Spring 5-4-2024

Evaluating the Effectiveness of Video Self-Monitoring on Staff Treatment Integrity during Covert and Overt Observations

Amber R. Paden
University of Nebraska Medical Center

EVALUATING THE EFFECTIVENESS OF VIDEO SELF-MONITORING ON STAFF TREATMENT INTEGRITY DURING COVERT AND OVERT OBSERVATIONS

by

Amber R. Paden

A DISSERTATION

Presented to the Faculty of

The University of Nebraska Graduate College

in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy

Medical Sciences Interdepartmental Area
Graduate Program

(Applied Behavior Analysis)

Under the Supervision of Professor Regina A. Carroll

University of Nebraska Medical Center
Omaha, Nebraska

May, 2024

Supervisory Committee:

Holly Roberts, Ph.D Nicole Rodriguez, Ph.D.

Amanda Zangrillo, Ph.D

ACKNOWLEDGMENTS

I could not have reached this goal without the help of many people in my life. I would like to take this opportunity to thank them for their support.

I would like to express my deepest gratitude to my dissertation advisor, Dr. Regina Carroll for her invaluable guidance, support, and encouragement throughout the entire process of completing my dissertation. Without her, I would have never started this journey and I would not be the clinician, researcher, and professional I am today.

I would also like to extend my sincere thanks to my dissertation committee, Holly Roberts, Ph.D., Nicole Rodriguez, Ph.D., and Amanda Zangrillo, Psy.D. Their guidance, support, and expertise throughout this project have been invaluable. I am deeply grateful for the time and dedication they provided.

Next, I'd like to acknowledge the emotional support provided by my coworkers and friends. They survived listening to my constant complaints and frustrations yet were still there to provide support and encouragement.

Last, but certainly not least, I'd like to thank my husband, Jordan, and boys Liam and Declan. They put up with my grumpy moods, took over some household chores, and helped me stay positive until the end. Now you can call me Dr. Mom!

ABSTRACT

Amber R. Paden, Ph.D.

University of Nebraska, 2024

Supervisor: Regina A. Carroll, Ph.D.

Previous research has shown low treatment integrity can lead to decreased effectiveness and efficiency of skill acquisition during discrete-trial instruction (DTI; Carroll et al., 2013). Pantermuehl and Lechago (2015) found that during covert observations, treatment integrity ranged from 18.6 to 76% whereas during overt observations, integrity was as high as 100%. This shows reactivity is an obstacle within DTI service delivery. It is important to investigate how reactivity affects staff performance and identify ways to increase and maintain high integrity. Finding a socially acceptable, effective, and efficient method to increase and maintain high levels of staff treatment integrity during covert observations is critical in the clinic setting. The purpose of the current study was to examine the effectiveness of video self-monitoring in increasing and maintaining high treatment integrity for staff implementing DTI during covert and overt observations. Participants included four staff working one-on-one with children with autism spectrum disorder (ASD) implementing DTI with less than 90% integrity during covert or overt observations. Results show video self-monitoring was effective at increasing staff treatment integrity and maintaining high integrity over time.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ASD	Autism Spectrum Disorder
BST	Behavioral Skills Training
DTI	Discrete Trial Instruction
EIBI	Early Intensive Behavioral Intervention
FR	Fixed Ratio
IOA	Inter Observer Agreement
TARF-R	Treatment Acceptability Rating Form - Revised
VMVO	Video Modeling with Voice-Over Instruction

CHAPTER 1: INTRODUCTION

Children diagnosed with autism spectrum disorder (ASD) are often referred to an early intensive behavior intervention (EIBI) program to help address skill deficits. EIBI focuses on increasing skills by focusing on areas such as communication, social behavior, and relationships (Cooper, 2021; Smith, 2001). The goal of EIBI is for the child to succeed in mainstream education when they enter elementary school and have minimal need for special education services. Cooper (2021) estimates that if a child fails to mainstream following EIBI, they will need another 15 years of special education services. It is important that teaching during EIBI is effective and efficient to maximize progress.

Discrete trial instruction (DTI) is a common procedure used in EIBI. Each trial is a discrete unit of instruction, typically consisting of five parts including a discriminative stimulus, prompt, response, consequence, and an intertrial interval (Smith, 2001). Correct DTI implementation is important to ensure children acquire the critical skills needed to be successful in mainstream education. Treatment integrity is the consistent and accurate implementation of intervention as it is intended by a treatment agent (Codding et al., 2005). During DTI, treatment integrity is often measured by collecting data on a therapist's correct or incorrect implementation of each component of a teaching trial. Previous research has shown poor treatment integrity can lead to slower progress toward mastering skills and high treatment integrity can positively affect client progress and behavior (Breeman et al., 2020; Carroll et al., 2013). High treatment integrity may lead to better clinical intervention outcomes and reduce levels of challenging behavior (Digennaro-Reed et al., 2010; Pelletier et al., 2010).

Carroll et al. (2013) evaluated the effects of three programmed treatment integrity errors (i.e., delivery of instructions, prompts, and reinforcement) on skill acquisition for

children with ASD during DTI. First, the researchers identified the most common treatment integrity errors teachers made during one-on-one instruction, by observing children with ASD in their school. They found the three most common treatment integrity errors included the teacher (a) omitting the delivery of a tangible item following a correct response, (b) withholding a prompt following an error, and (c) repeating the instruction more than once. Next, the researchers evaluated the effects of these integrity errors on skill acquisition by implementing the teaching procedures with high or low integrity. During the low integrity condition, integrity errors for the three most common errors were programmed during 67% of trials. This level was chosen because it approximated the level of treatment integrity observed in the natural environment and research has found integrity errors at this level are likely to lead to decreased performance. Overall, they found low levels of treatment integrity led to decreased effectiveness and efficiency of skill acquisition during DTI. Only one of the six participants mastered the targets in the low integrity condition and mastery took twice as long as in the high integrity condition.

Breeman et al. (2020) replicated and extended the work of Carroll et al. (2013) by evaluating the effects of different procedural integrity errors during DTI when teaching auditory-visual conditional discriminations. The researchers assessed the effects of different integrity errors and compared these effects to DTI delivered with perfect treatment integrity. The experimenters first conducted a descriptive assessment by observing clinicians implement DTI teaching procedures as they typically would. Results of the descriptive assessment were used to inform the programmed integrity errors during the low integrity condition. Programmed errors included incorrect tangible delivery, praise delivery, and error correction and occurred during 25% of trials. During the high integrity condition, all steps of the teaching procedure were conducted correctly. Skill acquisition was compared across control, low-integrity, and high-integrity conditions. This study found that mastery took as much as twice as long for targets

taught with low treatment integrity. These lower levels of treatment integrity can result in slower skill acquisition and potential long-term inefficiency of client programming and goal mastery. This, in turn, can result in the need for additional special education services once the child enters elementary school. The results of this study show that even a small decrease in integrity from 100% to 75% can double the teaching time needed to reach mastery level responding for the client. Therefore, identifying methods to ensure levels of treatment integrity remain as close to 100% as possible following training can be imperative to the success and progress of the client.

Behavioral skills training (BST) is a commonly used procedure for staff training (Reid et al., 2012; Sarokoff & Sturmey, 2004; Weston et al., 2020). During BST a supervisor first provides instructions to the staff member. Next, the supervisor models how to correctly implement the target skill followed by an opportunity to practice. Finally, the supervisor provides feedback to the staff member on their performance. Previous research has successfully used BST to teach staff members to implement teaching procedures with children with ASD. A study by Sarokoff and Sturmey (2004) evaluated the effectiveness of BST to train three special education teachers to implement DTI. Data were collected on teacher integrity with the 10 DTI components. During baseline, teachers were provided with a list and definitions of DTI components and were told to do their best to implement the procedure. Integrity was low for all teachers during baseline with mean scores of 43%, 49%, and 43%. Training consisted of BST and continued until teacher integrity reached 90% mastery criterion or higher across three consecutive sessions. During post-training, researchers provided no feedback to teachers. While integrity increased to a mean of 97%, 98%, and 99% for all teachers, it is unclear if this level of integrity was maintained over time. It is important to evaluate the maintenance of this performance over time. Maintaining high treatment integrity over time is crucial in ensuring the effectiveness of interventions and therapeutic approaches. Consistency in

implementing treatment protocols ensures individuals receive the intended benefits and achieve optimal outcomes.

Video modeling with voice-over instruction (VMVO) is another procedure commonly used to train staff to implement DTI procedures with high integrity. VMVO involves using a video recording to model skills with voice-over instructions that the viewer should then accurately imitate (Catania et al., 2009; Moore & Fisher, 2007; Vladescu et al., 2012). Catania et al. (2009) evaluated the effectiveness of VMVO on the integrity of three new staff conducting DTI. During baseline, staff were given a brief explanation of the teaching procedure and were then told to conduct the DTI session with an actor. Mean integrity for all staff was low (48%, 21%, and 63%). During the VMVO condition, staff watched a 7-min video with voiceover instructions depicting a DTI session. Within a few minutes of viewing the video, staff conducted a DTI session with an actor. VMVO sessions continued until staff performance stabilized, but no minimum criteria were noted. All staff showed an increase in mean integrity (90%, 85%, and 94%). During a 1-week follow-up, high integrity maintained. Data on integrity beyond one week were not reported so it is unclear if these results maintained.

Although we have several evidence-based procedures for teaching staff to implement DTI with high integrity (Catania et al., 2009; Reid et al., 2012; Sarokoff & Sturmey, 2004; Vladescu et al., 2012), few studies have examined the long-term maintenance of staff performance following initial training or procedures to improve performance when integrity declines over time. It is important to focus on increasing staff members' treatment integrity when implementing DTI and evaluating the long-term maintenance of high integrity.

Observer reactivity is another variable that needs to be considered when teaching staff to implement DTI with high integrity. The presence of a supervisor may influence staff behavior; thus, treatment integrity may be higher when a supervisor is

present (Brackett et al., 2007; Mowery et al., 2010; Pan et al., 2013; Pantermuehl & Lechago, 2015; Wu et al., 2018). This reactivity, or the influence an observer has on behavior (Brackett et al., 2007), may be a barrier to consistent high treatment integrity. Reactivity may be occurring because the supervisor has become a discriminative stimulus for performance feedback. Therefore, when the supervisor is present, the therapist changes their behavior to avoid being given corrective feedback. While researchers have noted concern with reactivity, this continues to be an obstacle in evaluating treatment integrity. It is important to investigate differences in treatment integrity with and without the supervisor present and identify ways to increase and maintain high levels of integrity at all times.

Mowery et al. (2010) investigated the influence of the presence of a supervisor on the effectiveness of staff management procedures in a group home setting. During treatment, BST was used to increase positive interactions between the staff and the clients. The experimenters evaluated the frequency of positive interactions when the supervisor was present and absent. Results of the study showed positive social interactions increased for two of the four staff only when the supervisor was present. This study shows that the presence of the supervisor can affect staff behavior and treatment integrity may decrease when the supervisor is absent. One limitation of this study is that while the experimenters collected data on positive interactions, they did not collect formal treatment integrity data on staff implementation of the treatment package.

Pantermuehl and Lechago (2015) compared the effects of overt observations, real-time internet-based observations, and covert observations on the treatment integrity of three staff during DTI. During baseline, covert observations occurred via inconspicuous video recording. The researchers found during baseline (covert observations only) treatment integrity ranged from 19 to 76% across all three participants. During overt observations, when the supervisor was present and feedback

was provided, integrity increased to as high as 100% (range 63–100%). Integrity during covert observations remained low for all participants, ranging between 36 to 67%. The increase in integrity during overt observations but not covert observations supports the occurrence of reactivity during overt observations. However, it is unclear if staff integrity was affected by the presence of the supervisor alone or because of the supervisor's presence and feedback. This presents a need to find methods that will lead to high treatment integrity in the absence of a supervisor. While the results of this study support the use of overt observations to increase staff treatment integrity, this is not a long-term solution because it is not possible for a supervisor to always be present.

