist to deliver a prompt; Clark & Green, 2004).

Low-integrity prompt. The purpose of the low-integrity prompt condition was to evaluate the effects of delivering a prompt with different levels of low integrity on skill acquisition with children with ASD. The therapist began teaching in this condition with two 0-s prompt-delay sessions. Following two 0-s prompt-delay sessions, the therapist increased the delay between the presentation of the instruction and the prompt to 2 s (Chad and Ulysses Sets 1 and 2) or 5 s (Ulysses Set 3).

During a low-integrity trial, the therapist presented a picture card (e.g., a picture of a card; Ulysses only), secured the participant's attention, and presented the instruction (e.g., "What do you do with it?"). On low-integrity 0-s prompt delay trials, the therapist did not immediately deliver a prompt following the instruction. Instead, the therapist waited 2 s (Chad and Ulysses Sets 1 and 2) or 5 s (Ulysses Set 3) to respond. On 2- or 5-s prompt delay trials, the therapist waited 2 or 5 s for the participant to respond. If the participant responded correctly, the therapist provided descriptive praise (e.g., "Excellent, drive!") and delivered brief access to a preferred item. If the participant responded incorrectly, or did not respond, the therapist ended the trial and began the next trial following a brief inter-trial interval (i.e., a 2- to 3-s pause between trials). During low-integrity trials, the therapist implemented all other components of a teaching trial using procedures described in the high-integrity condition. For example, if the participant responded correctly, the therapist delivered descriptive praise and brief access to a preferred item.

We programmed prompting errors prior to the start of the experiment and balanced them across target skills in each condition. For example, in the 50% Integrity condition, we programmed a prompting error on two trials for each of the three target skills, totaling six trials

with programmed prompting errors. We also balanced prompting errors across the beginning of sessions (i.e., trials 1-6) and the end of sessions (i.e., trials 7-12).

Low-integrity prompt (75% Integrity). In this condition, we programmed prompting errors during 3 of the 12 trials in a session. During a low-integrity trial, if the participant responded incorrectly or did not respond, the therapist did not provide a prompt, and ended the trial. The therapist implemented 9 of the 12 trials in a session using procedures identical to the high-integrity condition.

Low-integrity prompt (50% Integrity). In this condition, we programmed prompting errors during 6 of the 12 trials in a session. During a low-integrity trial, if the participant responded incorrectly or did not respond, the therapist did not provide a prompt, and ended the trial. The therapist implemented 6 of the 12 trials in a session using procedures identical to the high-integrity condition.

Low-integrity prompt (25% Integrity). In this condition, we programmed prompting errors during 9 of the 12 trials in a session. During a low-integrity trial, if the participant responded incorrectly or did not respond, the therapist did not provide a prompt, and ended the trial. The therapist implemented 3 of the 12 trials in a session using procedures identical to the high-integrity condition.

High-integrity procedure with low-integrity set(s). The purpose of this condition was to demonstrate whether a participant would acquire a set of target skills that the therapist initially taught with low treatment integrity when the therapist subsequently taught them with high-integrity procedures. In this condition, if the participants reached the early-termination criterion (i.e., the participant did not master target skills within twice the number of sessions it took to master target skills in another condition) before mastering target skills taught with low-integrity

procedures, the therapist taught the low-integrity target skills with high-integrity procedures. The therapist continued teaching with procedures used in the high-integrity condition until the participant mastered target skills in the low-integrity set(s).

Maintenance. The purpose of the maintenance condition was to assess the short-term maintenance of skills taught with high- and low-integrity teaching procedures. After the participants mastered target skills in a condition, we conducted maintenance sessions once per week for up to four weeks. During maintenance sessions, we used procedures identical to baseline. The therapist presented each target skill an equal number of times during a 12-trial session. The therapist interspersed mastered tasks throughout the session on an additional 6 trials. Each maintenance session consisted of 18 trials.

Results and Discussion

Tables 4 and 5 show the obtained percentage of prompting errors for the first five, last five, and all sessions across conditions for Chad (Set 1) and Ulysses (Sets 1, 2, and 3). Figures 1-4 show the percentage of correct responses for Chad and Ulysses across the high-integrity (100% Integrity), low-integrity (75%, 50%, and 25% Integrity), and control conditions for baseline and teaching sessions.

Across conditions, Chad did not respond correctly during any of the baseline or 0-s prompt-delay sessions (Figure 1). Chad mastered the target skills from all conditions during teaching sessions with a 2-s prompt-delay. First, he mastered the target skills from the 75% Integrity condition following 18 sessions and 107 min of teaching. Next, he mastered the target skills from the 100% Integrity (22 sessions; 151 min) and 50% Integrity (22 sessions; 126 min) conditions in a comparable amount of teaching time. It took Chad the longest to acquire the target skills from the 25% Integrity condition (26 sessions; 119 min). Chad never responded

correctly to the target skills from the control condition. Chad mastered the target skills from the high-integrity and all low-integrity conditions in a comparable amount of time. Chad's results suggest that teaching with varying levels of prompting errors may not impair skill acquisition.

Table 4 shows the average obtained percentage of integrity with the prompting procedure across conditions for Chad. During the 100% Integrity condition, the therapist implemented the prompting procedure with 100% integrity across sessions. Overall, the average obtained percentage of integrity across the first 5 sessions of teaching and all sessions was similar to the programed percentage of integrity. During the last five training sessions, the average obtained percentage of integrity did deviate considerably from the programmed percentage of integrity for the 50% Integrity (M = 83%) and 25% Integrity (M = 0%) conditions. We observed larger differences between the obtained percentage of integrity and the programmed integrity percentages during the last five teaching sessions, because Chad was engaging in a higher frequency of correct responses. Thus, the therapist had fewer opportunities to engage in prompting errors when compared to earlier teaching sessions.

Figures 2-4 show Ulysses' correct responding. He did not respond during any baseline sessions across all conditions during Set 1 (Figure 2) Ulysses responded correctly on some trials during 0-s prompt-delay sessions across the low integrity conditions (i.e., 75%, 50%, and 25% Integrity). Ulysses mastered the target skills from all conditions except 25% Integrity during teaching with a 2-s prompt delay. First, Ulysses mastered target skills from the 100% Integrity condition (7 sessions; 41 min) and the 50% Integrity condition (7 sessions; 39 min). Next, he mastered target skills from the 75% Integrity in 9 sessions and 55 min. Ulysses did not master the target skills from the 25% Integrity condition within the early-termination criterion. Following teaching with 100% Integrity procedures, Ulysses mastered the target skills from the

25% Integrity condition (23 sessions; 120 min). Ulysses never responded correctly to the target skills from the control condition.

During Set 2 (Figure 3), Ulysses did not respond correctly during any baseline sessions across all conditions. He responded correctly on some 0-s prompt-delay trials across all low-integrity conditions. Ulysses mastered the target skills from all conditions during teaching with a 5-s prompt delay. He mastered target skills from the 25% Integrity condition first, in 28 sessions and 146 min. He mastered target skills from the 100% Integrity condition in a similar amount of teaching time (29 sessions; 164 min). Next Ulysses mastered the target skills from the 75% Integrity condition in 31 sessions and 179 min, followed by the target skills from the 50% Integrity condition in 32 sessions and 178 min. Ulysses responded correctly to the target skills from the control condition on some trials.

During Set 3 (Figure 4), Ulysses did not respond correctly during baseline sessions across all conditions. He responded correctly on some 0-s prompt-delay trials across all low-integrity conditions. Ulysses mastered the target skills from all conditions during teaching with a 5-s prompt delay. First, he mastered the target skills from the 75% Integrity condition (9 sessions; 65 min) and the 50% Integrity condition in a comparable amount of teaching time (9 sessions; 55 min). Next, Ulysses mastered target skills from the 25% Integrity condition in 13 sessions and 80 min. He mastered the target skills from the 100% Integrity condition last (18 sessions; 136 min). Ulysses never responded correctly to the target skills from the control condition.

Table 5 shows the average obtained percentage of integrity with the prompting procedure across conditions for Sets 1, 2, and 3 for Ulysses. During the 100% Integrity condition, the therapist implemented the prompting procedure with 100% integrity across sessions for all sets of target skills. Overall, the average obtained percentage of integrity across the first 5 sessions of

teaching and all sessions across all sets was similar to the programed percentage of integrity. During the last five training sessions, the average obtained percentage of integrity deviated from the programmed percentage of integrity for the 75% Integrity (Set 1; *M* = 95%) and 50% Integrity (Set 2; *M* = 85%) conditions. As with Chad, we observed larger differences between the obtained percentage of integrity and the programmed integrity percentages during the last five teaching sessions. This was because Ulysses was engaging in a higher frequency of correct responses, therefore the therapist had fewer opportunities to engage in prompting errors compared to earlier teaching sessions.

