Evaluating the Effects of Feedback Type in a Computer-Managed Learning Program

by

Lisa Hunter

A dissertation submitted to the Faculty of Graduate Studies of

The University of Manitoba

In partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Psychology

University of Manitoba

Winnipeg

Abstract

Computer-Aided Personalized System of Instruction (CAPSI) is an online instructional system that can be used for teaching and training individuals in a variety of skills and behaviours. It has been found to be beneficial for teaching both declarative and procedural knowledge. Given the usefulness of CAPSI the following question arises: how can CAPSI be made more effective? The present research evaluated whether different types of feedback differentially affect declarative and procedural learning while training a behavioural assessment procedure called a functional analysis (FA). Each of eight university students were presented with each type of feedback, alternated within each FA-training condition, in a multi-element reversal treatments design. The study evaluated and compared textual feedback of two interventions. Intervention A was Elaborative Knowledge of Results (i.e., an explanation of why the answer was correct) and Intervention B was Simple Knowledge of Results (i.e., "correct"). Results demonstrated benefits of both types of intervention. Declarative knowledge written test results demonstrated that Elaborative Knowledge of Results yielded slightly higher test scores when presented as the first intervention, but not when presented as the second intervention. Procedural knowledge demonstrated no large differences in procedural accuracy of implementing an FA; however, Simple Knowledge of Results feedback procedural accuracy results were slightly higher for the Alone, Control and Demand condition. The small sample size precluded the use of inferential statistics; however, visual analysis revealed individual differences that would likely not be found using statistics.

Key words: CAPSI, online learning, computer-managed learning, computer-aided learning, feedback, procedural and declarative knowledge, functional analysis training, e-learning

Acknowledgements

I would like to express my deepest thanks and gratitude to all the individuals who assisted in making my research successful. My advisor, Dr. Pear for his guidance and support throughout my doctoral work, as well as my committee members, Dr. Yu, Dr. Docker, and Dr. Ramon for their input and recommendations throughout this project. I would like the thank the participants who volunteered their time. Finally, I would like to thank my research assistants who dedicated their time and effort to this project, Maria Pongoski and Karli Pedreira who came to help assist with research sessions, Amy Pankewich for assisting in writing the DTT chapter for this project, and Morena Miljkovic for helping with coding data.

Table of Contents

Evaluating the Effects of Feedback Type in a Computer-Managed Learning Program

Introduction

Computer-Aided Personalized System of Instruction (CAPSI) is an online learning (tool to help individuals learn a variety of materials and skills. CAPSI has been used to teach university course material, and various behavioural assessment procedures. It has also been used as a therapeutic technique (e.g., Simister et al., 2018). The system was developed by Joseph J. Pear and Witold Kinsner (1988) to take advantage of new computer technologies to teach undergraduate students course material. CAPSI is based on Keller's (1968) Personalized System of Instruction (PSI). Over the years modifications have been made to increase the efficiency of the program, to facilitate interactions between students and between the instructor and students, and to implement modern technological advances (Pear et al., 2011). One of the main objectives of creating CAPSI was to increase instructor efficiency in PSI courses, as PSI requires considerable instructor time to set up and implement the system (Pear & Crone-Todd, 1999). Benefits of CAPSI for students include the convenience of accessing the system from any location, including rural areas with internet connection, and ability to proceed at one's own pace. The system also results in a reduction of time and resources from instructors' and administrators' perspective to improve cost effectiveness (Hu et al., 2012; Zaragoza Scherman et al., 2015). These reasons are particularly important given the current emphasis on virtual learning and working from home. This raises the following question: can CAPSI be made more effective to train higher level skills than it currently is? The present study manipulated the feedback component of CAPSI to determine whether some types of feedback are more effective than others; specifically, whether elaborative knowledge of results (EKR) feedback is more effective than simple knowledge of results (SKR) at increasing both declarative and procedural

knowledge. Declarative knowledge refers to the written understanding of material, whereas procedural knowledge refers to the application of behaviours in practice.