Brackett et al. (2007) decreased the reactivity of two staff working to support adults with multiple disabilities during observations. The staff were trained to allow the workers to complete tasks independently. During overt observations, staff allowed independent completion of tasks. In contrast, during covert observations, staff completed tasks for the workers. The researchers implemented a self-monitoring procedure, which involved the staff collecting data on the level of prompting they provided. Following the implementation of this procedure, tasks completed by the staff during covert observations were reduced to levels like those during overt observations. This study shows the supervisor's presence influenced staff performance and a self-monitoring procedure effectively improved staff behavior. One limitation of the study was that the supervisor was present during all observations, although they attempted to observe discreetly during covert observations. Research is needed to identify methods to covertly observe that do not require the supervisor to be physically present (e.g., use of a video camera). Another limitation is that the supervisor provided feedback following overt observations but not following covert observations. Therefore, it is possible the feedback influenced staff behavior. It may also be the case that just the presence of the supervisor was a signal to the staff that feedback was forthcoming. It is important to parse apart

whether the change in staff behavior is due to feedback from the supervisor or the supervisor's presence alone. Brackett et al. (2007) is one of the few studies to measure long-term maintenance (up to 25 weeks). The results suggest self-monitoring may be an effective intervention to promote maintenance of high integrity and overcome reactivity.

Researchers have evaluated numerous procedures to improve treatment integrity (Hawkins & Heflin, 2010; Pelletier et al., 2010; Weston et al., 2020). One potential intervention that may improve treatment integrity, require less time from the supervisor, and help overcome reactivity is video self-monitoring. Video self-monitoring involves video recording one's own behavior and viewing the video later to collect data on the target behavior (Alexander et al., 2017; Pelletier et al., 2010; Hager, 2018). It can be used to improve many behaviors, including treatment integrity of staff providing DTI in the EIBI setting. Staff are trained to collect relevant data from videos of themselves implementing DTI, review the data, and identify skills to target for improvement. Video self-monitoring is a procedure that requires minimal time from a supervisor to implement and previous research has shown video self-monitoring to be an effective procedure for increasing treatment integrity (Alexander et al., 2017; Belfiore et al., 2008; Hawkins & Heflin, 2010; Pelletier et al., 2010; Weston et al., 2020).

Pelletier et al. (2010) showed a video self-monitoring treatment package was effective at increasing the integrity of behavioral program implementation for three school staff members. The treatment package consisted of an instructional video, video self-monitoring, and experimenter feedback. During video self-monitoring, staff watched a video from baseline of their implementation of the target behavioral program while working with a student and they recorded data on their own treatment integrity. Shortly after viewing their video and receiving feedback from the supervisor, the staff were again video recorded while implementing the target behavioral program. The experimenter was present while the staff worked with the student and scored treatment integrity. Integrity

increased and remained high for two of the three participants; therefore, video self-monitoring sessions were discontinued. The third participant completed a second video self-monitoring session and an increase in treatment integrity followed. The experimenters conducted a one-month follow-up probe for two participants and treatment integrity remained high. During treatment and maintenance sessions, the experimenter was present and collected treatment integrity data, therefore it is unclear if the increase and maintenance of treatment integrity was due to reactivity.

Belfiore et al. (2008) also found improvements in treatment integrity with the use of video self-monitoring. Following baseline, three school paraprofessional staff were video recorded implementing DTI while a supervisor watched from behind a one-way mirror. No feedback was provided during or after the baseline observations. The researchers then taught the paraprofessional staff to collect integrity data using a self-monitoring checklist. Once the staff completed the checklist with 90% or higher integrity, video self-monitoring was implemented. Following each DTI session, the staff watched the video from the session and used the self-monitoring checklist to collect treatment integrity data on themselves. The supervisor compared the staff checklist to their own to ensure data collection accuracy. When the staff met the criteria of 90% correct during DTI sessions across three out of four sessions, self-monitoring was no longer required. Maintenance was conducted 4 weeks after mastery and high integrity was maintained for one out of three participants. While the results of this study show video self-monitoring increased treatment integrity for all three participants, once video self-monitoring ended, high integrity was not maintained for two of the three participants.

Weston et al. (2020) evaluated the effects of video self-monitoring versus performance feedback on treatment integrity with preference assessment procedures. Four graduate students pursuing a degree in applied behavior analysis participated. The experimenters used an alternating treatments design embedded within a multiple

baseline across participants design. Data were collected on the percentage of steps in each preference assessment conducted correctly. During the pre-experimental phase, the students went through the first three steps of BST (i.e., instruction, modeling, rehearsal) to learn how to implement the preference assessments. During baseline, the experimenter provided a data sheet and instructed the student to conduct the target preference assessment with the child. Average treatment integrity ranged from 58% to 68%. Following baseline, video self-monitoring and performance feedback were randomly assigned to each of the preference assessments for each participant and counterbalanced across participants. The procedures used in the video self-monitoring condition were like those used by Pelletier et al. (2010). During the feedback sessions, the experimenter met with the student to review the treatment integrity data and provided praise and corrective feedback. While both procedures resulted in improved treatment integrity for all participants, the experimenters noted the time-intensive nature of the feedback procedure. During feedback, the supervisor is required to observe the entire session, collect data on treatment integrity, and meet with the participant to provide feedback after the session. In contrast, the video self-monitoring procedure could be set up to require minimal supervision time. This study also did not control for reactivity. The experimenter was present either collecting treatment integrity data or video recording the session. It is unclear if high integrity would be maintained without the experimenter present. The experimenters also did not collect data on integrity across time, therefore it is unclear whether high treatment integrity maintained. Research is needed to identify strategies that will lead to the maintenance of high treatment integrity. These results present a need to identify a procedure that will not only increase treatment integrity but will promote maintenance of high integrity over time. It is important to assess how often staff should be required to self-monitor to maintain high integrity during DTI sessions.

Another study that found video self-monitoring can lead to increases in integrity was conducted by Alexander et al. (2017). The researchers demonstrated video self-monitoring can be used to improve the instructional performance of two teachers. The teachers video recorded 5- to 15-min teaching sessions. After watching their videos, the teachers chose their own target behavior (e.g., praise rates, opportunities to respond). During video self-monitoring, the teachers recorded teaching sessions and later watched the video recording to collect data on their target behavior. For both participants, performance increased with the use of video self-monitoring. Alexander and colleagues reported both teachers indicated they found the procedure to be beneficial; however, they did not report any formal procedures for assessing social validity. In addition, none of the other studies reviewed on video self-monitoring assessed social validity. It is important to assess social validity to ensure procedures are acceptable, and outcomes are meaningful in the natural environment.

Finding a socially acceptable, effective, and efficient method to increase and maintain high levels of staff treatment integrity when no supervisor is present is critical in the clinic setting. The use of covert video recording methods could help reduce reactivity. Previous research has conducted 1-month follow-up sessions, but it is unclear if treatment integrity maintained beyond one month. In addition, only one of the reviewed studies (Alexander et al., 2017) assessed the social validity, but not formally. The purpose of the current study was to examine the effectiveness of video self-monitoring in maintaining high treatment integrity for staff implementing DTI during covert and overt observations by a supervisor. First, we assessed the extent to which treatment integrity varied during covert and overt observations. Second, we compared the effects of video self-monitoring during covert and overt observations on staff treatment integrity. Third, we assessed whether we could systematically fade the schedule of video self-monitoring observations while still maintaining high treatment integrity during both covert and overt

observations. Finally, we assessed the social validity of the video self-monitoring intervention.

CHAPTER 2: METHOD

Participants, Setting, and Materials

Participants included four female therapists who worked one-on-one with children with ASD at a community-based early intervention clinic. Each therapist was paired with one child who they worked with regularly at the clinic. The therapists were Associate Behavior Technicians who had completed clinic training and obtained certification as a Registered Behavior Technician®. Participant 1 had a high school degree and had 11 months of experience before participating in the evaluation. Participant 2 had obtained a bachelor's degree and had been working in the clinic for 6 months before participating. Participant 3 also had 6 months of experience and had a high school degree. Participant 4 was enrolled in an undergraduate program and had 1 year and 7 months of experience.

Before beginning the study, the experimenter informed all therapists in the clinic that video recording could take place at any time throughout the day. This is not uncommon for the setting as video recordings are sometimes used for data collection purposes. Therapists were told to contact the experimenter if they wanted to opt out of video recording. No therapists opted out of video recording.

All sessions were conducted in the child's normal treatment area in the clinic. Typical treatment areas were 1.5 m by 1.5 m and included two small chairs and a table. Partitions were present on at least three sides of the child to minimize distractions. Materials present included teaching stimuli (i.e., flashcards), timers, data sheets, and preferred items. The therapist sat across the table from the participant.

All sessions (covert and overt) were video recorded using a GoPro® HERO video camera. Video cameras were placed in the treatment area at least two weeks before beginning the study to desensitize the therapists to its presence. The recording light was disabled on the camera so the staff could not see when the camera was recording, and the video camera was turned on and off when the therapist was not present.

Dependent Measures and Data Collection

The experimenter collected integrity data from the video recordings. Data were collected on each component of DTI. The components included (a) inter-trial interval, (b) establishing ready behavior, (c) delivering the instruction, (d) delivering the reinforcer, (e) delivering the prompt, (f) implementing error correction, (g) ignoring problem behavior, (h) reinforcement duration, and (i) collecting data (see Table 1 for definitions). The primary dependent variable was therapist treatment integrity with DTI. The experimenter summarized treatment integrity by calculating the total number of components implemented correctly for all 12 trials divided by the total number of opportunities to implement a component for all 12 trials and multiplied by 100. The mastery criterion for treatment integrity during video self-monitoring was two consecutive sessions with 90% correct or greater. The mastery criterion during fading was one session with 90% correct or greater integrity for the fading step. In addition, we calculated the integrity for each DTI component. We did so by dividing the number of times the component was implemented correctly by the total number of opportunities to implement the component and multiplied by 100. This allowed us to look at the therapist's integrity with each individual component of DTI.

Interobserver Agreement (IOA) and Procedural Integrity

A trained observer recorded data on therapist treatment integrity during covert and overt observations. The observer scored the therapist's accuracy in implementing

the nine components for each discrete trial. A second observer collected data during an average of 35% of the sessions (range, 31–37%) in each condition. We calculated trial-by-trial IOA by comparing primary and secondary observers' data for each of the nine components for all 12 trials. We scored an agreement if both observers scored the component the same. A disagreement was scored if observers recorded different responses for a component. We divided the number of components across all trials in agreement by the total number of components across all trials with agreements plus disagreements and converted the ratio to a percentage. Overall, IOA averaged 95% (range, 90–100%).

A second observer scored the experimenter's procedural integrity with conducting video self-monitoring training during 100% of sessions. The observer scored each behavior as correct, incorrect, or not applicable. The experimenter responses for video self-monitoring training included (a) provided instructions and materials for protocol review, (b) provided the instruction to view the training PowerPoint®, (c) provided an opportunity for questions, (d) provided the instruction to score procedural integrity from the video, (e) scored the therapist's integrity data sheet, and (f) provided feedback to the therapist on data collection accuracy. Experimenter integrity was 100% for all sessions.

A trained observer also scored procedural integrity during an average of 72% (range, 43–100%) of covert and overt observation sessions. The observer scored each behavior as correct, incorrect, or not applicable. The experimenter's responses for overt observations included (a) entering the treatment area, (b) informing the therapist an observation was beginning, (c) sitting approximately 6 ft from the therapist, (d) collecting integrity data, (e) refraining from providing corrective feedback, (f) answering any questions asked, and (g) providing a general statement before exiting the treatment area. For covert observations, the experimenter's response included not being present in

the treatment area. For video self-monitoring, experimenter responses included (a) providing data sheet and operational definitions, (b) presenting a laptop with the PowerPoint®, (c) providing verbal instructions for PowerPoint®,(d) starting the PowerPoint®, (e) asking if the therapist had questions and answered questions, (f) providing verbal instructions for video, (g) starting the video, (h) comparing data, and (i) providing positive and constructive feedback. The experimenter calculated procedural integrity by dividing the number of components the experimenter implemented correctly divided by the number of components implemented correctly plus the number of components implemented incorrectly and multiplying by 100. Experimenter integrity was 100% for all covert and overt observation sessions scored.