During Set 1, we found that teaching with 25% Integrity may have influenced skill acquisition for Ulysses. However, his correct responding during Sets 2 and 3 was inconsistent with his responding during Set 1. That is, we were not able replicate Ulysses' pattern of responding across the three sets of target skills. His results suggest that teaching with varying levels of prompting errors may not interfere with skill acquisition.

During Experiment 1, we evaluated the effects of varying levels of programmed prompting errors on skill acquisition for two children with ASD. Overall, the results of Experiment 1 suggest that teaching with varying levels of prompting errors may not interfere with skill acquisition for these participants. Previous research (e.g., Holcombe et al., 1994; Carroll et al., 2013) suggest that delivering prompts with low treatment integrity may interfere with skill acquisition for some children with ASD. Although the results of Experiment 1 do not show similar effects with our participants, delivering prompts with varying levels of integrity may impair skill acquisition for other children with ASD.

The results of Experiment 1 may be limited because our obtained integrity measures varied from our programmed integrity measures. For example, during the last 5 sessions in the

50% and 25% Integrity conditions (Chad) and the last 5 sessions in the 75% (Set 1) and 50% (Set 2) Integrity conditions (Ulysses), the obtained integrity values deviated from our programmed values. However, during the first 5 sessions and all sessions across all conditions and sets for both participants the average obtained integrity values were the same or similar to our programmed integrity values.

We wanted to hold the number of trials in a session and trials with exposure to programmed prompting errors consistent across sessions and conditions. If we had ensured that the integrity of each session for each condition was at the programmed level, we would have had to increase the number of trials in a session. Thus, the number of trials and exposure to prompting errors would have been inconsistent across sessions and conditions. To control for this, future researchers could increase the number of trials in a session if the obtained integrity does not match the programmed integrity within a 12-trial session (e.g., add three trials to make a 15-trial session). The therapist would then run all subsequent sessions in a condition with the same number of trials as the previous session. Future research could also consider using simultaneous prompting (i.e., 0-s prompt delay procedures; Swain, Lane, & Gast, 2015) throughout the entirety of teaching. To assess skill acquisition, the therapist could intersperse probe sessions (i.e., sessions with no differential consequences for correct, incorrect, or no responses).

During Experiment 1, we only evaluated the effects of programmed prompting errors on skill acquisition and found that it did not interfere with skill acquisition for both participants. Prompting is only one of the five main components of DTI. Implementing other components of DTI (e.g., delivering a reinforcer following a correct response) may influence skill acquisition differently.

Experiment 2: A Parametric Analysis of Reinforcement Errors During Discrete-Trial Instruction

Jenkins et al. (2015) investigated the effects of delivering a reinforcer following incorrect responses and Carroll et al. (2013) investigated the effects of withholding a reinforcer following correct responses. It is unclear how withholding a reinforcer following correct responses on more or less trials in a session influences skill acquisition. The purpose of Experiment 2 was to evaluate the effects of delivering a reinforcer when the therapist taught target skills with high integrity and varying levels of low-integrity (i.e., 25%, 50%, and 75% Integrity) during DTI.

Dependent Measures and Data Collection

During each session, we collected data on the participant's correct, incorrect, no responses, and problem behavior as described in Experiment 1. We converted each dependent measure to a percentage of trials and measured the total number of sessions and training time required for a participant to reach a pre-determined mastery criterion.

Interobserver Agreement and Treatment Integrity

A secondary observer independently collected data on participant responses for an average of 62% of sessions (range, 50% to 86%). We calculated interobserver agreement by taking the number of trials with an agreement in a session, dividing it by the number of agreements plus disagreements, and multiplying by 100. The average interobserver agreement for Chad was 99% (range, 92% to 100%) and the average interobserver agreement for Ulysses was 99% (range, 93% to 100%).

The secondary observer also collected data on the therapist's correct implementation of the teaching protocol for an average of 62% (range, 50% to 86%) of sessions. During all conditions, the therapist's correct implementation of a trial included (a) securing attention,

defined as waiting to present the instruction until the participant was sitting upright and oriented toward the table; and (b) presenting the instruction exactly as outlined in the teaching protocol. During baseline, maintenance, and control trials, the therapist's correct implementation of a trial was identical to Experiment 1, except that the therapist delivered descriptive praise and tokens following Chad's correct responses to mastered tasks. The therapist's correct implementation of a high-integrity trial was identical to Experiment 1, except that the therapist provided praise and tokens following Chad's correct responses. During low-integrity trials, correct implementation of a trial included (a) delivering a prompt following an incorrect or no response and (b) withholding praise and brief access to a preferred item or a token following a correct response. Following sessions in all conditions, Chad exchanged the tokens he earned. The therapist's correct implementation of a token exchange (Chad only) included (a) delivering the item Chad selected during the pre-session preference assessment, (b) delivering the preferred item for the appropriate amount of time (i.e., providing 25 s of access to the item for each token earned), and (c) delivering a behavior-specific statement once during every 25-s interval.

We scored trials as either correct or incorrect. We calculated treatment integrity for each session by taking the number of trials implemented correctly, dividing it by the total number of trials in a session, and multiplying by 100. Average treatment integrity for Chad's sessions was 100% and average treatment integrity for Ulysses was 99% (range, 92% to 100%).

Preference Assessment

Before the first session of each day, the therapist conducted a brief multiple stimulus without replacement preference assessment (Higbee et al., 2000), as described in Experiment 1 to identify preferred edible and tangible items to use during teaching. During Experiment 2, the

therapist conducted preference assessments before every session with both participants and used only the first tangible item that the participant chose as a preferred item during sessions.

Experimental Design

We used an adapted alternating treatments design (Sindelair et al., 1985) to compare skill acquisition across high-integrity reinforcement, low-integrity reinforcement, and control conditions with children with ASD. Specifically, we compared the effects of a therapist delivering reinforcers with high treatment integrity (i.e., 100% Integrity), varying levels of low treatment integrity (i.e., 25%, 50%, and 75% Integrity), and not directly teaching target skills (i.e., control condition) on skill acquisition.

General Procedures

Pretest and target assignment. We conducted a pretest that was identical to the one we used in Experiment 1, to ensure that the participants did not know the target skills before the start of the experiment. For each participant we identified 15 target skills that were not associated with correct responding during the pretest. We assigned three target skills to each condition.

Teaching. The therapist presented target skills an equal number of times during a 12-trial session. The therapist conducted two sessions for the high- and low-integrity conditions and up to one control session per day. The therapist conducted sessions with Chad three days a week and with Ulysses two days a week. Thus, the maximum number of sessions that the therapist conducted per week for Chad equaled 25 and the maximum number of sessions that the therapist conducted per week for Ulysses equaled 17.

The therapist used a constant prompt-delay procedure (Gast et al., 1988) to teach target skills as described in Experiment 1. Following two 0-s prompt-delay sessions, the therapist

increased the delay between the presentation of the instruction and the delivery of the prompt to 3 s (Chad) or 5 s (Ulysses).

Mastery and early termination criteria. We considered a set of target skills mastered once the participant responded correctly during at least 92% (11 of 12) of trials for two consecutive sessions. We kept the same early-termination criterion, as described in Experiment 1. However, both participants acquired all of the target skills before reaching the early-termination criterion.

Token Economy

During Experiment 2, we used a token economy with Chad. We used a token economy because the early intervention clinic that Chad was attending was also using a token economy with him. Following Chad's correct responses, the therapist placed a token (i.e., a small, plastic, green square) on a laminated sheet of paper. At the end of each 12-trial session, Chad exchanged the tokens he earned for the tangible item he selected during the pre-session preference assessment. Each token that he earned equaled 25 s of access to the item he selected. For example, if Chad earned 12 tokens, he was able to play with the item he selected for 5 min.

Procedures

Baseline. The purpose of the baseline condition was to ensure that the participants did not know the target skills prior to teaching. The therapist conducted baseline sessions with procedures identical to Experiment 1. However, in Experiment 2, the therapist delivered praise and tokens following Chad's correct responses to mastered tasks. We used the same mastered tasks for each participant that we used in Experiment 1.

Control. The purpose of the control condition was to monitor correct responding in the absence of direct teaching. The therapist used procedures identical to baseline sessions for

sessions in the control condition. The therapist conducted control sessions in the same manner as described in Experiment 1.

High-integrity reinforcement (100% Integrity). The purpose of the high-integrity reinforcement condition was to evaluate the effects of delivering reinforcers with high integrity on skill acquisition with children with ASD. The therapist began teaching in this condition with two 0-s prompt-delay sessions. The therapist used procedures identical to the high-integrity prompt condition in Experiment 1 for 0-s prompt-delay and 3- or 5-s prompt-delay sessions, with the exception of delivering praise and tokens following Chad's correct responses.