CAPSI

CAPSI evaluates a student's mastery of material being taught by evaluating student written responses of randomly assigned instructor questions graded by the instructor or peer reviewer and records the student's progress in learning the material. Pear and Crone-Todd (1999) conducted research on CAPSI that supported the system as an effective teaching method when used to teach four different undergraduate courses. This study was the first to test a new version of CAPSI that could be remotely accessed via computer. Prior to that, the original version of CAPSI involved students and instructors (and teaching assistants, if there were any) accessing CAPSI through connections to a mainframe computer (e.g., Pear & Novak, 1996). For the newer version student's computers required special software to support the program to access the system remotely. Students were also given access to a computer laboratory 24 hours per day, 7 days a week. The CAPSI components of the Pear and Crone-Todd study included: course material to study, study questions based on those readings, and computer-generated unit tests consisting of randomly assigned short essay type questions, and an option to peer review other students unit tests for extra course credit. The course also included two midterms and a final exam. No lectures were given during these courses, and all tests and exams were administered via CAPSI. Answers to unit tests were marked (i.e., graded) either by the instructor, a teaching assistant (TA), or two student peer reviewers. Students would be assigned as peer reviewers for a unit test once they had mastered and passed a computer-generated test on that particular unit. Once a student completed a unit test, CAPSI would assign their answers to two student peer reviewers to mark the unit test; if no students were available to mark the test, the instructor or

TA were assigned to review the questions. Students who were assigned to peer review a unit test were given 24 hours to mark the test; otherwise, they received a small late-penalty mark. Students accrued points for completing peer reviews as well as for passing unit tests. Units were sequential. For students to advance from one unit to the next, they had to demonstrate mastery of the material in the preceding unit. Overall, the results indicated that students received average final marks for all four courses; however, one of the four courses reported lower marks and it was hypothesized this was due to the level of difficulty of the learning material. It therefore appears that CAPSI is at least as efficient as a teaching method when compared to traditional methods, and on average is more preferred by students. The departmental student course evaluation at the end of the course indicated that 53.7% of students rated the CAPSI-taught course as good or very good, and 37% of students rated the course as average. Seventy-one percent of students agreed that peer-reviewing the unit tests aided in their understanding of the material. Similar results from student satisfaction evaluations suggested that CAPSI-taught courses were as good as or better than traditionally taught courses (Pear & Novak, 1996; Svenningsen, Bottomley & Pear, 2018). Additional studies evaluated the ability of PSI to increase higher order thinking and suggests that CAPSI can be used to teach critical thinking (Grant & Spencer, 2003; Reboy & Semb, 1991; Svenningsen & Pear, 2011). These studies demonstrate that CAPSI is effective at increasing declarative knowledge.

Previous research has found CAPSI to be an effective method to teach not only university material but also to conduct behavioral skills training (BST), including discrete trial teaching (DTT) and the Assessment of Basic Learning Abilities Test (ABLA) (Zaragoza Scherman et al., 2015; Hu et al., 2012; Hu & Pear, 2016). These studies suggested CAPSI's effectiveness not only to teach declarative knowledge but to teach procedural knowledge as well. In a comparison

between a teaching method for teaching DTT that contained a CAPSI component and a self-instructional manual with a method that contained only the self-instructional manual, Pedreira and Pear (2015) found that students showed higher preference and were more motivated with the method containing the CAPSI component although these differences did not meet a statistical significance criterion and therefore can only be considered suggestive. In addition, there was a positive correlation between motivation and performance scores in the CAPSI condition suggesting that motivation is a predictor of performance on tests of learning. In their discussion, the authors tentatively concluded:

"…CAPSI tends to increase participants motivation levels and, in addition, motivation is a significant predictor of performance. It therefore seems to follow that CAPSI is potentially effective in increasing performance levels, although this was not tested directly in the present study" (p. 49).

Functional Analysis

The learning material used for the current study to evaluate declarative and procedural knowledge included a functional analysis self-instructional manual (SIM) to learn to conduct an FA. Saltel and Yu (2015) developed the SIM to teach undergraduate students to conduct FAs. The SIM was shown to be effective in teaching undergraduate students to conduct FAs (Saltel, 2016).