Pre-study Training

Before participating in the current study all therapists had completed initial training provided to all new staff at the clinic during their first 60 days of employment. During initial training, therapists viewed a 15-min PowerPoint® presentation with voiceover instructions outlining the nine components of DTI. The presentation contained video examples of each component and reviewed how to use the datasheet for data collection. The therapists then watched five different videos depicting a therapist implementing DTI with an actor. The therapists collected data on child responses for each session and the trainer checked the therapists' datasheet for accuracy. The trainer provided feedback to the therapist regarding any errors made. The therapists then participated in five, 12-trial DTI role-play sessions with the trainer where the trainer played the role of an actor child. The trainer provided both positive and constructive feedback during the role-plays as well as additional opportunities to practice following

errors. Following the completion of the role-plays, therapists began conducting DTI sessions with children under the observation of the trainer. The trainer collected integrity data on the therapists' implementation of DTI. Following the observation, the trainer reviewed the integrity data with the therapist. The trainer provided corrective feedback for steps performed incorrectly, as well as praise for steps conducted correctly. Therapists were required to conduct DTI with 90% or higher integrity for two sessions for five different teaching procedures before training ended.

Pre-Assessment and Target Identification for Child Participants

A receptive identification task was identified for each child participant based on their current treatment goals. The experimenter conducted a pre-assessment to identify a minimum of 30 targets for teaching. During the pre-assessment, the experimenter first ensured ready behavior (i.e., child sitting with bottom in the chair, oriented toward the therapist or instructional material, no disruptive movements). The experimenter placed three cards in an array on the table approximately 6 inches in front of the child and approximately 2 inches apart and secured attending by gesturing to the array if needed. The position of the target card was counterbalanced quasi-randomly across trials. After the child had looked at the cards, the experimenter provided the instruction (e.g., touch apple) and waited 5 s for the child to respond. No consequences were provided for correct or incorrect responses. Following every two to three trials, the experimenter presented a mastered receptive identification task. If the child responded correctly to the mastered task, the experimenter provided praise and access to a preferred item for 30 s. If the child responded incorrectly or did not respond, the experimenter used hand-over-hand guidance to prompt the correct response and moved to the next trial. Each target

was presented up to four times during the pre-assessment. The purpose of the pre-assessment was to identify unknown targets for each child.

Targets were included in the teaching sessions if the participant responded correctly during one or fewer presentations. Three to six targets were taught at a time (based on what was typical for each participant). When a target reached mastery criterion (i.e., 100% correct for two consecutive sessions) that target was removed and replaced with a new target.

Discrete-trial Instruction with Perfect Integrity

Below is a description of the DTI procedure that was used throughout this study. The teaching procedure was a constant 5-s prompt delay with error correction. The therapist first ensured ready behavior (i.e., sitting with their bottom in the chair, oriented toward the therapist, and not engaging in any disruptive movements). The therapist placed three cards in an array on the table approximately 6 inches in front of the child and approximately 2 inches apart and secured attending by gesturing to the array if needed. The position of the target card was counterbalanced quasi-randomly across trials. After the child had looked at the cards, the therapist provided the instruction (e.g., touch apple) and waited 5 s for the child to respond. If the child responded correctly, the therapist provided descriptive verbal praise and an edible or 30-s access to a preferred item on a fixed ratio (FR) 1 schedule. If the child responded incorrectly or did not respond within the prompt delay, the therapist hand-over-hand guided the correct response. The therapist then began a re-present until independent error-correction procedure (Carroll et al., 2015). Following hand-over-hand guidance, the therapist re-represented the trial and continued to re-present the trial until the child responded correctly to the instruction or until three error-correction trials were presented without a correct independent response. If the child responded correctly to the instruction during

error correction, the therapist provided descriptive praise. If the child did not respond correctly during error correction, then the therapist ended the trial.

Simulated Videos for Training

Simulated-session videos were created and used to train the therapist to collect treatment integrity data. During the simulated session videos, the actor therapist and the actor child sat at a table in a small 2 m by 2 m room. All materials needed for the session were present (e.g., flashcards, datasheet, leisure items). Six different 10-min simulated-session videos were created. In each video, the actor therapist conducted 12 trials of a receptive identification task with an actor child. During each video, the actor therapist implemented between four to six component skills accurately and three to five component skills incorrectly. For example, an error for *prompt delivery* may have included the actor therapist waiting more than 5s to physically guide the correct response following an error or no response. An error for *ignoring problem behavior* may have included the actor therapist providing attention following a disruptive behavior displayed by the actor child. Errors were predetermined for each session. The number and type of errors for each component were varied and balanced across videos. The actor therapist engaged in both correct and incorrect components of DTI during the videos to ensure the therapist was exposed to trials with high integrity and trials with low integrity for data collection practice.

During each session, the actor child also engaged in predetermined responses to the actor therapist's instructions. The actor child engaged in predetermined responses such as complying with the initial instruction (range, 6–8 trials), engaging in an error (range, 1–3 trials), or not responding to the initial instruction (range, 2–4 trials). For incorrect trials, a predetermined number of error-correction trials were conducted and the actor child either complied during error correction before the third error-correction

trial or did not comply during error correction. Instances of challenging behavior (e.g., disruptions, aggression) were predetermined for the simulated session as well. Challenging behavior occurred on 1 to 3 trials during three of the six videos.

Experimental Design and General Procedures

We used a multi-element design embedded within a concurrent multiple baseline across participants design to evaluate the effects of video self-monitoring on treatment integrity during covert and overt observations. We also evaluated long-term maintenance as we systematically faded the schedule of self-monitoring. Each DTI session consisted of 12 trials and each target was presented an equal number of times per session. Throughout the study, the therapist conducted one session during each shift the therapist was scheduled with the target child. The order of the observations was randomized across sessions. The teaching procedure was a constant 5-s prompt delay with error correction.

Overt observations

During overt observations, the experimenter entered the treatment area carrying a clipboard or notebook and pen. The experimenter made a general statement about conducting an observation such as “I’d like to watch you conduct the receptive identification program during this sit.” Overt observations by a supervisor were typical and occurred a few times per week. The experimenter sat approximately 2 m from the therapist and observed the therapist implement DTI. The experimenter collected integrity data during the observation but did not provide positive or corrective feedback. If the therapist asked a question, the experimenter provided an answer because that is what they would typically do during an observation. The experimenter documented the number and types of questions asked. At the end of the observation, the experimenter

provided a general statement thanking the therapist for conducting the program and exited the treatment area.

Covert observations

Covert observations consisted of collecting integrity data via a video recording of a session where the experimenter was not present in the treatment area during the implementation of the session. The video camera was turned on and off without the therapist present and the recording light was disabled so the therapist did not know when the camera was recording.

Baseline

A minimum of three baseline sessions were conducted for both covert and overt observations before beginning video self-monitoring. Therapist integrity was scored for each session. To participate in the study, integrity during baseline needed to be below 90% during covert observations.

Video Self-Monitoring

The therapists were trained on how to collect integrity data on a therapist's performance implementing DTI. Training on data collection took place after baseline and before video self-monitoring (Pelletier et al., 2010). Training took place in a small conference room with a table and chairs where both the therapist and experimenter were present. The experimenter gave the therapist a data sheet and operational definitions for treatment integrity data collection and provided instructions that included a description of the materials. The experimenter started the PowerPoint® with voice-over instructions and stayed in the room to watch the video with the therapist. The PowerPoint® contained the operational definition of each component, an example of the data sheet with descriptions of each component of the data sheet, and a description of how to collect treatment integrity data on the data sheet. After the therapists viewed the

PowerPoint®, the experimenter asked the therapist if there were any questions and answered all questions. Next, the experimenter provided the therapist with additional integrity data sheets (if needed) and instructed the therapist to collect treatment integrity data on the therapist's performance while watching the simulated session video of two actors (i.e., an actor playing the therapist and an actor playing the child) implementing DTI. Following each trial, the experimenter paused the video and reviewed the therapist's data for that trial. The experimenter provided positive feedback for components with agreement (e.g., nice job with [list correctly scored components]) and corrective feedback for components with disagreements. If corrective feedback was needed, the experimenter replayed the trial, described the error, and provided a rationale for change. The experimenter provided another opportunity for questions and answered all questions.

The experimenter then provided instructions and a handout to the therapist about independently collecting integrity data for two different simulated session videos along with two data sheets. The therapist was given 48 hours to score both videos and return the data sheets to the experimenter. Within 48 hours the experimenter scored the therapist's accuracy with data collection by comparing the therapist's completed data sheet with the video key. The experimenter scored an agreement if both observers recorded the same response for a component and a disagreement if observers recorded different responses. Within 48 hours after checking the data sheets, the experimenter met with the therapist to provide feedback. The experimenter provided positive feedback (i.e., behavior-specific praise) for agreements and corrective feedback (i.e., described the error and provided rationale) for any disagreements following each session. If the therapist did not reach mastery after the first two videos, additional videos were assigned, and training continued. Training was completed when the therapist scored two consecutive videos with 90% accuracy.

Within three days of completing the training on data collection, therapists were asked to watch a video recording of their own implementation of DTI during a covert observation and collect integrity data. The therapist sat in a small conference room equipped with a table and chairs. A pen, timer, datasheet, and operational definitions were on the table next to the laptop. The experimenter informed the therapist they could review the operational definitions and data sheet before scoring the video, and they then placed a laptop on the table with a video of the therapist completing DTI on the screen. The experimenter provided instructions to the therapist to watch the video and collect integrity data. The therapist was told they could re-watch any portion of the video if needed, and to open the door when they were finished. The experimenter independently scored the same video recordings to obtain interobserver agreement. For the first video self-monitoring session, the therapist scored their most recent baseline session. Within 72 hours of the video self-monitoring session, a covert and overt observation was conducted. The therapist then scored integrity for their most recent covert observation session, while the experimenter scored integrity for both the covert and overt sessions. Sessions continued until the therapist's treatment integrity met mastery criteria of two consecutive sessions at 90% or higher for covert sessions.

If the therapist's accuracy with scoring their sessions fell below 90% for two consecutive sessions, the experimenter provided feedback to the participant regarding errors with data collection during the previous sessions.

Fading

The schedule the therapists collected self-monitoring data was faded to weekly, bi-weekly, and then monthly. The therapist scored their integrity from a recent video recording of a session. The schedule was faded if the therapists' integrity was 90% or higher across two consecutive sessions.

Maintenance

A covert observation maintenance probe was conducted one month after the final fading session. Procedures were the same as they were in baseline, and no video self-monitoring sessions were conducted before the observation.

Generalization

Generalization probes were scheduled to occur during baseline, following mastery in the self-monitoring condition, and following the final fading step. During generalization sessions, the therapist was paired with a different child and conducted the same DTI procedure. The children identified for generalization were children the therapist worked with regularly. Sessions were conducted using the procedures described in the covert observation section.

Debrief

After completion of the follow-up session, participants were debriefed. The experimenter explained the purpose of the evaluation, outlined covert and overt observations, reviewed the graph of their performance, provided an explanation of skills to continue to focus on based on errors made, and provided contact information for further questions.

Social validity

After completing the study, we asked the therapists to complete a social-validity questionnaire. This was done to assess the social acceptability of the procedures used in this study. We used a modified version of the Treatment Acceptability Rating Form Revised (TARF-R; Reimers et.al., 1992). The questionnaire consisted of 16 questions using a Likert-type scale to assess the effectiveness of the procedure, the therapist's willingness to use the procedure again, and the disruptiveness of the procedure (see

Appendix A). Therapists indicated their agreement or disagreement with each statement with higher scores indicating greater agreement and acceptability.

CHAPTER 3: RESULTS

Figure 1 displays the data for Participants 1 and 2. Mean treatment integrity during baseline for Participant 1 (top panel of Figure 1) was an average of 76% (range, 68–83%) during overt observations and slightly lower at an average of 65% (range, 64 – 67%) during covert observations. Treatment integrity during the generalization probe was 78%. Following the introduction of video self-monitoring, we observed only a slight increase in treatment integrity. Participant 1 was provided with feedback on their data collection accuracy during video self-monitoring sessions. They were consistently making errors with data collection for prompt delivery and error correction. Following feedback, treatment integrity increased to at or above the mastery criteria of 90%. After the mastery criterion was met, a generalization probe was conducted, and treatment integrity remained low at 67%. Video self-monitoring sessions were then systematically faded out. Treatment integrity remained above the criteria for all covert sessions throughout the fading process. Following the last monthly session, a generalization probe was conducted, and treatment integrity was 65%. A maintenance probe for Participant 1 was not able to be scheduled due to the therapist going on long-term leave from their position at the clinic.