Low-integrity reinforcement. The purpose of the low-integrity reinforcement condition was to evaluate the effects of delivering reinforcers with different levels of low integrity on skill acquisition with children with ASD. The therapist began teaching in this condition with two 0-s prompt-delay sessions. During a low-integrity 0-s prompt-delay trial, the therapist secured the participant's attention and presented the instruction (e.g., "What has hooves?") immediately followed by a model of the correct response (e.g., "horse"). If the participant echoed the model correctly, the therapist withheld praise and a preferred item or a token and ended the trial. Following a brief inter-trial-interval (i.e., a 2- to 3-s pause between trials), the therapist presented the next trial. After conducting two 0-s prompt delay sessions, the therapist increased the delay between the presentation of the instruction and the prompt from 0 s to 3 s (Chad) or 5 s (Ulysses).

The therapist began a low-integrity 3- or 5-s prompt-delay trial with securing the participant's attention and delivering the instruction. The therapist waited either 3 s (Chad) or 5 s (Ulysses) for the participant to respond to the instruction. If the participant responded correctly, the therapist withheld praise and a preferred item or a token and ended the trial. The therapist

presented the next trial following a brief inter-trial interval. During low-integrity reinforcement trials, the therapist implemented all other components of a teaching trial using procedures described in the high-integrity prompt condition in Experiment 1. We programmed and balanced reinforcement errors in the same manner as described in Experiment 1.

Low-integrity reinforcement (75% Integrity). In this condition, we programmed reinforcement errors during 3 of the 12 trials in a session. During a low-integrity trial, if the participant responded correctly, the therapist did not provide praise and a preferred item or a token, and ended the trial. The therapist implemented 9 of the 12 trials in a session using procedures identical to the high-integrity condition.

Low-integrity prompt (50% Integrity). In this condition, we programmed reinforcement errors during 6 of the 12 trials in a session. During a low-integrity trial, if the participant responded correctly, the therapist did not provide praise and a preferred item or a token, and ended the trial. The therapist implemented 6 of the 12 trials in a session using procedures identical to the high-integrity condition.

Low-integrity prompt (25% Integrity). In this condition, we programmed reinforcement errors during 9 of the 12 trials in a session. During a low-integrity trial, if the participant responded correctly, the therapist did not provide praise and a preferred item or a token, and ended the trial. The therapist implemented 3 of the 12 trials in a session using procedures identical to the high-integrity condition.

Maintenance. The purpose of this condition was to assess the short-term maintenance of skills taught with high- and low-integrity teaching procedures. After the participants mastered target skills in a condition, we conducted weekly maintenance sessions for up to four weeks. The maintenance procedures used in Experiment 2 were identical to those described in Experiment 1.

Results and Discussion

Figures 5 and 6 show the percentage of correct responses for Chad and Ulysses across the high-integrity (100% Integrity), low-integrity (75%, 50%, and 25% Integrity), and control conditions for baseline and all teaching sessions. We kept the percentage of reinforcement errors constant throughout this experiment, so we did not calculate the obtained percentage of integrity for each condition.

Chad did not respond correctly during baseline and 0-s prompt-delay sessions across all conditions (Figure 5). During Experiment 2, we did not give the participants an opportunity to respond independently during 0-s prompt-delay sessions, so we expected responding to be 0% correct. Chad mastered the target skills from all conditions during teaching with a 3-s prompt delay. First, he mastered the target skills from the 25% Integrity condition in 7 sessions and 20.4 min. Next, he mastered the target skills from the 100% Integrity (8 sessions; 25.7 min) and 75% Integrity (8 sessions; 27.2 min) conditions. Finally, Chad mastered the target skills from the 50% Integrity condition in 12 sessions and 36.8 min. Chad never responded correctly to the target skills from the control condition. Chad mastered the target skills from all conditions in a comparable amount of time. Notably, he mastered the target skills from the 25% Integrity condition first. These results suggest that varying levels of reinforcement errors may not interfere with skill acquisition.

Ulysses did not respond correctly during baseline and 0-s prompt delay sessions across all conditions (Figure 6). Ulysses mastered the target skills from all conditions during teaching with a 5-s prompt delay. He mastered the target skills from the 100% Integrity condition first (4 sessions; 52.6 min). Then, he mastered the target skills from the 75% Integrity (7 sessions; 51 min) and 25% Integrity (7 sessions; 28 min) conditions. It took Ulysses the longest to master the

target skills from the 50% Integrity condition (8 sessions; 45.7 min). He never responded correctly to the target skills from the control condition. As was the case with Chad, Ulysses mastered the target skills from all conditions in a comparable amount of time. Notably, he mastered the target skills from the 25% Integrity condition second (in the same amount of time as the 75% Integrity condition). These results suggest that varying levels of reinforcement errors may not interfere with skill acquisition for these participants.

During Experiment 2, both participants mastered the target skills from all conditions in a comparable amount of time. These results are similar to those of Carroll et al.'s (2013) Study 3, which found that all three participants acquired skills when a therapist delivered a reinforcer for correct responses with low integrity, although one of the three participants showed delayed acquisition. However, other research (e.g., Jenkins et al., 2015) has evaluated the effects of delivering a reinforcer for incorrect responses and found that, in this case, participants did not acquire skills. Had we delivered reinforcers for incorrect responses, we may have seen results similar to Jenkins et al. (2015).

General Discussion

In Experiment 1, we evaluated the effects of delivering a prompt with varying levels of integrity on skill acquisition for two participants with ASD. In Experiment 2, we evaluated the effects of delivering a reinforcer following correct responses with varying levels of integrity on skill acquisition for the two participants from Experiment 1. In the current study, delivering prompts and reinforcers with varying levels of integrity did not appear to interfere with skill acquisition for two participants with ASD. Specifically in Experiment 1, we found that both participants acquired the target skills from all conditions, in comparable amounts of time. These results are inconsistent with those of previous studies that found that delivering prompts with low

treatment integrity interferes with skill acquisition for children with intellectual disabilities (e.g., Holcombe et al., 1994) and for typically developing children (Noell et al., 2002). For example, Holcombe and colleagues found that providing prompts with 50% integrity interfered with skill acquisition for 5 of the 6 participants. During Experiment 2, we found that both participants acquired the target skills from all conditions in a comparable amount of time. The results from Experiment 2 are also inconsistent with previous research that has found that delivering reinforcers with low treatment integrity impairs skill acquisition for children with ASD (Jenkins et al., 2015). Jenkins and colleagues examined the effects of delivering reinforcers following incorrect responses on 100%, 50%, and 0% of trials in a session. They found that delivering a reinforcer for incorrect responses on 100% of trials in a session impaired skill acquisition for all 4 participants and delivering a reinforcer for incorrect responses on 50% of trials in a session impaired skill acquisition for 3 of 4 participants.

There are a number of potential explanations for why participants in the current study were able to acquire skills with low levels of treatment integrity. First, in the current study, we evaluated the effects of implementing only one component of DTI with low levels of treatment integrity in isolation. That is, we evaluated the effects of delivering a prompt with low integrity in isolation, and we evaluated the effects of delivering reinforcers with low integrity in isolation. In Study 2, Carroll et al. (2013) investigated the effects of implementing three components of DTI with 67% integrity simultaneously. They found that when a therapist implemented multiple components of DTI with low integrity simultaneously, skill acquisition was impaired for 5 of 6 participants. Had we implemented multiple components of DTI with varying levels of low integrity, we may have seen similar results as Carroll and colleagues. Future research should continue to investigate the effects of simultaneously implementing multiple components of DTI

with low treatment integrity on skill acquisition. In Study 3, Carroll and colleagues isolated the components of DTI and implemented each with 67% integrity. They found that when the therapist delivered prompts with low integrity, skill acquisition was impaired for only 1 of 3 participants, and when they delivered reinforcers for correct responses with low integrity, all 3 participants acquired the target skills, albeit with one participant showing delayed acquisition. Carroll and colleagues' results from Study 3 are consistent with the findings for the participants in the current study, who acquired the target skills even when procedures were implemented with low integrity.

A second potential explanation for the results of the current study may be the participants' histories with DTI. Before the start of the study, Ulysses had six months of experience with DTI. It is possible that he did not acquire the skills from the 25% Integrity condition during Set 1 because of his short history with DTI. However, by the time we began teaching the target skills in Set 2, Ulysses had an additional 5 months of experience with DTI. This increase in exposure to DTI may have influenced why we were unable to replicate the pattern of responding we saw in Set 1. Unlike Ulysses, Chad had an extensive history with DTI (i.e., 3 years 6 months) before the start of the current study. It is possible that Chad acquired the target skills from all conditions in a comparable amount of time because of his extensive history to DTI. Implementing components of DTI with low treatment integrity may interfere skill acquisition for learners with little to no exposure to DTI. Future research should compare the effects of implementing components of DTI with low integrity on skill acquisition for children with ASD who have had no exposure to DTI to children with ASD who have had extensive exposure to DTI (e.g., for at least one year).