FAs are one of the most effective behavioural assessments to evaluate variables maintaining problem behaviours in both adults and children with ASD and developmental disabilities. The importance of understanding the function maintaining behaviour is the necessity for designing interventions to reduce the target problem behaviour based on its function (Baer et al., 1968). However, this study focused on teaching the procedures required to conduct an FA; consequently, participants were not be taught how to interpret results from an FA or how to create function-based interventions.

It is tempting to make assumptions about functions of behaviour based solely on observations. The problem with descriptive observations, however, is they have not been found to be as accurate as FAs in determining functions of behaviour (Thompson & Iwata, 2007). FAs help distinguish consequences that contribute to problem behaviours, such as different types of attention that strengthen the target behaviour (e.g., praise, reprimands) (Iwata, 1994).

Behaviours that have been assessed using an FA include: self-injurious behaviour, aggression, echolalia, elopement (in the sense of "leaving without authorization"), tantrums, and property destruction (Hanely et al., 2003). An FA consists of manipulating the environment and establishing operations to maximize the probability of evoking the problem behaviour to best determine the function maintaining it (Iwata et al., 2000). A standard FA assesses three functions of problem behaviour which are explained below: automatic reinforcement, attention, escape, and a control condition (Iwata et al., 2000).

The purpose of the automatic/alone condition is to evaluate whether the target behaviour occurs in the absence of attention and demands. The automatic/alone condition involves ignoring the individual regardless of the behaviour they emit or the client can be left alone in the room and observed behind a two-way mirror to observe the target behaviour. The client is not provided any materials to engage with. These environmental factors increase the probability that the client will engage in the problem behaviour if the target behaviour is automatically maintained (i.e., sensory stimulation) (Iwata et al., 2000).

The conditions have the following purposes: The purpose of the attention condition is to assess whether rates of behaviour increase when the individual is provided attention by the assessor as a consequence for engaging in the problem behaviour. The attention condition consists of the assessor providing moderately preferred items on the table for the client to engage

with, while the assessor remains in the room and pretends to be busy. The assessor provides attention to the client after each occurrence of the problem behaviour being assessed, by for example saying “don’t do that”. The assessor ignores all other behaviour (Iwata et al., 2000).

Finally, the purpose of the demand condition is to assess whether the removal of demands functions as the maintaining variable of the target behaviour (Iwata et al., 2000). The demand condition involves placing simple demands consistently to the client throughout the condition. For example, simple demands may include imitation (e.g., copying motor movements such as “wave”) receptive and expressive language skills (e.g., asking the client to show or tell you what an item is, such as pointing to an object and asking, “what’s this?”), or completing a brief task such as solving a puzzle. In the event that the client does not respond within four seconds, the assessor prompts the client to complete the task by modelling, gesturing, or hand-over-hand guidance. The assessor then praises the client right away for completing the tasks and issues the next demand after two-seconds. If the client engages in the problem behaviour the assessor removes the stimuli and postpones demands for 30 seconds. At the end of the 30 second interval, the assessor resumes issuing demands and continues the above procedures. All other behaviours are ignored.

The purpose of play/control procedure is to act as a baseline in which potential controlling variables of the target behaviours are absent so the target behaviour can be observed in a more natural setting (Iwata et al., 2000). The procedure involves providing positive attention, access to preferred items, and no demands are placed on the individual. The assessor stays in the room with the client and ignores the occurrence of the problem and other non-target inappropriate behaviours. The client is provided positive social attention on a fixed-time schedule of at least once per 30 seconds; for example, “I love the way you’re playing” or “it is

nice to see you today" (Iwata et al., 2000). A fixed-time schedule is defined as a pre-determined time set after which reinforcement occurs, such as every 5 minutes the child was praised for staying calm. A differential reinforcement of low rates (DRL) may also be added. If the individual engages in the target behaviour during the delivery interval time, i.e., end of the 30 second interval, the assessor delays attention for 5 seconds before re-administering attention on the fixed-time schedule (Iwata et al., 2000). If the client tries to engage with the assessor, the assessor engages back in a positive way.