During baseline sessions for Participant 2 (bottom panel of Figure 1), treatment integrity was an average of 65% (range, 59–71%) during overt observations and an average of 56% (range, 38–68%) during covert observations. Treatment integrity during the generalization probe was 78%. Following the introduction of video self-monitoring, an immediate increase in treatment integrity was observed. During this time, Participant 2’s attendance became sporadic due to repeated illnesses, and sessions were conducted

weekly. The decision was made to continue fading video self-monitoring sessions because treatment integrity was above the mastery criteria of 90%. Participant 2 was not at the clinic for a month, therefore fading moved from weekly to monthly. Following the final monthly session, integrity during the generalization session was 99%.

Figure 2 shows the data for participants 3 and 4. Mean treatment integrity during baseline for Participant 3 (top panel of Figure 2) averaged 49% (range, 46–51%) during overt observations and averaged 46% (range, 36–52%) during covert observations. Integrity during the generalization probe was 59%. Following the introduction of video self-monitoring, treatment integrity immediately increased to above mastery criteria. Mastery was met after 2 video self-monitoring sessions. Treatment integrity during generalization was above mastery criteria as well. Treatment integrity remained at or above criteria for most sessions, with only 2 covert (86% and 87%) and 1 overt (79%) falling below criteria. A final follow-up session was conducted, and treatment integrity remained above the criteria. During fading, the child paired with therapist 3 was discharged from services, so a new child was identified for the remainder of the study. A similar skill acquisition program using the target procedure was used for all sessions.

Mean treatment integrity during baseline for Participant 4 (bottom panel of Figure 2) was an average of 47% (range, 35–54%) during overt observations and averaged 38% (range, 25–54%) during covert observations. Treatment integrity during the generalization probe was 63%. Following the introduction of video self-monitoring, treatment integrity immediately increased but was below mastery for 2 covert sessions. As with Participant 1, feedback was provided to Participant 4 on data collection accuracy during video self-monitoring sessions. Following feedback, integrity during covert observations increased and was above the mastery criteria of 90%. During fading, treatment integrity remained above the mastery criteria for all sessions. During fading,

the child paired with Participant 4 ended services at the clinic, and a new child was identified for the remainder of the study.

Figures 3 and 4 show participants' accuracy in scoring their own treatment integrity from video. Overall, participants' accuracy in scoring their own treatment integrity was high for all participants. Participants 1 and 4 received feedback regarding data collection accuracy following 2 sessions below 90% accuracy. Following feedback, accuracy immediately increased and remained above 90%. Once fading began, no feedback was provided if accuracy dropped below 90%.

Social Validity

Social validity data were collected for all participants following the completion of the study (see Table 2). Responses were reviewed to gain an understanding of the therapists' views of acceptability, effectiveness, and ease of implementation of the intervention. Each question was scored on a scale of 1 to 7 with a total of 112 possible points. Total acceptability scores ranged from 84 to 102 (M = 95), indicating general intervention acceptability. The mean item rating across all participants was 5.9 (out of 7). The range of scores was 3 to 7. Only one question was rated 3 across all participants. The question that received this rating was a question indicating participants were not confident others would be willing to help them implement the video self-monitoring procedure. Two participants rated the question regarding the ease of carrying out the procedure in their existing routine as 4 (neutral), whereas the other two participants rated it as 7 (very easy). Consistent highly rated areas included a willingness to carry out the procedure, the likeliness of the procedure to make permanent improvements in behavior, experiencing no discomfort, and low levels of disruption resulting from the procedure.

Post-hoc Error Analysis

After the evaluation, an error analysis was conducted. For each session, we analyzed treatment integrity for each component of DTI. Black boxes indicate a skill that the participant implemented accurately during 100% of opportunities during the session, striped boxes indicate a skill the participant implemented with accuracy between 80% and 99%, gray boxes indicate accuracy between 50% and 79%, and white boxes indicate accuracy below 50%.

The first graphs in Appendix B show the results for Participant 1. During baseline, Participant 1 was at or near mastery for skills 2 (establish ready behavior), 3 (instruction), and 9 (data collection). The lowest integrity occurred with skills 4 (reinforcer delivery), 5 (prompt delivery), and 6 (error correction). Following video self-monitoring training, integrity for skills 1 (intertrial interval), 2 (ready behavior), 3 (instruction), and 7 (ignoring problem behavior) increased to 100%. There was some variability with integrity for other skills, but integrity was at or above 80% for skills during most sessions, except for generalization sessions. During the final generalization session, 6 of the 9 skills were below 80% integrity.

The second set of graphs in Appendix B show the results for Participant 2. For participant 2, integrity with the 9 skills was variable during baseline. The lowest integrity occurred with skills 4 (reinforcer delivery), 5 (prompt delivery), 6 (error correction), and 8 (reinforcement duration). Following video self-monitoring training, integrity with all skills increased, although variability with integrity was seen for all skills.

The third set of graphs in Appendix B show the results for Participant 3. During baseline, Participant 3 had the lowest integrity with skills 2 (establishes ready behavior), 4 (reinforcer delivery), 5 (prompt delivery), 6 (error correction), 7 (ignoring problem behavior), and 8 (reinforcement duration). Following video self-monitoring training,

integrity increased across all skills. Integrity was the most variable for skills 5 (prompt delivery), 8 (reinforcement duration), and 9 (data collection).

The final set of graphs in Appendix B show the results for Participant 4. During baseline, Participant 4 was not at mastery level for any skill. Skills 1 (establishes ready behavior), 4 (prompt delivery), and 7 (reinforcement duration) were at or near mastery for some sessions. Following video self-monitoring training, integrity with all skills increased, although variability continued. Most skills were at or above 80% with the most variability occurring with skills 1 (intertrial interval) and 8 (reinforcement duration).

CHAPTER 4: DISCUSSION

The current study examined the effectiveness of video self-monitoring in increasing and maintaining high treatment integrity for staff implementing DTI during covert and overt observations by a supervisor. All participants exhibited treatment integrity below our target of 90%, with some sessions as low as 25%, during baseline. Prior to the study, all participants had the skill of performing DTI sessions with high integrity in their repertoires. All therapists were previously trained to criteria (i.e., 90% or higher integrity for two sessions across five teaching procedures) before being allowed to independently conduct DTI sessions with children. For all participants, we observed low treatment integrity during both covert and overt observations during baseline suggesting a skill deficit with correct DTI implementation. Similar to the results found by Pantermuehl and Lechago (2015), Participant 1 displayed lower treatment integrity during covert observations when compared to overt observations during baseline, suggesting some reactivity to the supervisor's presence. Taken together these results suggest Participant 1 also exhibited a performance deficit with some clear reactivity. For Participant 1, the supervisor could be a setting event that changed how she responded to the potential for reinforcement, or punishment, based on her history with the

supervisor and her expectations of the outcome of the observation. The supervisor may have served as a discriminative stimulus signaling to the participant that their behavior could lead to certain consequences. Positive reinforcement (e.g., praise, approval) or negative reinforcement (e.g., avoiding criticism) is more likely to occur when a supervisor is present, therefore an increase in integrity with the DTI protocol could have been due to the direct or indirect social cues provided by the supervisor. Reactivity is likely influenced by historical interactions between the participant and supervisor, immediate social context, and the complex interaction of reinforcement and punishment contingencies. Although skill and performance deficits were factors for Participant 1, all four participants demonstrated a skill deficit. Therefore, we examined the effectiveness of a video self-monitoring procedure to increase treatment integrity.

It is important for treatment integrity with implementing DTI procedures to remain as close to 100% as possible to ensure the child makes progress toward treatment goals as research has shown even small degradations in integrity can negatively impact outcomes for children with ASD (e.g., Breeman et al., 2020; Carroll et al., 2013). As noted by Breeman et al., delivering treatments with limited effectiveness goes against consumers' rights to receive the most impactful interventions. Additionally, as per the Ethics Code for Behavior Analysts (Behavior Analyst Certification Board, 2022), behavior analysts are required to implement effective treatments and address any hindering environmental factors. The current study found video self-monitoring was effective in increasing treatment integrity for all four participants. These results are consistent with previous studies finding video self-monitoring to be effective at increasing staff performance (Alexander et al., 2017; Belfiore et al., 2008; Pelletier et al., 2010; Weston et al., 2020). It is possible that continued video self-monitoring increased their awareness of DTI components, and it allowed them to receive feedback by engaging in self-monitoring and reflection (Alexander et al., 2017). During video self-monitoring

sessions, staff had to discriminate between their accurate and inaccurate implementation of DTI components. It is possible this could aid staff in distinguishing between correct and incorrect components enabling them to adjust their performance accordingly (Belfiore et al., 2008).

Not only did we find improved treatment integrity following video self-monitoring sessions, results were also maintained during follow-up sessions. Limited research has been conducted on the long-term maintenance of high treatment integrity following training or strategies to ensure performance does not diminish over time (Belfiore et al., 2008; Catania et al., 2009; Pelletier et al., 2010; Sarokoff & Sturmey, 2004). For example, Sarokoff and Sturmey found BST to be effective at increasing teacher integrity with implementing DTI, but they did not report on the maintenance of high integrity. Catania et al. found VMVO was effective at increasing staff integrity with DTI during training and the researchers conducted a maintenance probe one week after training ended. Although they found integrity remained high after one week, it is unclear if these results were maintained past one week. It is important to know if the behavior change has generality and will persist over time. To measure generality, the current study assessed the long-term maintenance of increased treatment integrity. Integrity with DTI was measured continuously for up to 18 weeks following the conclusion of treatment. The schedule of video self-monitoring sessions was faded to weekly, bi-weekly, and once per month. Integrity during both overt and covert observations between video self-monitoring sessions remained above 90% for all participants. The goal was to fade the frequency of video self-monitoring sessions to a schedule that would be manageable in a clinic setting. Although conducting sessions once per month should be a manageable schedule, future research could continue to fade video self-monitoring sessions to see how lean the schedule could be while still promoting high treatment integrity. Variable schedules could also be assessed as these schedules promote steady responding and

are more resistant to extinction. It would be beneficial to assess the social acceptability of the schedule of video self-monitoring sessions with both staff and supervisors to ensure the schedule that promotes maintenance is also acceptable by those who will need to implement it.

The current study's results also support video self-monitoring as a socially acceptable method to increase staff treatment integrity. No previous studies employing a video self-monitoring procedure to increase treatment integrity have formally measured social validity. Alexander et al. (2017) informally assessed the social validity of video self-monitoring by asking the participants if they found the procedure to be beneficial. This study contributes to the literature by formally assessing the social validity of the video self-monitoring intervention (Bishop et al., 2015; Hager et al., 2012). The results of the social validity questionnaire for the current study were favorable for all four participants suggesting the video self-monitoring procedure is not only effective at increasing and maintaining high treatment integrity but also a method accepted by therapists. Training procedures that have high social validity are more likely to be sustained over time (Austin & Carr, 2000; Geller & Lehman, 1990, Wolf, 1978). Ensuring therapists find the training procedures socially acceptable and can see the positive effect they have on their behavior and the services they provide may be important for maintaining behavior change. Assessing social validity will help ensure training is not only effective but also meaningful, acceptable, and useful.

The current investigation has several limitations. First, we conducted the self-monitoring intervention with only one child and one DTI program. It is possible that although we took steps to reduce reactivity by placing video cameras in the work areas weeks before the start of the study, disabling the recording light on the cameras, and turning the cameras on and off covertly, participants were able to predict which program was being covertly observed because those were the only sessions that they scored

during video self-monitoring. This only seemed to be a problem for Participant 1, who had previously demonstrated reactivity during overt observations. Participant 1 even disclosed during the debriefing session after the study that she thought she knew when the experimenter would be recording sessions because the videos that she scored during self-monitoring were always with the same client and the same DTI program. Thus, for individuals who demonstrate reactivity in baseline, it may be important to use indiscriminable contingencies to support maintenance and generalization (Freeland & Noell, 2002). Experimenters could use a variety of clients and programs when conducting video self-monitoring sessions so the individual is not able to discriminate the conditions under which sessions may be occurring as well as employing intermittent schedules for session occurrence. For the remaining participants, the increase in integrity proved to have generality as the behavioral change occurred with novel children (Baer et al., 1968, Stokes & Baer, 1977).