A third explanation for the results of the current study could have to do with the difficulty of the target skills that we taught. During Experiment 1, Ulysses did not acquire the target skills from the 25% Integrity condition within the early-termination criterion in Set 1. We attempted to replicate this pattern of responding with a new set of target skills. During Set 2 of Experiment 1, it took Ulysses substantially longer to acquire the target skills across the high- and low-integrity conditions, compared to Set 1. This increase in number of sessions and time to acquire the target skills from Set 2 may have been influenced by an increased difficulty of the target skills, despite our efforts to equate the difficulty of the target skills within and across sets. Following Set 2, we attempted to replicate with a third set for Ulysses, and again observed a different pattern of responding. Previous studies that have evaluated the influence of treatment integrity errors on skill acquisition have not conducted within participant replications (e.g., Holcombe et al., 1994; Carroll et al., 2013; Jenkins et al., 2015). The results of the current experiments highlight the importance of conducting within participant replications when comparing instructional procedures with an adapted alternating treatments design. Within participant replication increases the likelihood that the difficulty of the target skills was equated and the observed differences were due to the different instructional procedures and not extraneous variables (Wolery, Gast, & Hammond, 2010).

The current study had several limitations. First, we did not conduct within participant replications with Chad in Experiment 1 and with either participant in Experiment 2. During Experiments 1 and 2, Chad acquired the target skills from all conditions in a similar amount of time. During Experiment 2, Ulysses acquired the target skills from all conditions in a similar amount of time. Because we did not see substantial differences in skill acquisition between conditions for Chad (Experiments 1 and 2) or Ulysses (Experiment 2), we did not attempt to

replicate these patterns of responding with new sets of target skills. As previously mentioned, it is important to conduct within participant replications. Future research should conduct additional studies with teaching with low treatment integrity where they conduct within participant replications.

Second, in the current study we used an adapted alternating treatments design to compare skill acquisition across high and varying levels of low integrity conditions. With this design, the therapist rapidly alternates between conditions to simultaneously compare the effects of each condition (Sindelair et al., 1985). Given the alternation between high- and low- integrity conditions, it is possible that responding during a low-integrity condition was influenced by a preceding condition of higher integrity (i.e., carryover effect). In the current study we took several steps to minimize potential carryover effects. First, we only conducted one to two sessions for each condition in a day. We also incorporated brief breaks between sessions. Although we took steps to minimize carryover between conditions, our results may have been different if we used distinct external stimuli to signal each condition. For example, we could have paired each condition with a different color card, or the therapist could have worn a different colored shirt when conducting sessions in each condition. We may have seen different results if we had only conducted one session in each condition per day. Future research should evaluate the ideal presentation of sessions in order to ensure that participants discriminate between the different conditions and to minimize carryover effects.

The results from the current study only apply to two participants with ASD. This small number of participants was a limitation of the current study. We may have seen different patterns of responding if we had enrolled a larger number of participants. Additionally, the language skills of both participants were similar. Chad's age-equivalent score on the Peabody Picture

Vocabulary Test-4 was 3 years 3 months and his age-equivalent score on the Expressive Vocabulary Test-2 was 2 years 6 months. These scores were similar to Ulysses, whose age equivalent scores were 3 years 1 month on the Peabody Picture Vocabulary Test-4 and 3 years 5 months on the Expressive Vocabulary Test-2. We may have gotten different results if our participants had different levels of language skills. Future research should investigate the effects of implementing DTI with low levels of integrity on skill acquisition for children who have less language skills.

The results of the current study suggest several areas for future research. First, future research should investigate the effects of varying levels of low integrity on skill acquisition with skills of varying difficulty. It may not be necessary for a therapist to implement DTI with high integrity when they are teaching target skills that are easy for a child to learn. However, the therapist might have to implement DTI with high integrity when teaching a child a more difficult skill. Second, future research should assess the long-term maintenance of skills that are taught with low levels of treatment integrity. In the current study, we only assessed maintenance of target skills for up to four weeks following mastery. Children with ASD may maintain target skills that are taught with low treatment integrity shortly after they master them, however, it is possible that they will not maintain those skills after a longer period of time following mastery.

A third area for future research concerns the components of DTI that are implemented with high and low treatment integrity. In the current study, we only evaluated the effects of delivering prompts and reinforcers in isolation with varying levels of low integrity on skill acquisition. It is possible that implementing both prompts and reinforcers with low integrity together will impair skill acquisition. It is also possible that implementing other components of DTI (e.g., the instruction, the inter-trial interval) with varying levels of low integrity will impair

skill acquisition. Future research should investigate the effects of implementing different components of DTI in combination and in isolation with varying levels of low integrity.

In the current study, we found that delivering prompts and reinforcers with varying levels of integrity did not interfere with skill acquisition for two participants with ASD. However, these results only apply to the two children with ASD who participated in the current study. Future researchers should continue to investigate if implementing components of DTI with varying levels of low integrity interferes with skill acquisition.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

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*, 379-394. doi:10.1002/jaba.49

Carroll, R. A., Kodak, T., & Adolf, K. J. (2016). Effect of delayed reinforcement on skill acquisition during discrete‐trial instruction: Implications for treatment‐integrity errors in academic settings. *Journal of Applied Behavior Analysis*, *49*, 176-181. doi:10.1002/jaba.268

Centers for Disease Control and Prevention. (2015). *Autism spectrum disorders*. Retrieved from http://www.cdc.gov/ncbddd/autism/research.html

Clark, K. M., & Green, G. (2004). Comparison of two procedures for teaching dictated-word/symbol relations to learners with autism. *Journal of Applied Behavior Analysis*, *37*, 503–507. doi: 10.1901/jaba.2004.37-503

Dib, N., & Sturmey, P. (2007). Reducing student stereotypy by improving teachers' implementation of discrete-trial teaching. *Journal of Applied Behavior Analysis*, *40*, 339-343. doi:10.1901/jaba.2007.52-06

DiGennaro Reed, F. D., Reed, D. D., Baez, C. N., & Maguire, H. (2011). A parametric analysis of errors of commission during discrete-trial training. *Journal of Applied Behavior Analysis*, *44*, 611-615. doi:10.1901/jaba.2011.44-611

Dunn, L. M., & Dunn, D. M. (2007). *Peabody picture vocabulary test manual* (4th ed.). Minneapolis, MN: Pearson.

Gast, D. L., Ault, M. J., Wolery, M., Doyle, P. M., & Belanger, S. (1988). Comparison of constant time delay and the system of least prompts in teaching sight word reading to students with moderate retardation. *Education and Training in Mental Retardation*, *23*(2), 117-128. Retrieved from http://www.jstor.org/stable/

Geiger, K. B., Carr, J. E., LeBlanc, L., Hanney, N. M., Polick, A. S., & Heinicke, M. R. (2012). Teaching receptive discriminations to children with autism: A comparison of traditional and embedded discrete trial training. *Behavior Analysis in Practice*, *5*, 49-59. doi: 10.1007/BF03391823

Green, V. A., Pituch, K. A., Itchon, J., Choi, A., O'Reilly, M., & Sigafoos, J. (2006). Internet survey of treatments used by parents of children with autism. *Research in Developmental Disabilities*, 27, 70-84. doi:10.1016/j.ridd.2004.12.002

Grow, L. L., Carr, J. E., Gunby, K. V., Charania, S. M., Gonsalves, L., Ktaech, I. A., & Kisamore, A. N. (2009). Deviations from prescribed prompting procedures: Implications for treatment integrity. *Journal of Behavioral Education*, *18*, 142-156. doi:10.1007/s10864-009-9085-6

Grow, L. L., Kodak, T., & Clements, A. (2016). An evaluation of instructive feedback to teach play behavior to a child with autism spectrum disorder. *Behavior Analysis in Practice*, doi:10.1007/s40617-016-0153-9

Higbee, T. S., Carr, J. E., & Harrison, C. D. (2000). Further evaluation of the multiple-stimulus preference assessment. *Research in Developmental Disabilities*, *21*, 61-73. doi: 10.1016/S0891-4222(99)00030-X

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*, 49–73.

doi:10.1007/BF01560509

Howard, J. S., Sparkman, C. R., Cohen, H. G., Green, G., & Stanislaw, H. (2005). A comparison of intensive behavior analytic and eclectic treatments for young children with autism. *Research in Developmental Disabilities, 26*, 359-383. doi:10.1016/j.ridd.2004.09.005

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

Lovaas, O. I. (1987). Behavioral treatment and normal education and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55*, 3-9.