The order of the conditions of an FA is important as it may influence the occurrence of a behaviour by manipulating establishing operations. An example of this is the alone/automatic condition occurring before the attention condition, thus limiting access to attention, which may increase the problem behaviour. If an FA demonstrates that a child engaged in the target behaviour most frequently in the attention condition, one may conclude the target behaviour is maintained by attention and could therefore provide the basis for a function-based intervention. Typically, each FA condition is 10 to 15-minutes long; however, there are studies that have examined brief FAs of 5 minutes (Beavers & Iwata, 2013; Iwata et al., 2000).

Training individuals to conduct FAs was first studied by Iwata et al. (2000). A typical study involving a multi-component training package to train undergraduate students to conduct FAs involved a package consisting of reading material, videotapes of simulated FAs, a written quiz, and feedback on performance during sessions. The results of the study demonstrated that the training package was successful at teaching undergraduates to conduct FAs with confederates role-playing individuals whose challenging behaviours are being functionally analyzed. Various studies have since replicated the results of this study with the use of "real clients" instead of

confederates and direct care staff as well as student participants (Moore et al., 2002; Philips & Mudford, 2008; Wallace et al., 2004).

FAs are conducted or supervised by trained behaviour analysts. FAs are supported by evidence-based research and have been demonstrated to be the most effective assessments for evaluating the antecedents and consequences maintaining problem behaviour (Beavers et al., 2008; Iwata et al., 2000; Thompson & Iwata, 2007). It is important to train individuals effectively to conduct these procedures to ensure that individuals being supported are getting the best treatments possible. Effective training to conduct FAs is important to minimize the likelihood of potential risk associated with conducting this type of assessment procedure given that the nature of the procedure is intended to assess challenging and potentially harmful behaviours. Since the Iwata et al. (2000) study on training students to conduct FAs there have been many successful training packages that have effectively trained support staff, parents, and supervisors to conduct FAs (Lambert et al., 2014; Phillips & Mudford, 2008; Saltel & Yu, 2015; Stokes & Luiselli, 2008). Two studies in particular demonstrated the importance of feedback: Both Stokes and Luiselli (2008) and Trahan and Wordsell (2011) successfully trained individuals to conduct FAs using a BST package. The results of the both studies demonstrated that all participants reached the mastery criterion after receiving the addition of feedback. This indicated the importance of feedback as a component of BST packages.

Some research has looked at what exactly about the multi-component training package was effective in a behavioural skills training (BST) package. BST packages typically consist of written instructions, modelling, rehearsal, and feedback. Various BST packages have been found effective to teach a variety of behavioural skills including: preference assessments, functional analysis (FA), ABLA, DTT, and picture exchange communication systems (Arnal, 2013; Boris,

2016; Lavie & Sturmey, 2002; Rosales et al., 2009; Wightman, et al., 2012; Ward-Horner & Sturmey, 2012). Ward-Horner and Sturmey (2012) completed a component analysis of a BST of an FA to evaluate the effectiveness of these components separately. FAs are a behavioural assessment of function of behaviour which will be further explained later on. The study trained direct care staff with varying educational backgrounds as participants to conduct FAs. The effectiveness of each of the previously mentioned typical BST package components was assessed using an alternating treatments design (ABC design for one participant and ABCD for two participants). The results demonstrated that feedback was the most effective component of the package overall for all three participants at increasing performance to mastery criterion. This increase in performance was found most consistently with feedback compared to any other component evaluated. Modelling was found to be the second most effective component leading to mastery. These research studies strongly suggest the importance of feedback as a component of increasing procedural knowledge.