A related limitation was that we only observed generalization for three of the four participants. For Participant 1, who clearly demonstrated a performance and skill deficit, treatment integrity was low when DTI was conducted with a novel child. As mentioned above, it is likely that the self-monitoring intervention was not sufficient to decrease reactivity. Thus, future research could focus on a functional approach to intervention selection. Researchers should identify therapists who show a difference in integrity during covert and overt observations (i.e., lower integrity during covert observations), which would suggest reactivity may be at play, as well as therapists who show low integrity during both covert and overt observations, which would suggest a skill deficit. One approach used in the field of organizational behavior management is the performance analysis. This is often done using the Performance Diagnostic Checklist-Human Services and is used to identify the variables that are influencing a staff member's unsatisfactory job performance (Austin, 2000). The performance analysis

results are then used to develop an intervention linked to the function of the performance problem. Therefore, the intervention should be chosen to match the therapist's response in baseline (Carr et al., 2013). Additional training, such as video self-monitoring, would be appropriate for a skill deficit as it would teach the therapist the skills they are implementing incorrectly and allow them to track their progress toward implementation with high integrity. If a performance deficit is identified, it may be more appropriate to provide direct feedback regarding the difference in the therapist's integrity during covert and overt observations (Austin et al., 1999). Working with the therapist to identify and address barriers to correct performance could also be effective.

Another limitation was that generalization sessions were not conducted following the intended schedule. Due to issues with conducting research in an applied setting (e.g., attendance, ending services, long-term leave), the generalization children were not the same as in baseline and after treatment for all participants. However, we replaced the original generalization child with a novel child not paired with the video self-monitoring treatment. Three of the four participants continued to implement DTI with high integrity when working with a novel child, thus demonstrating generality.

The current study suggests several avenues for future research related to reducing reactivity and maintaining treatment integrity over time with the implementation of behavioral interventions. Future research should prioritize exploring strategies to minimize reactivity in clinic settings, as it can significantly impact the validity of observations, quality of data collected, and integrity of services delivered. This could involve investigating methods to reduce staff awareness of being observed, such as video recording, live streaming for in vivo observation, or one-way mirrors. They could also investigate implementing naturalistic observation protocols such as unstructured observations and time sampling. These naturalistic observation protocols are valuable

tools for studying behavior in applied settings, and they allow researchers to capture accurate behaviors and interactions as they naturally occur.

The current study adds to the literature by monitoring the long-term maintenance of high treatment integrity with DTI implementation following mastery (Belfiore et al., 2008; Pelletier et al., 2010; Weston et al., 2020). Maintenance sessions were conducted for up to 18 weeks and high treatment integrity was maintained as we reduced the frequency of video self-monitoring sessions to monthly. Future research could focus on determining an appropriate schedule of covert observations and treatment sessions to ensure the intervention remains effective and maintains high treatment integrity.

After the completion of the study, we implemented a post-hoc error analysis to review the data for any error patterns. Although this analytical approach was not used to inform in-the-moment treatment decisions throughout the study regarding mastery or systematic fading of the video self-monitoring sessions, the analysis did not reveal any significant error patterns for any of the participants. Future research could benefit from incorporating real-time error analysis as a methodological approach. This practice could enable responsive adjustments based on participant performance, guiding decisions on when to consider a skill as mastered and when to initiate fading. This approach aligns with best practices in behavioral research, ensuring interventions are individualized based on participant performance.

This study specifically examined the contingencies influencing staff behavior, rather than focusing on the children’s skill acquisition. We aimed to analyze the factors that influence staff behavior without directly observing the children's learning outcomes. An appropriate next step for future research would be to investigate the effects of staff integrity on children’s skill acquisition. Improvements in child performance that occur in conjunction with the increase in treatment integrity can act as a significant indicator of the intervention’s effectiveness and validity (Martin et al., 2015). This would build upon

the findings of the current study by examining the link between staff treatment integrity and the subsequent changes in children's skill acquisition.

As clinicians and researchers, we must recognize different forms of supervision impact treatment integrity. By addressing these areas, future research can offer comprehensive insights into optimizing treatment integrity across various interventions and contexts, ultimately enhancing the effectiveness of clinical practices.

REFERENCES

Alexander, M., Williams, N. A., & Nelson, K. L. (2017). When you can't get there: Using video self-monitoring as a tool for changing behaviors of pre-service teachers. *Rural Special Education Quarterly*, *31*(4), 18-24. https://doi.org/10.1177/875687051203100404

Austin, J. (2000). Performance analysis and performance diagnostics. In J. Austin & J.E. Carr (Eds.), *Handbook of Applied Behavior Analysis*, pp. 321-349

Austin, J., & Carr, J. E. (2000) Practical strategies for increasing treatment integrity in school settings. *Behavior Analysis in Practice, 2*(1), 17-24.

Austin, J., Carr, J. E., & Agnew, J. L. (1999) The need for assessment of maintaining variables in OBM. *Journal of Organizational Behavior Management,19*(2), 59-87. https://doi.org/10.1300/J075v19n02_05

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*(1), 91-97. https://doi.org/10.1901/jaba.1968.1.91

Behavior Analyst Certification Board. (2020). *Ethics code for behavior analysts.* https://bacb.com/wp-content/ethics-code-for-behavior-analysts/

Belfiore, P. J., Fritts, K. M., & Herman, B. C. (2008). The role of procedural integrity: Using self-monitoring to enhance discrete trial instruction. *Focus on Autism and Other Developmental Disabilities*, *23*(2), 95-102. https://doi.org/10.1177/1088357607311445

Brackett, L., Reid, D. H., & Green, C. W. (2007). Effects of reactivity to observations on staff performance. *Journal of Applied Behavior Analysis*, *40*, 191–195. https://doi.org/10.1901/jaba.2007.112-05

Breeman, S. L., Vladescu, J. C., DeBar, R. M., Grow, L. L., & Marano, K. E. (2020). The effects of procedural integrity errors during auditory-visual conditional

discrimination training: A preliminary investigation. *Behavioral Interventions*, *35*(2), 203-216. https://doi.org/10.1002/bin.1710

Carr, J. E., Wilder, D. A., Majdalany, L., Mathisen, D., & Strain, L. A. (2013), An assessment-based solution to a human service employee performance problem. *Behavior Analysis in Practice, 6*(1), 16-32. https://doi.org/10.1007/BF03391789

Carroll, R. A., Joachim, B. T., St Peter, C. C., & Robinson, N. (2015). A comparison of error-correction procedures on skill acquisition during discrete-trial instruction, *Journal of Applied Behavior Analysis*, *48*(2), 257-273. https://doi.org/10.1002/jaba.205

Carroll, R. A., Kodak, T., & Fisher, W. W. (2013). An evaluation of programmed treatment integrity errors during discrete trial instruction, *Journal of Applied Behavior Analysis*, *46*(2), 379-394. https://doi.org/10.1002/jaba.49

Catania, C. N., Almeida, D., Liu-Constant, B., & Digennaro Reed, F. D. (2009). Video modeling to train staff to implement discrete-trial instruction. *Journal of Applied Behavior Analysis*, *42*(2), 387– 392. https://doi.org/10.1901/jaba.2009.42-387

Codding, R. S., Feinberg, A. B., Dunn, E. K., & Pace, G. M. (2005). Effects of immediate performance feedback on implementation of behavior support plans. *Journal of Applied Behavior Analysis*, *38*(2), 205– 219. https://doi.org/10.1901/jaba.2005.98-04

Cooper, M. (2021). Reducing special education costs by providing early intervention for autistic children. *Behavioral Interventions*, *37*(2), 397-414. https://doi.org/10.1002/bin.1839

Digennaro-Reed, F. D., Codding, R., Catania, C. N., & Maguire, H. (2010). Effects of video modeling on treatment integrity of behavioral interventions. *Journal of Applied Behavior Analysis*, *43*(2), 291-295. https://doi.org/10.1901/jaba.2010.43-291

Freeland, J. T. & Noell, G. H. (2002). Programming for maintenance: An investigation of delayed intermittent reinforcement and common stimuli to create indiscriminable contingencies. *Journal of Behavioral Education, 11*(1), 5-18. https://doi.org/10/1023/A:1014329104102

Geller, E. S. & Lehman, G. R. (1990). The social validity of behavioral interventions: A social ecology perspective. *Behavior Analysis and Social Action, 8*(2), 15-23. https://doi.org/10.1007/s10803-024-06297-3

Hager, K.D. (2012) Self-monitoring as a strategy to increase student teachers' use of effective teaching practices. *Rural Special Education Quarterly, 31*(4), 9-17. https://doilorg/10.1177/875687051203100403

Hager, K.D. (2018) Teachers' use of video self-monitoring to improve delivery of effective teaching practices. *Teaching Exceptional Children*, *50*(5), 283-290. https://doi.org/10.1177/0040059918765749

Hawkins, S. M. & Heflin, L. J. (2010). Increasing secondary teachers' behavior-specific praise using a video self-monitoring and visual performance feedback intervention. *Journal of Positive Behavior Interventions*, *13*(2), 97-108. https://doi.org/10.1177/1098300709358110

Martin, C.A., Drasgow, E., and Halle, J.W. (2015). Training teachers to enhance the play skills of young children with developmental disabilities during outdoor time by embedding instructional interactions. *Journal of Early Intervention*, *37*(4), 247-269. https://doi.org/10.1177/1053815115620209

Moore, J. W., & Fisher, W. W. (2007). The effects of videotape modeling on staff acquisition of functional analysis methodology. *Journal of Applied Behavior Analysis*, *40*(1), 197-202. https://doi.org/10.1901/jaba.2007.24-06

Mowery, J. M., Miltenberger, R. G., & Weil, T. M. (2010). Evaluating the effects of reactivity to supervisor presence on staff response to tactile prompts and self-

monitoring in a group home setting. *Behavioral Interventions*, *25*(1), 21–35. https://doi.org/10.1002/bin.296

Pan, S. C., Tien, K. L., Hung, I. C., Lin, Y. J., Sheng, W. H., Wang, M. J., Chang, S. C., Kunin, C. M., Chen, Y. C. (2013). Compliance of health care workers with hand hygiene practices: Independent advantages of overt and covert observers. *PLOS ONE 8*(1): e53746. https://doi.org/10.1371/journal.pone.0053746

Pantermuehl, R. M., & Lechago, S. A., (2015). A comparison of feedback provided in vivo versus an online platform on the treatment integrity of staff working with children with autism. *Behavior Analysis in Practice*, *8*, 291-222. https://doi.org/10.1007/s40617-015-0059-y

Pelletier, K., McNamara, B., Braga-Kenyon, P., & Ahearn, W. H. (2010). Effect of video self-monitoring on procedural integrity. *Behavioral Interventions*, *25*(4), 261-274. https://doi.org/10.1002/bin.316

Reid, D. H., Parsons, M. B., & Green, C. W. (2012). Monitoring staff performance. Supervisor's guidebook: Evidence-based strategies for promoting work quality and enjoyment among human service staff.