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*, 51– 67. doi:10.1023/A:1014385321849

Peterson, L., Homer, A. L., & Wonderlich, S. A. (1982). The integrity of independent variables in behavior analysis. *Journal of Applied Behavior Analysis, 15*, 477– 492. doi:10.1901/jaba.1982.15-477

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

Reichow, B. (2012). Overview of meta-analyses on early intensive behavioral intervention for young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 42*, 512-520. doi:10.1007/s10803-011-1218-9

Reichow, B., Barton, E. E., Boyd, B. A., & Hume, K. (2012). Early intensive behavioral intervention (EIBI) for young children with autism spectrum disorders (ASD): A systematic review. *Cochrane Database of Systematic Reviews*, *10*. doi: 10.1002/14651858.CD009260.pub2

Rivard, M., Terroux, A., & Mercier, C. (2014). Effectiveness of early behavioral intervention in public and mainstream settings: The case of preschool-age children with autism spectrum disorders. *Research in Autism Spectrum Disorders*, *8*, 1031-1043. doi:10.1016/j.rasd.2014.05.010

Shillingsburg, M. A., Bowen, C. N., & Shapiro, S. K. (2014). Increasing social approach and decreasing social avoidance in children with autism spectrum disorder during discrete trial training. *Research in Autism Spectrum Disorders*, *8*, 1443-1453. doi:10.1016/j.rasd.2014.07.013

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. Retrieved from http://www.jstor.org/stable/42898888

Smith, T. (2001). Discrete trial training in the treatment of autism. *Focus on Autism and Other Developmental Disabilities*, *16*, 86–92. doi: 10.1177/108835760101 600204

Sundberg, M. (2008). *Verbal behavior milestones assessment and placement program: The VB-MAPP*. Concord, CA: AVB Press.

Swain, R., Lane, J. D., & Gast, D. L. (2015). Comparison of constant time delay and simultaneous prompting procedures: Teaching functional sight words to students with intellectual disabilities and autism spectrum disorder. *Journal of Behavioral Education*, *24*, 210-229. doi:10.1007/s10864-014-9209-5

Symes, M. D., Remington, B., Brown, T., & Hastings, R. P. (2006). Early intensive behavioral intervention for children with autism: Therapists' perspectives on achieving procedural fidelity. *Research in Developmental Disabilities*, *27*, 30-42. doi:10.1016/j.ridd.2004.07.007

Taubman, M., Brierley, S., Wishner, J., Baker, D., McEachin, J., & Leaf, R. B. (2001). The effectiveness of a group discrete trial instructional approach for preschoolers with developmental disabilities. *Research in Developmental Disabilities*, *22*, 205-219. doi:10.1016/S0891-4222(01)00068-3

Vollmer, T. R., Sloman, K. N., & Pipkin, C. P. (2008). Practical implications of data reliability and treatment integrity monitoring. *Behavior Analysis in Practice*, *1*, 4-11 doi: 10.1007/BF03391722

Weiss, M. J. (1999). Differential rates of skill acquisition and outcomes of early intensive behavioral intervention for autism. *Behavioral Interventions*, *14*, 3-22. doi:10.1002/(SICI)1099-078X(199901/03)14:1<3::AID-BIN25>3.0.CO;2

Williams, K. T. (2002). *Expressive vocabulary test manual* (2nd ed.). Minneapolis, MN: Pearson.

Wolery, M., Gast, D. L., & Hammond, D. (2010). Comparative intervention designs. In D. L. Gast (Ed.). *Single subject research methodology in behavioral sciences* (pp. 329-381). New York: Routledge.

Table 1

Participant characteristics

Participant	Age	History with DTI	VBMAPP	PPVT-4	EVT-2
Chad	7:11	3:6	30-48 months	3:3	2:6
Ulysses	3:5	0:6	18-30 months	3:1	3:5
Hannah	5:9	2:9	0-30 months	*	*

Note: Age provided in years: months; DTI = discrete-trial instruction, history with DTI provided in years: months; VBMAPP = Verbal Behavior Milestones Assessment Placement Program (Sundberg, 2008); PPVT-4 = Peabody Picture Vocabulary Test, 4th Edition (Dunn & Dunn, 2007) age provided in years: months; EVT-2 = Expressive Vocabulary Test, 2nd Edition (Williams, 2007) age provided in years: months; * = the participant was untestable on an assessment

Table 2

Target responses for each participant by condition for Experiment 1

Participant	Skill	Target Skills
Chad	Set 1 Labeling Categories of Items	100% Integrity: What is an ocean animal? (a crab, a whale); What is a zoo animal? (a lion, a zebra); What is a hot drink? (tea, coffee) 75% Integrity: What is bedroom furniture? (a bed, a dresser); What is living room furniture? (a couch, a bookshelf); What is an emotion? (happy, sad) 50% Integrity: What is a cleaning tool? (a vacuum, a sponge); What is a garden tool? (a rake, a hose); What is a kind of weather? (rain, snow) 25% Integrity: What are summer clothes? (a swimsuit, a t-shirt); What are winter clothes? (pants, a jacket); What is on the playground? (swings, a slide) Control: What is a dinner food? (tator tots, hot dog); What is a breakfast food? (eggs, bacon); What is a school supply? (a pencil, a book)
Ulysses	Set 1 Labeling Functions of Items	100% Integrity: play (piano), tie (shoes), crack (eggs) 75% Integrity: smell (nose), pour (milk), write (pencil) 50% Integrity: buy (money), lick (ice cream), ring (bell) 25% Integrity: cook (pan), dig (shovel), peel (banana) Control: sail (boat), talk (mouth), clean (vacuum)
	Set 2 Labeling Functions of Items	100% Integrity: lock (key), fold (clothes), carry (backpack) 75% Integrity: pound (hammer), dry (towels), kick (ball) 50% Integrity: sleep (bed), fly (airplane), wash (soap) 25% Integrity: pull (wagon), wear (pants), frost (cupcake) Control: sing (microphone), float (lifejacket), push (stroller)
	Set 3 Labeling Categories of Items	100% Integrity: What is a rope? (tool); What is a duck? (farm animal); What is a shirt? (clothing) 75% Integrity: What is milk? (drink); What is a flower? (outside thing); What is a bus? (vehicle) 50% Integrity: What is an apple? (fruit); What is a bathtub? (bathroom thing) What is a ring? (jewelry) 25% Integrity: What is a worm? (bug); What is a book? (school supply); What is a carrot? (school supply) Control: What is a star? (shape); What is a toaster? (kitchen thing); What is a chair? (furniture)

Table 3

Target responses for each participant by condition for Experiment 2

Participant	Skill	Target Skills
Chad	Set 1 Rotating Wh-Questions	100% Integrity: Why do you take a bath? (to get clean); Where do you use a shovel? (in the garden); What is a shovel? (a tool) 75% Integrity: Why do you wear a coat? (to stay warm); When do you brush your teeth? (in the morning); Where do you brush your teeth? (the bathroom) 50% Integrity: Why do you use a tissue? (to blow your nose); Where do you put turkey? (on a sandwich); When do you eat turkey? (at lunch) 25% Integrity: Why do you use an umbrella? (to stay dry); What do you cut grass with? (a lawn mower); When do you cut grass? (when it is long) Control: Why do birds have wings? (to fly); What is an alligator? (a reptile); Where do alligators live? (in a swamp)
Ulysses	Set 1 Identifying Features of Items	100% Integrity: What has hooves? (horse); What has a screen? (computer); What has a lid? (jar) 75% Integrity: What has a shell? (turtle); What has pedals? (bike); What has a bulb? (light) 50% Integrity: What has a trunk? (tree); What has laces? (shoes); What has stripes? (zebra) 25% Integrity: What has a strap? (purse); What has cushions? (couch); What has a plug? (vacuum) Control: What has a stem? (leaf); What has pockets? (pants); What has fur? (bunny)

Table 4

Average percentage of obtained prompting errors and average number of incorrect or no responses for the first 5 sessions, last 5 sessions, and all sessions for Chad for Experiment 1

	First 5 Sessions	Last 5 Sessions	All Sessions
100% Integrity	100 (7)	100 (1)	100 (5)
75% Integrity	76 (8)	72 (2)	73 (6)
50% Integrity	51 (9)	83 (1)	55 (4)
25% Integrity	25 (8)	0 (1)	22 (6)

Note: The number in parenthesis represents the average number of incorrect or no responses that the participant engaged in across sessions

Table 5

Average percentage of obtained prompting errors and average number of incorrect or no responses for the first 5 sessions, last 5 sessions, and all sessions for Ulysses for Experiment 1 for Sets 1, 2, and 3

	First 5 Sessions	Last 5 Sessions	All Sessions
100% Integrity	Set 1: 100 (2)	Set 1: 100 (2)	Set 1: 100 (1)
	Set 2: 100 (4)	Set 2: 100 (1)	Set 2: 100 (3)
	Set 3: 100 (3)	Set 3: 100 (2)	Set 3: 100 (2)
75% Integrity	Set 1: 72 (4)	Set 1: 95 (2)	Set 1: 83 (3)
	Set 2: 75 (6)	Set 2: 74 (2)	Set 2: 80 (4)
	Set 3: 80 (3)	Set 3: 83 (2)	Set 3: 81 (3)
50% Integrity	Set 1: 60 (3)	Set 1: 37 (2)	Set 1: 50 (2)
	Set 2: 53 (3)	Set 2: 85 (2)	Set 2: 58 (4)
	Set 3: 55 (5)	Set 3: 53 (3)	Set 3: 54 (4)
25% Integrity	Set 1: 34 (6)	Set 1: 26 (3)	Set 1: 29 (5)
	Set 2: 43 (4)	Set 2: 15 (3)	Set 2: 27 (4)
	Set 3: 32 (6)	Set 3: 27 (3)	Set 3: 27 (5)

Note: The number in parenthesis represents the average number of incorrect or no responses that the participant engaged in across sessions

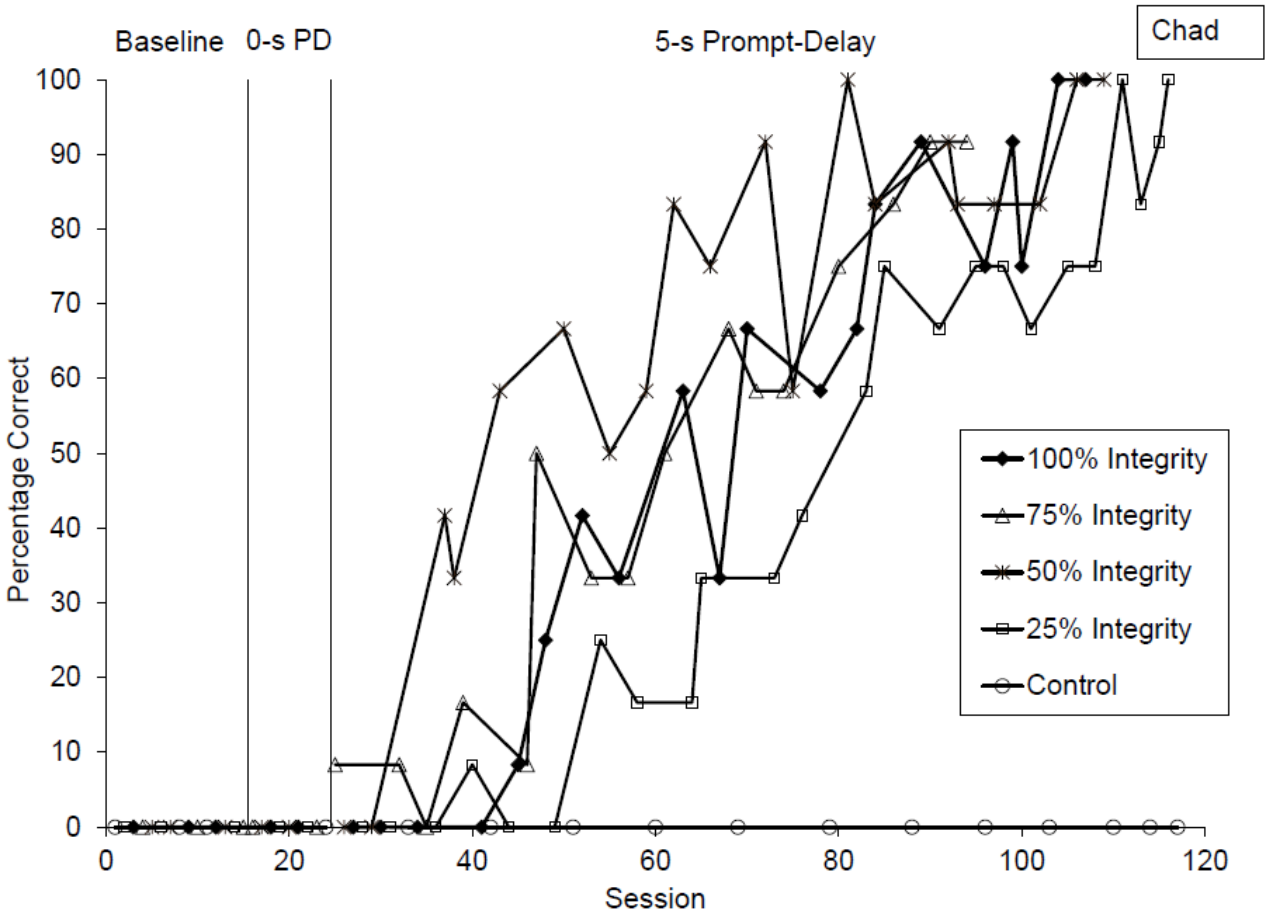

Figure 1. Percentage of correct responses for Chad for the control condition and the 100% Integrity, 75% Integrity, 50% Integrity, and 25% Integrity conditions for baseline, 0-s prompt-delay, and 5-s prompt-delay sessions during the comparison of prompting errors (Experiment 1).

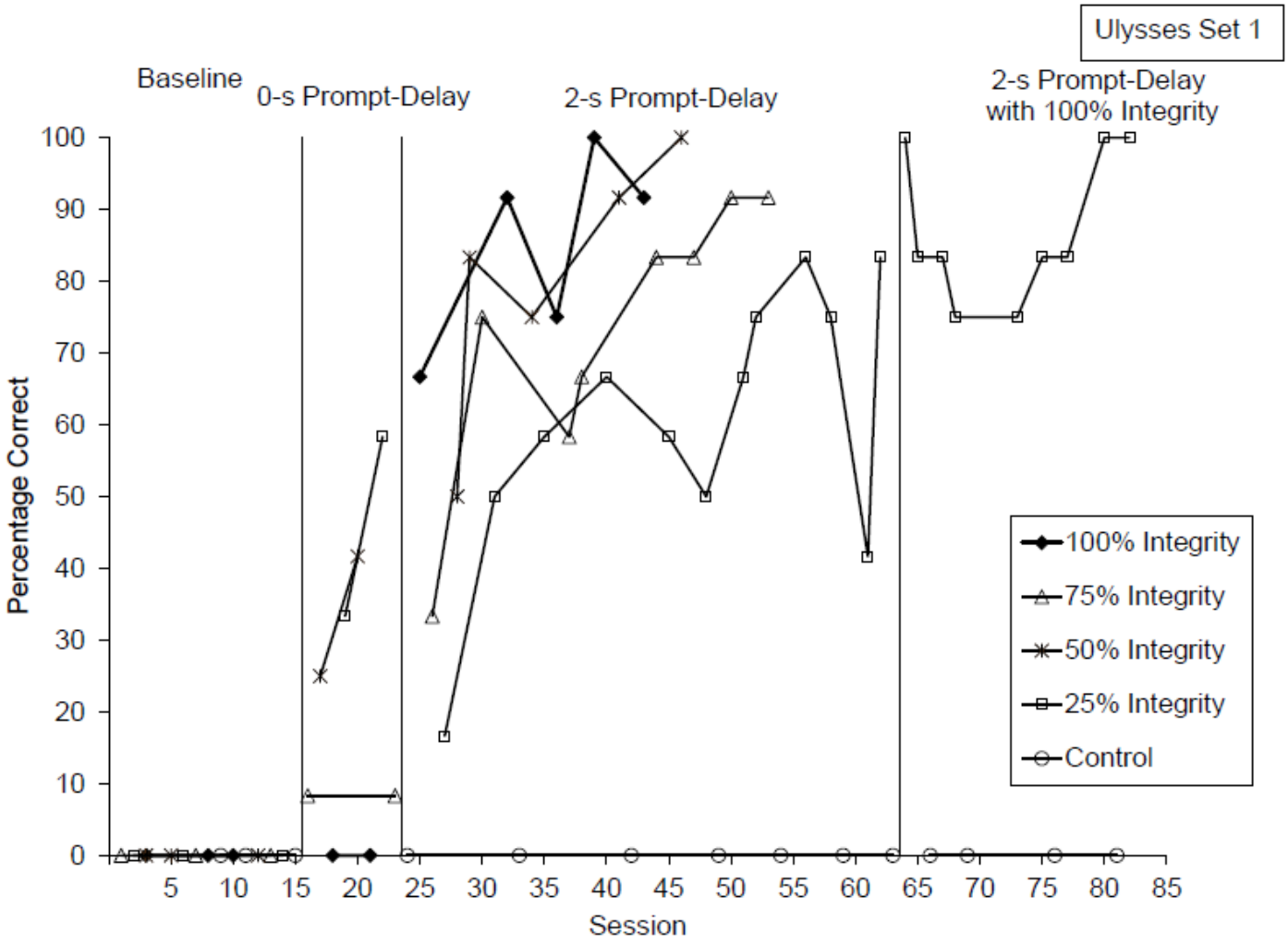

Figure 2. Percentage of correct responses for Ulysses for the control condition and the 100% Integrity, 75% Integrity, 50% Integrity, and 25% Integrity conditions for baseline, 0-s prompt-delay, 2-s prompt-delay sessions, and 2-s prompt-delay with 100% Integrity sessions for set 1 during the comparison of prompting errors (Experiment 1).