Reimers, T. M., Wacker, D. P., Cooper, L. J., & De Raad, A. O. (1992). Acceptability of behavioral treatments for children: Analog and naturalistic evaluations by parents. *School Psychology Review*, *21*(4), 628–643. https://www.tandfonline.com/action/showCitFormats?doi=10.1080%2F02796015.1992.12087371

Sarkoff, R. A., & Sturmey, P. (2004). The effects of behavioral skills training on staff implementation of discrete-trial teaching. *Journal of Applied Behavior Analysis*, *37*(4), 535-538. https://doi.org/10.1901/jaba.2004.37-535

Smith, T. (2001). Discrete trial training in the treatment of autism. *Focus on Autism and Other Developmental Disabilities*, *16*(2), 86-92. https://doi.org/10.1177/108835760101600204

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*(2), 349-367. https://doi.org/10.1901/jaba.1977/10-349

Vladescu, J. C., Carroll, R., Paden, A., & Kodak, T. M. (2012). The effects of video modeling with voiceover instruction on accurate implementation of discrete-trial instruction. *Journal of Applied Behavior Analysis*, *45*(2), 419-423. https://doi.org/10.1901/jaba.2012.45-419

Weston, R., Davis, T. N., Radhakrishnan, S., O'Guinn, N., & Rivera, G. (2020). Comparing performance feedback and video self-monitoring within a BST package to train pre-service behavior analysts to conduct preference assessments. *Journal of Behavioral Education*, *29*, 571-583. https://doi.org/10.1007/s10864-019-09333-w

Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis, 11*(2), 203-214. https://doi.org/10.1901/jaba.1978.11-203

Wu, K. S., Lee, S. S. J., Chen, J. K., Chen, Y. S., Tsai, H. C., Chen, Y. J., Huang, Y.H., & Lin, H. S. (2018). Identifying heterogeneity in the Hawthorne effect on hand hygiene observation: A cohort study of overtly and covertly observed results. *BMC Infectious Diseases*, *18*(1), 1-8. https://doi.org/10.1186/s12879-018-3292-5

Table 1: *The Nine Steps of DTI*

	Dependent Measure	*Operational Definition*
1.	Inter-trial Interval	Therapist initiates the start of the trial within 2 s (+/- 2 s) of the end of the last trial or the removal of a preferred item following a reinforcement interval.
2.	Establishes Ready Behavior	Therapist waits to present the instruction until the child is sitting with his/her bottom in the chair, is oriented toward the therapist or instructional material, and is not engaging in any disruptive movements with his/her hands and feet.
3.	Instruction	Therapist delivers the instruction exactly as it is specified in the protocol. The instruction does NOT include additional words or the child's name.
4.	Reinforcer Delivery	Therapist provides praise and a tangible item (within 2 s) following a correct response to the instruction. If tangible reinforcement is removed for error-correction trials, then the therapist provides praise only for a correct response during an error-correction trial.
5.	Prompt Delivery	Therapist delivers a correct prompt (based on the protocol) immediately following an error (within 1s) or following no response within the scheduled prompt delay (+/- 1s).
6.	Error Correction	Following a prompt, the therapist removes instructional materials (if applicable), and then re-presents the trial. The therapist continues to re-present the trial until the child responds correctly to the instruction or until the therapist has re-presented the trial 3 times without a correct response.
7.	Ignore Problem Behavior	Therapist attempts to block problem behavior (e.g., prevent the child from sweeping materials off the table). Following problem behavior, the therapist does not comment on the problem behavior. If problem behavior occurs during an inter-trial interval, the therapist does not delay the onset of the next demand (i.e., a demand is presented within 2s). If problem behavior occurred when a demand was in place, the therapist does not remove the demand.
8.	Reinforcement Duration	Therapist lets the child play with the tangible item for the correct duration (i.e., 25s +/- 5 s).
9.	Data collection	Therapist collects data following the end of one trial and before the start of the next trial. Therapist collects data accurately.

Table 2: *Social Validity*

Participant	*Total*	*Average Rating*	*Minimum Rating*	*Maximum Rating*
1	102	6.38	5	7
2	91	5.69	4	7
3	103	6.48	4	7
4	84	5.25	3	7

Figure 1: *Percentage Correct Integrity for Participants 1 and 2*

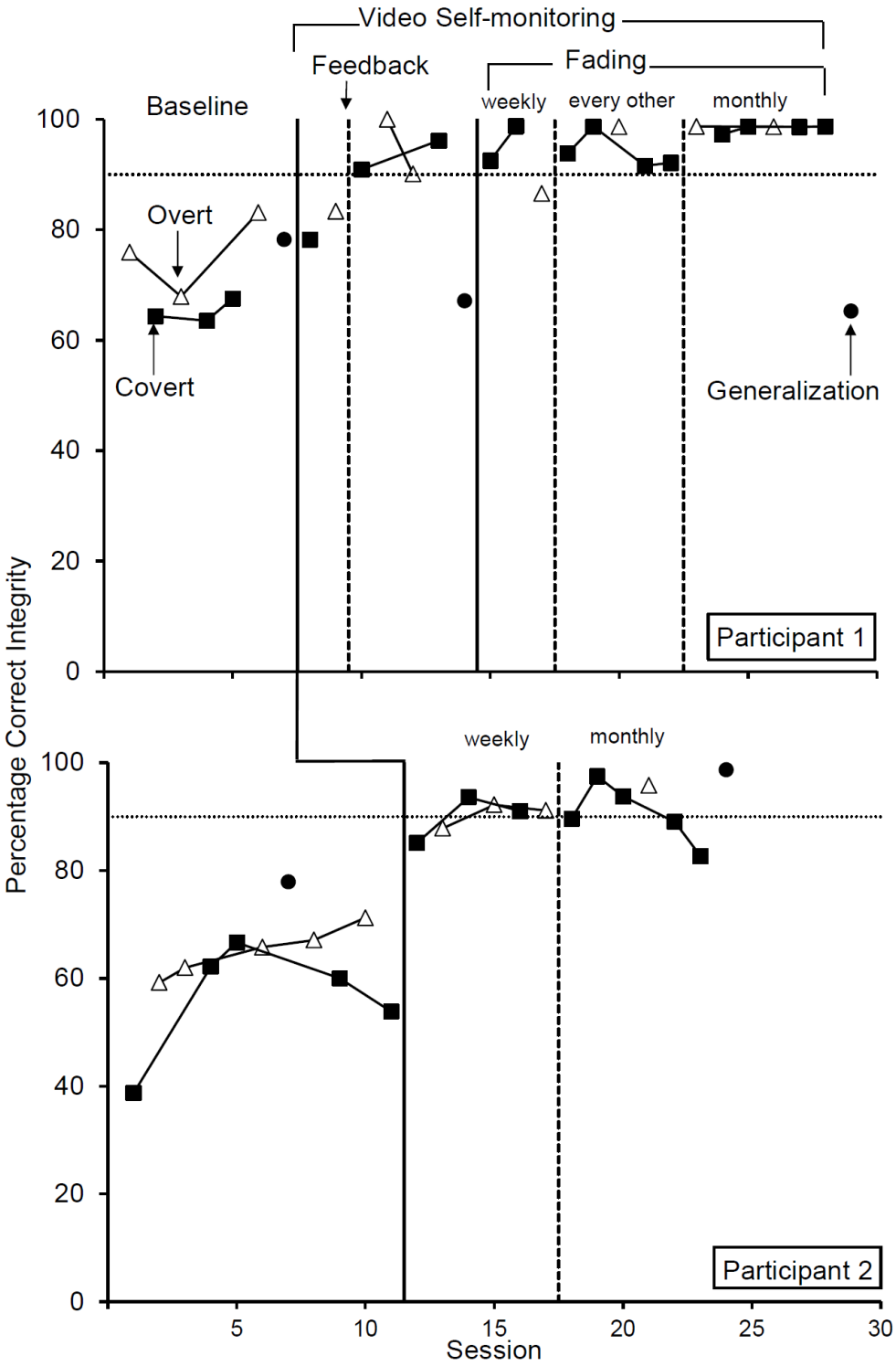


Figure 2: *Percentage Correct Integrity for Participants 3 and 4*

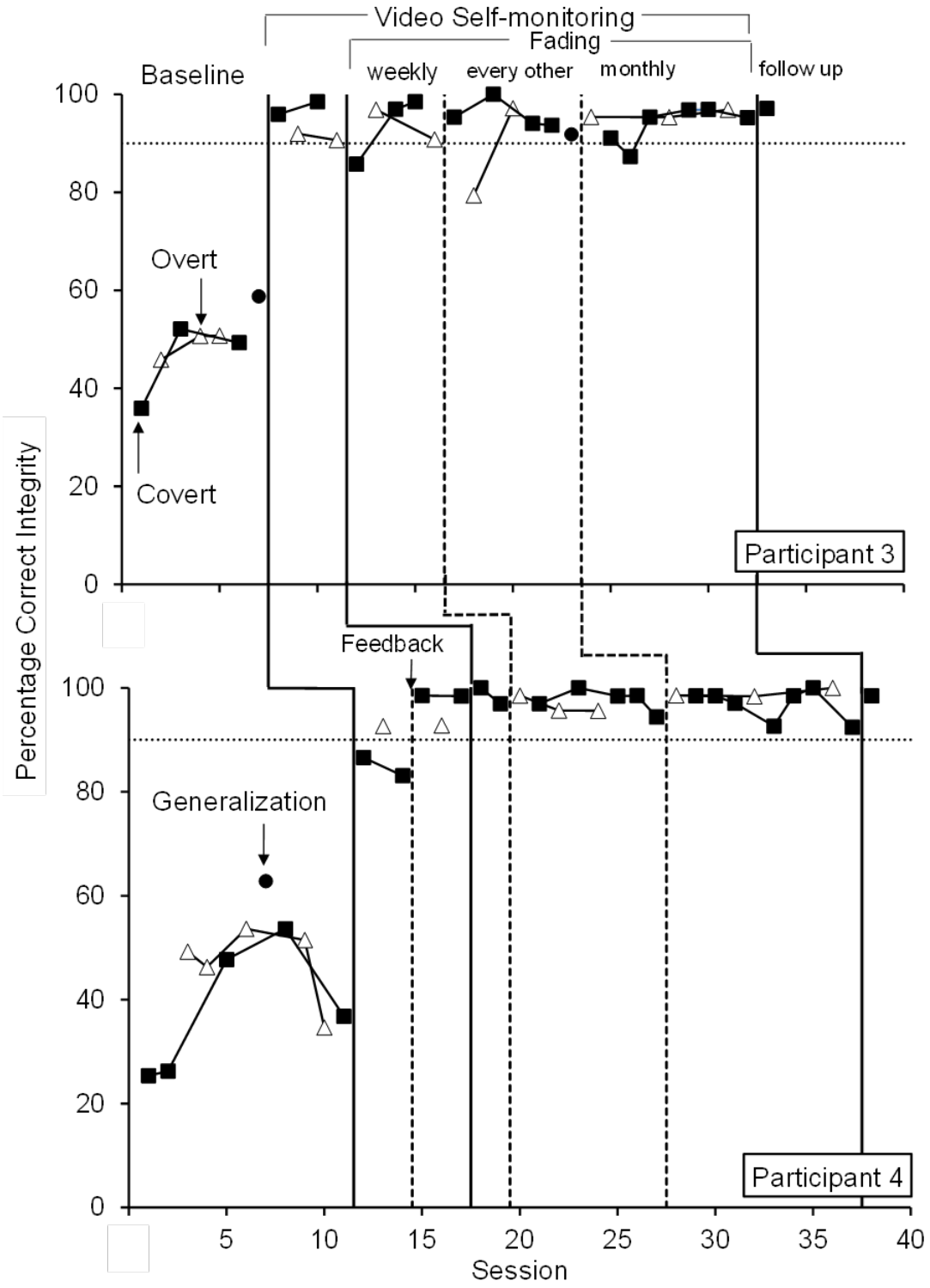


Figure 3: *Percentage Accuracy with Integrity Data Collection for Participants 1 and 2*