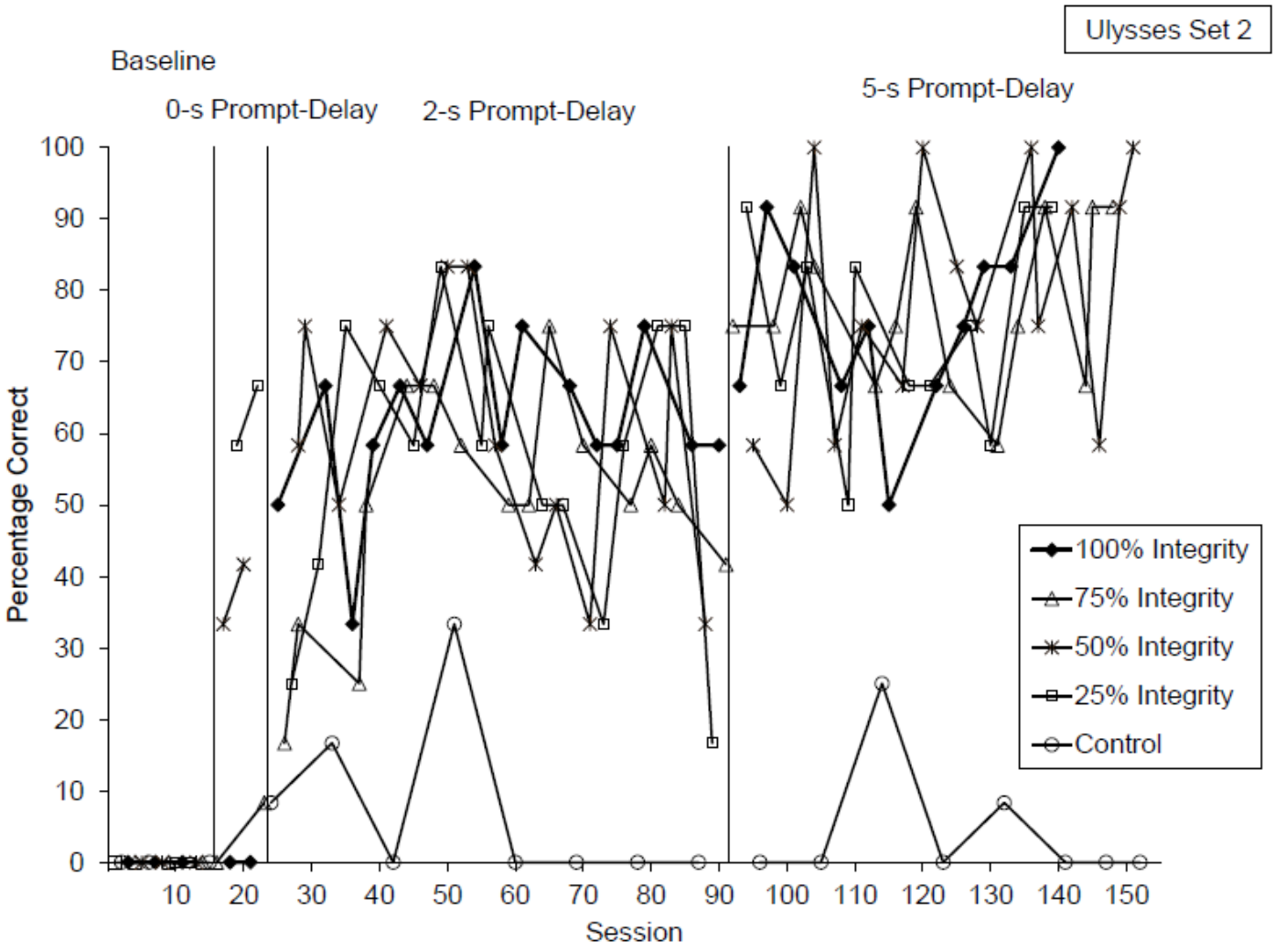

Figure 3. Percentage of correct responses for Ulysses for the control condition and the 100% Integrity, 75% Integrity, 50% Integrity, and 25% Integrity conditions for baseline, 0-s prompt-delay, 2-s prompt-delay sessions, and 5-s prompt-delay sessions for set 2 during the comparison of prompting errors (Experiment 1).

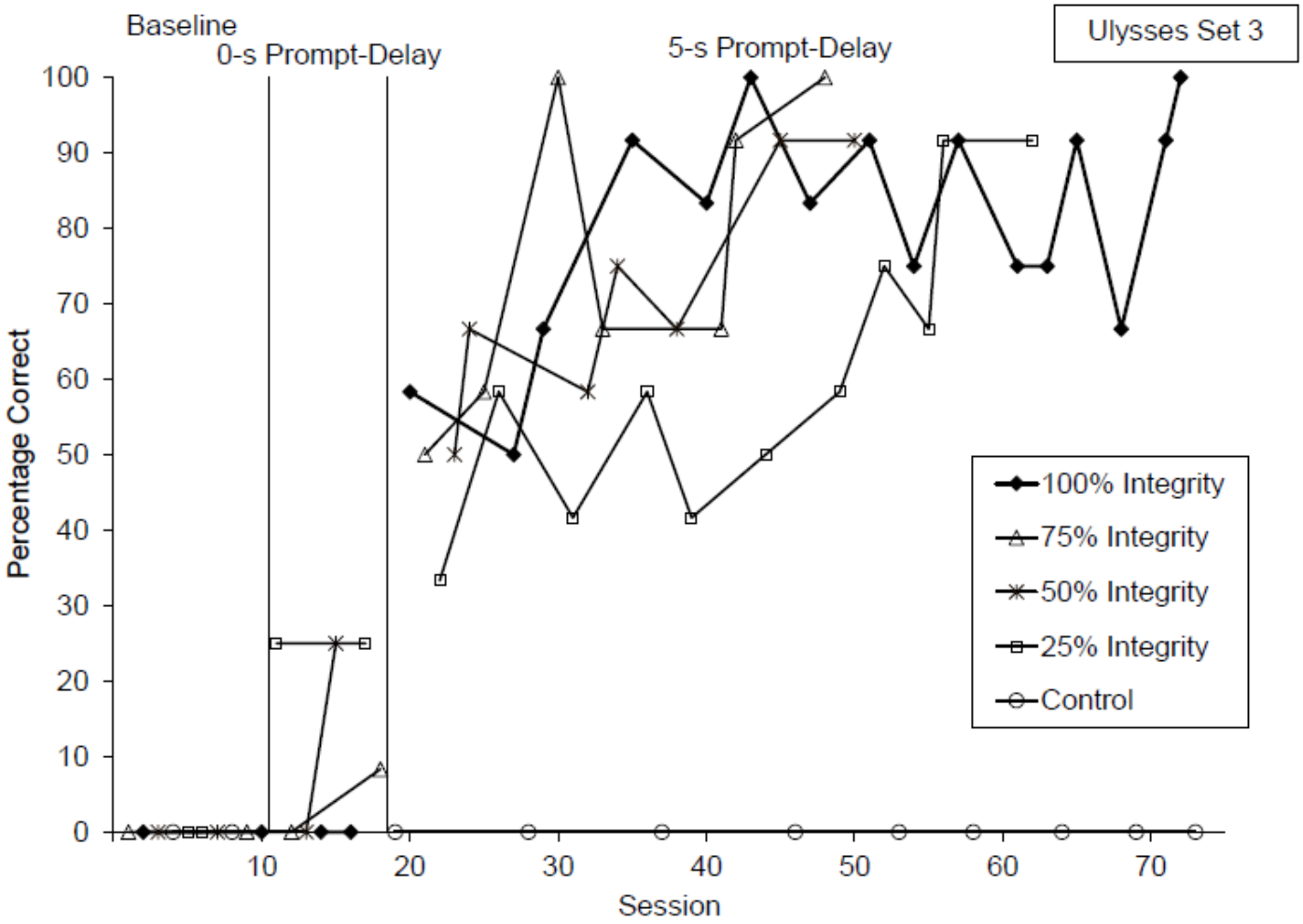

Figure 4. Percentage of correct responses for Ulysses for the control condition and the 100% Integrity, 75% Integrity, 50% Integrity, and 25% Integrity conditions for baseline, 0-s prompt-delay, and 5-s prompt-delay sessions for set 3 during the comparison of prompting errors (Experiment 1).