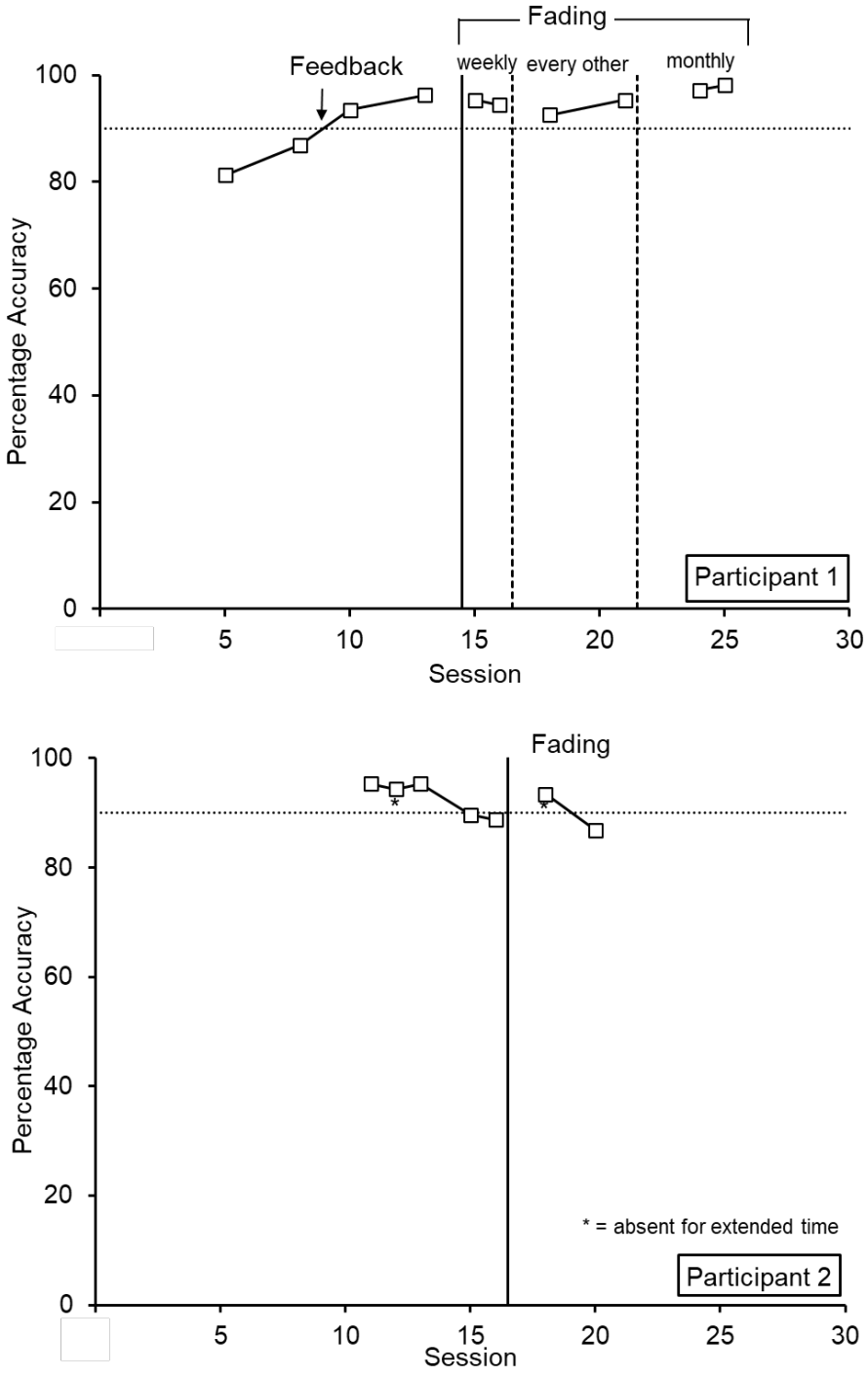
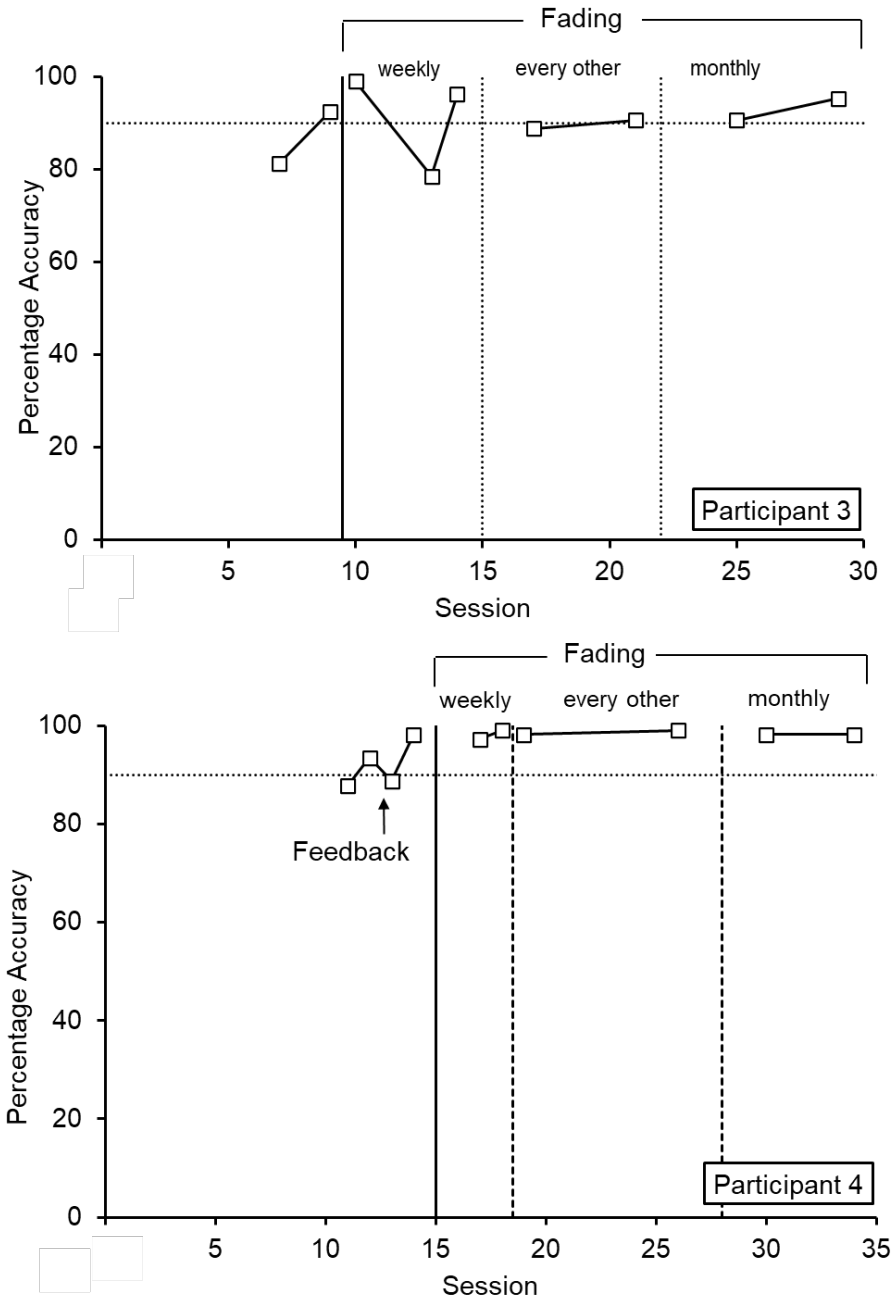


Figure 4: *Percentage Accuracy with Integrity Data Collection for Participants 3 and 4*



Appendix A: Social Acceptability Rating Form

Modified from the TARF-R

(Reimers, T., Wacker, D., Cooper, L., & DeRaad, A., 1992)

Please complete the items listed below by placing a check mark in the box under the question that best indicates how you feel about the procedure.

1. How acceptable do you find the procedure to be regarding increasing your treatment integrity?

Not at all acceptable			Neutral			Very acceptable

2. How willing are you to carry out this procedure?

Not at all willing			Neutral			Very willing

3. How costly will it be to carry out this procedure?

Not at all costly			Neutral			Very costly

4. To what extent do you think there might be disadvantages in following this procedure?

None are likely			Neutral			Many are likely

5. How likely is this procedure to make permanent improvements in your behavior?

Unlikely acceptable			Neutral			Very likely

6. How much time will be needed each day for you to carry out this procedure?

Little time needed			Neutral			Much time needed

7. How confident are you that this procedure will be effective?

Not at all confident			Neutral			Very confident

8. How disruptive will it be to carry out this procedure?

Not at all disruptive			Neutral			Very disruptive

9. How effective is this procedure likely to be for you?

Not at all effective			Neutral			Very effective

10. How affordable is this procedure?

Not at all affordable			Neutral			Very affordable

11. How much do you like the procedures used?

Do not like them at all			Neutral			Like them very much

12. How willing will other staff members be to help carry out this procedure?

Not at all willing			Neutral			Very willing

13. To what extent are undesirable side effects likely to result from this procedure?

No side effects likely			Neutral			Many side effects likely

14. How much discomfort are you likely to experience during this procedure?

No discomfort at all			Neutral			Very much discomfort

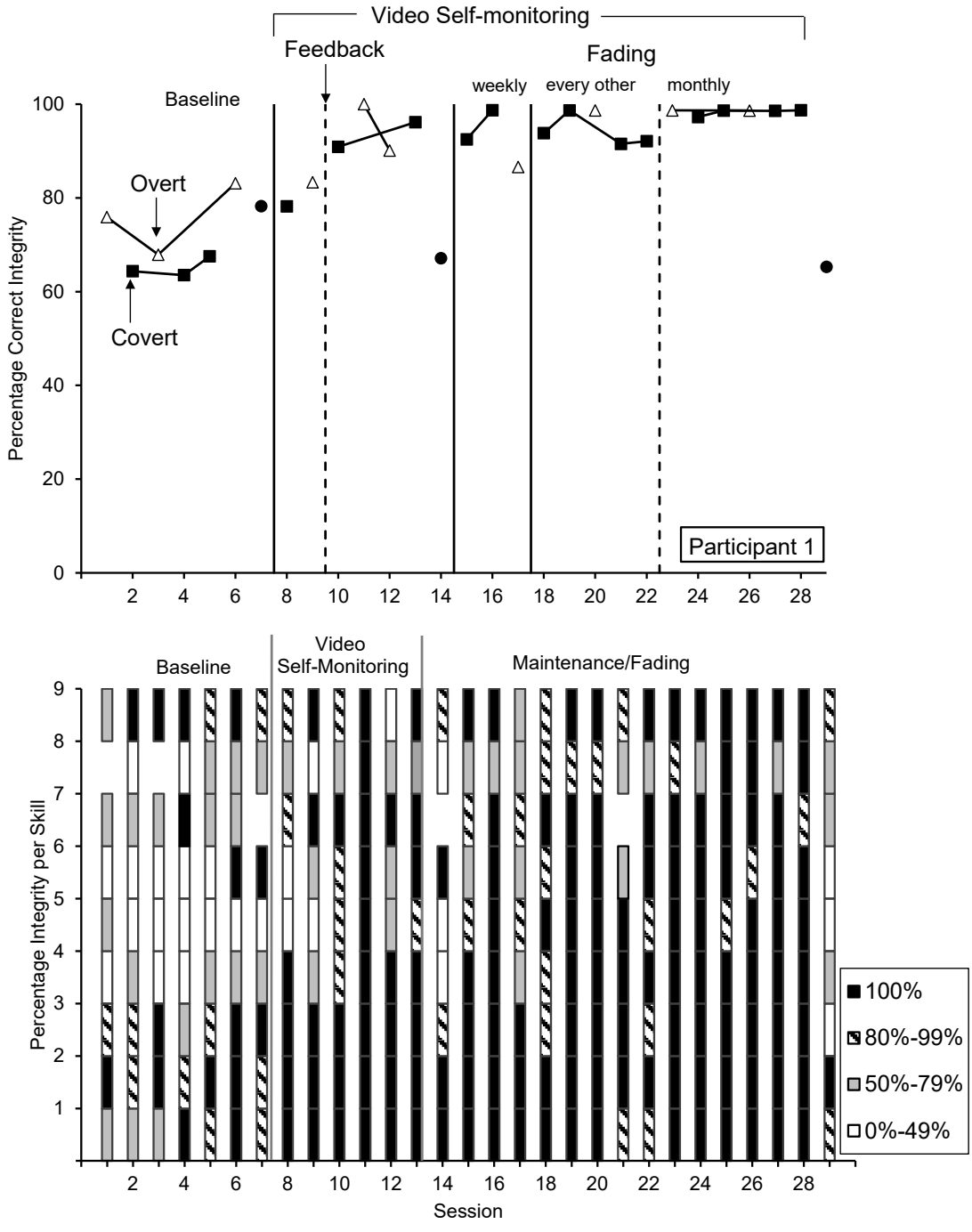
15. How willing would you be to change your routines to carry out this procedure?