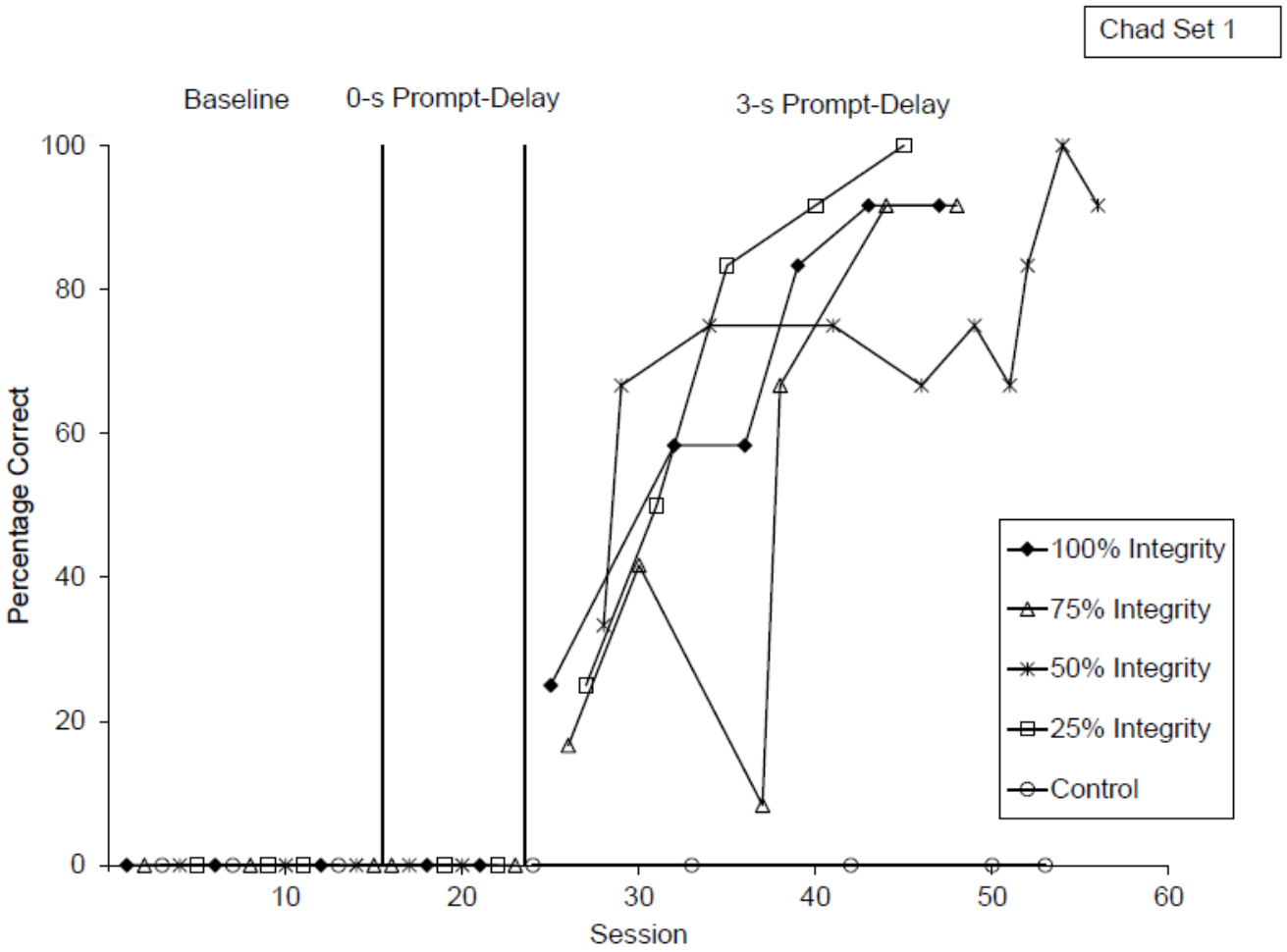

Figure 5. Percentage of correct responses for Chad for the control condition and the 100% Integrity, 75% Integrity, 50% Integrity, and 25% Integrity conditions for baseline, 0-s prompt-delay, and 3-s prompt-delay sessions during the comparison of reinforcement errors (Experiment 2).

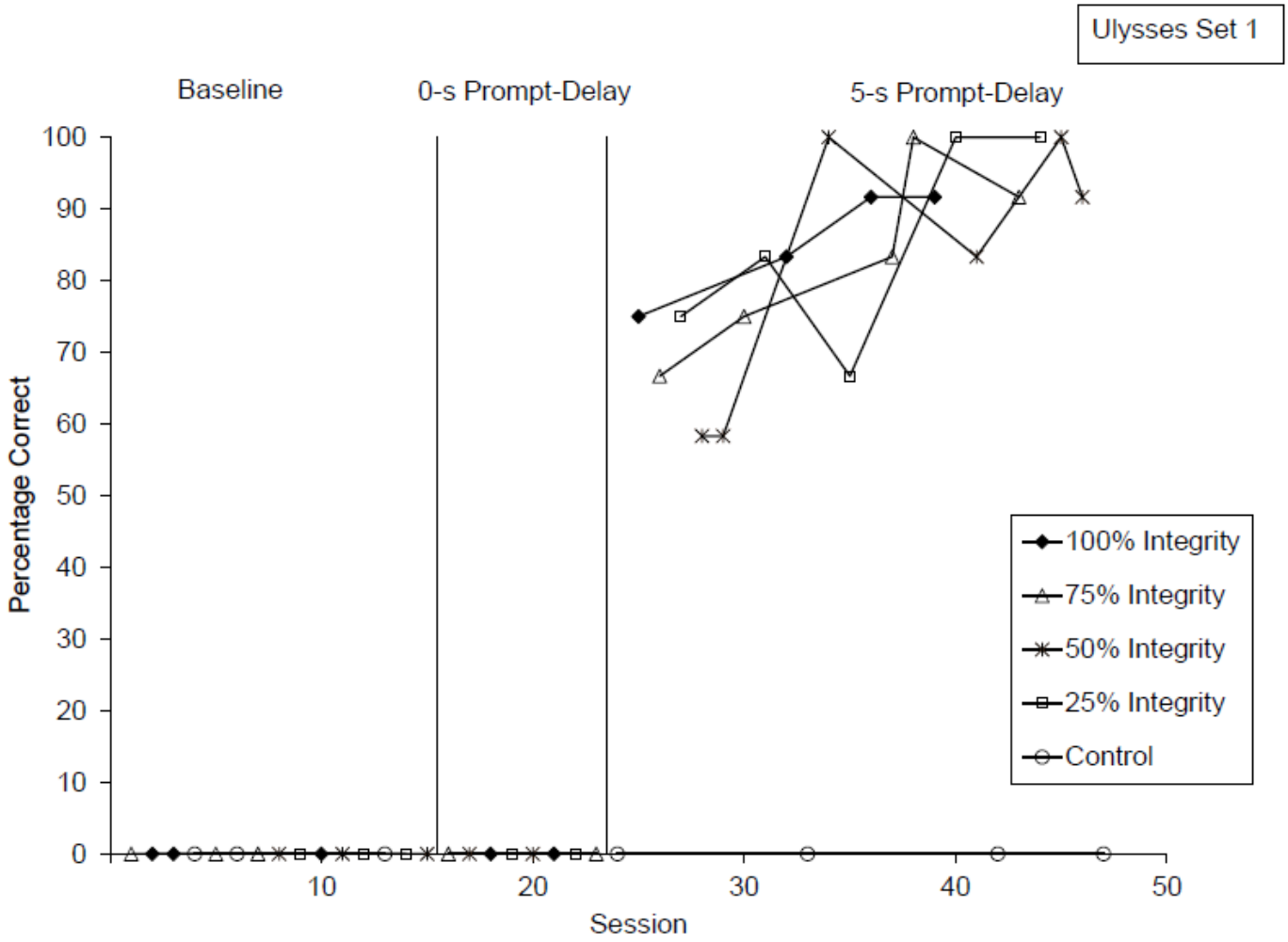

Figure 6. Percentage of correct responses for Ulysses for the control condition and the 100% Integrity, 75% Integrity, 50% Integrity, and 25% Integrity conditions for baseline, 0-s prompt-delay, and 5-s prompt-delay sessions during the comparison of reinforcement errors (Experiment 2).