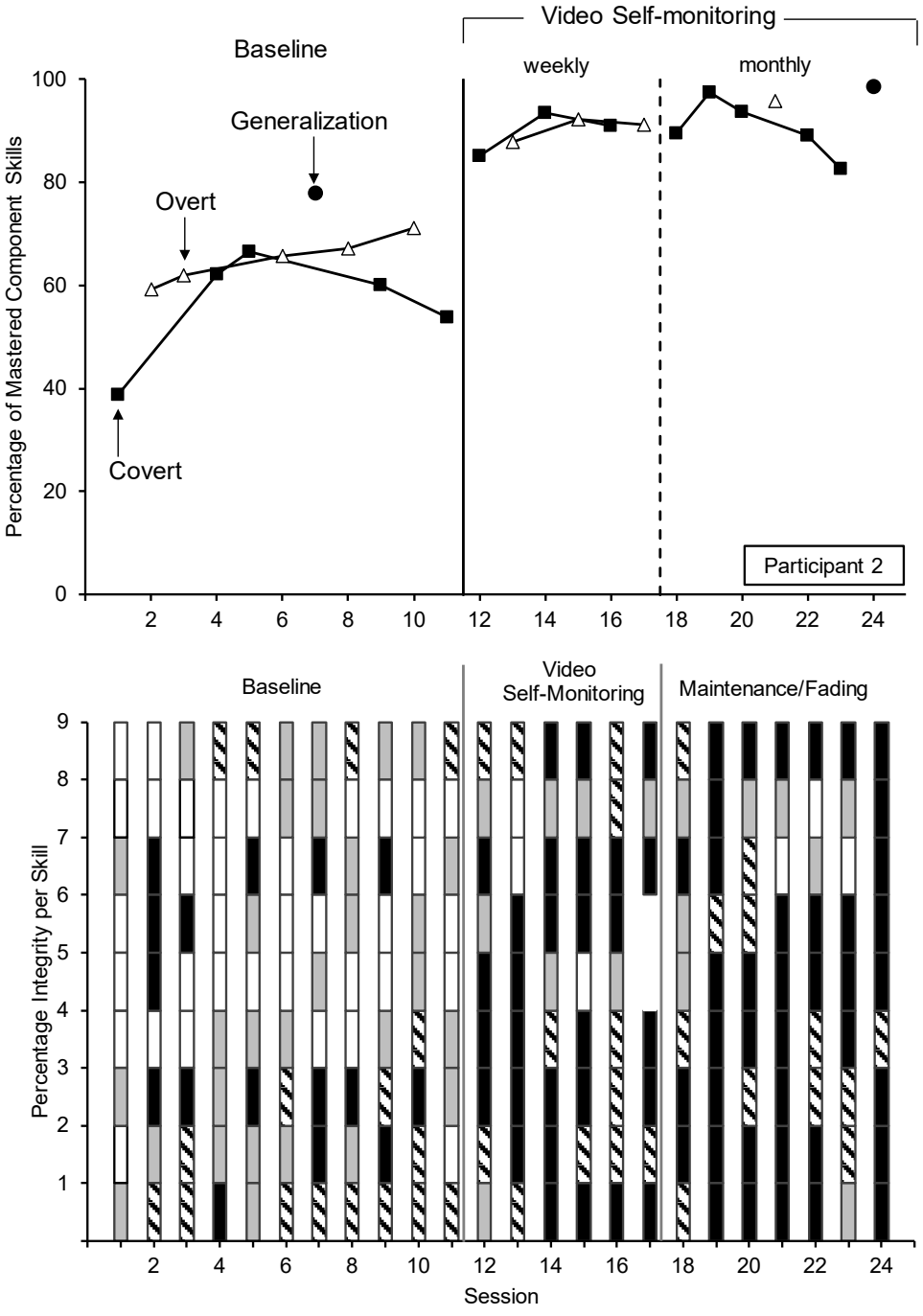
Not at all willing			Neutral			Very willing

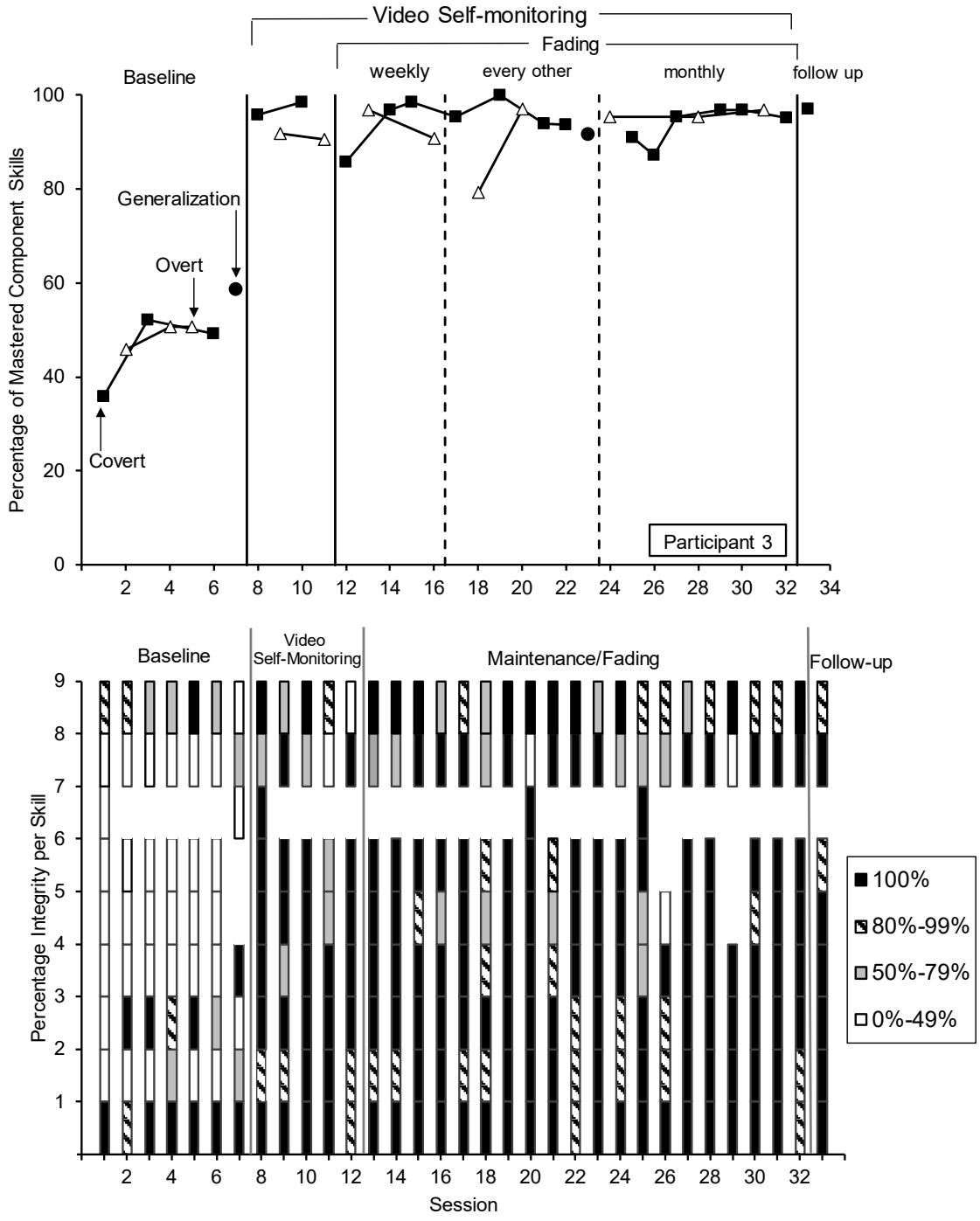
16. How well will carrying out this procedure fit into your existing routine?

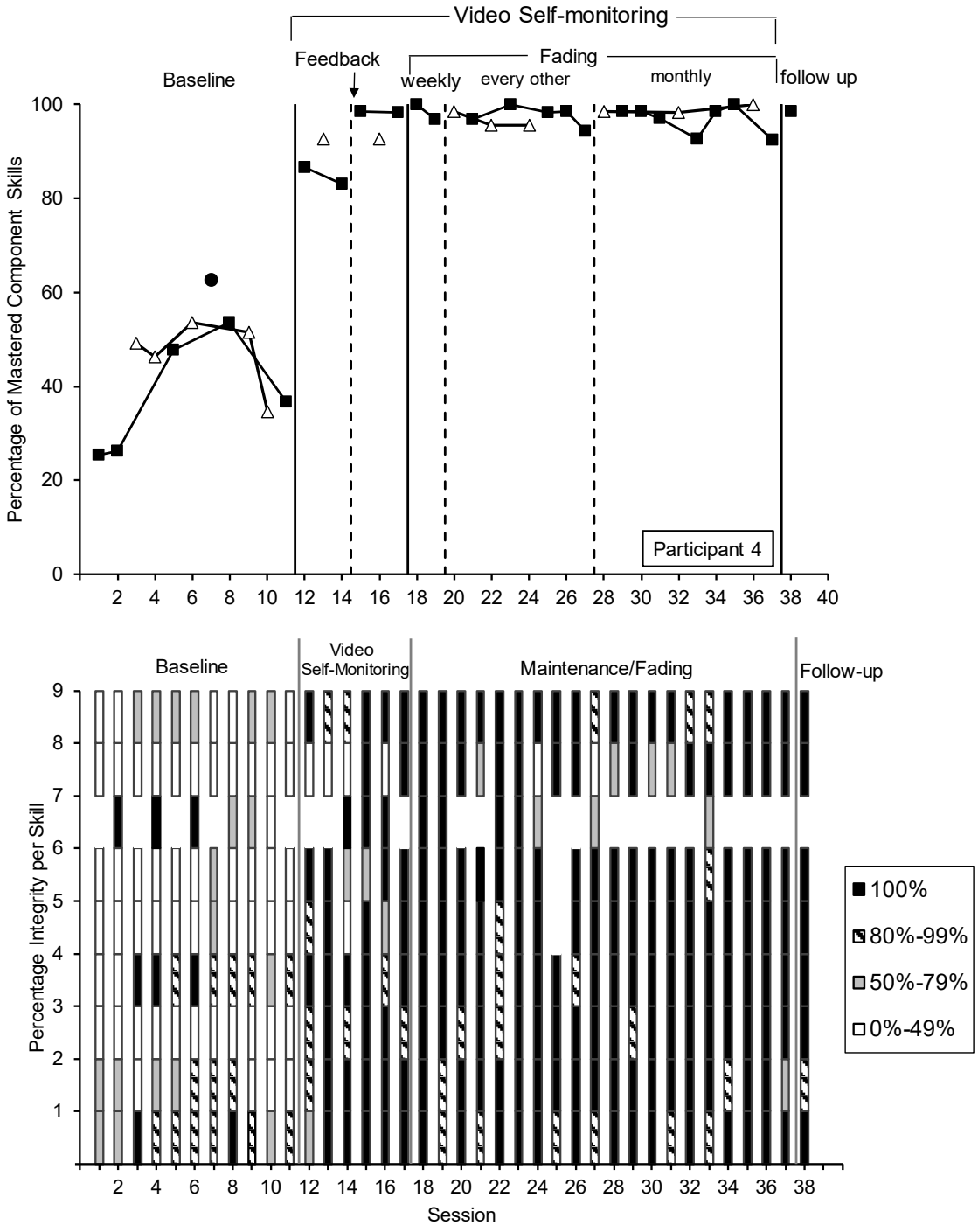
Not at all			Neutral			Very well

Appendix B: Post-hoc Error Analysis




Video Self-monitoring
Baseline
weekly
monthly
Generalization
Overt
Covert
Percentage of Mastered Component Skills
Participant 2
Baseline
Video Self-Monitoring
Maintenance/Fading
Percentage Integrity per Skill
Session
100%
80%-99%
50%-79%
0%-49%


Video Self-monitoring
Fading
Baseline
weekly
every other
monthly
follow up
Generalization
Overt
Covert
Percentage of Mastered Component Skills
Participant 3
Baseline
Video Self-Monitoring
Maintenance/Fading
Follow-up
Percentage Integrity per Skill
Session
100%
80%-99%
50%-79%
0%-49%


Video Self-monitoring
Feedback
Fading
Baseline
weekly
every other
monthly
follow up
Percentage of Mastered Component Skills
Participant 4
Baseline
Video Self-Monitoring
Maintenance/Fading
Follow-up
Percentage Integrity per Skill
100%
80%-99%
50%-79%
0%-49%
Session