Feedback

Feedback was used as a component of Keller's (1968) original PSI, it involved individuals called *proctors* who usually were students in a more advanced course who evaluated students' quizzes to determine whether the student had mastered the material or needed to restudy it. PSI uses a mastery system, which means that students must demonstrate mastery of the material in a unit before proceeding to the next unit. Feedback involves providing either positive or negative information based on a student's performance (Hattie & Timperly, 2007). There is some discussion in the ABA literature about what the function of feedback is, the hypothesis of the present study was that feedback can both be seen as a reinforcer or a punisher (Mangiapanello & Hemmes, 2015). When positive feedback is delivered, it works as a positive

reinforcer if it increases the future occurrence of answers written with similar quality. However, feedback can also be a punisher (i.e., it may suppress the occurrence of incorrect or low-quality answers). Majer, Hansen, and Dick (1971) evaluated whether specific praise words, which were based on the students' preference, would increase learning when teaching algebra to high school students using computer-assisted learning (CAL). They found that the words chosen did not affect student performance. However, when positive feedback was not provided, students showed an increased task completion time in comparison to students who had received positive feedback in the form of praise (Majer et al., 1971). Gallien and Oomen-Early (2008) examined whether personalized versus collective feedback from an instructor teaching a distance education course would affect student satisfaction, academic performance, and connectedness with the instruction. Personalized feedback was described as tailored feedback delivered based on the individual responses of the student and delivered directly to that student; whereas, collective feedback was described as a summary of feedback delivered to the entire class based on the responses of the all students. Based on the results of Hansen (1974), it could be suspected that the forms of feedback would not make a difference. However, Gallien and Oomen-Early (2008) found that students who received personalized feedback rated higher in satisfaction with the course and scored higher on academic performance . Based on student qualitative reports, the quick response from the instructor to answer questions and deliver feedback was reported as one of the most beneficial components of a course. No difference was found in perceived connectedness to the instructor.

Martin, Pear, and Martin (2002a) examined the types of feedback given via peer reviewers. The study examined a class of 33 students from a behaviour modification course taught using CAPSI. In this study the feedback was given through peer review. Two students

that had previously passed a given unit test were assigned to mark the unit tests of a fellow student. If two students had not yet passed the unit test, the instructor was assigned to mark it. Despite assessing different types of textual feedback (i.e., written feedback) to improve acquisition, 55% of students implemented the feedback they were provided. This raises the question how to improve the implementation of textual feedback to improve knowledge acquisition?

In a second study, Martin et al. (2002b) analyzed the accuracy of marking by peer reviewers. The answers were assessed by two independent students, one that previously passed the course with a high grade and the other had considerable experience as a TA. The scoring from the independent assessors were compared with those of the participants in the class that peer reviewed the questions, both false positives and false negatives were reported. The results showed that peer reviewers produced false negatives 66.8% of the time. When the same answers were marked by a second peer reviewer, it reduced the percentage of marking incorrect answers as correct to 46.4%. The rate of false positives were lower, answers that were scored as correct by the peer reviewer, that were actually answered incorrectly occurred 6.7%.

The authors detected two possible reasons for the inaccuracy in detecting incorrect answers as correct. One possible reason is that it is more difficult to detect errors than correct responses. The second possible reason is lack of motivation or reinforcement to deliver accurate feedback. Another study indicated that the peer-review component did not significantly increase skill acquisition over that of students who did not peer review unit tests (de Oliveira et al., 2016). For these reasons, the peer feedback component was not used in the present study.

A systematic review by Jaehnig and Miller (2007) analyzed different types of feedback and categorized them as follows: knowledge of results (KR), knowledge of correct response

(KCR), elaboration feedback, delayed feedback, review feedback. For the purpose of the present study, knowledge of results will be referred to as simple knowledge of results (SKR), which involves feedback that includes whether an answer was correct or incorrect and does not give the correct answer. For example, if the question asked "Name the principle that involves increasing future occurrence of a target behaviour", if the learner incorrectly answers, they will receive feedback saying "You are incorrect". KCR is a type of feedback that presents the correct answer when incorrect; for example, "You are incorrect, the correct answer is reinforcement". EKR feedback, provides the learner with additional information about the answer, usually including an example; for example, "You are incorrect, the answer is reinforcement, reinforcement is defined as …". Delayed feedback, refers to feedback being delivered after a certain amount of time has elapsed or after a certain number of responses has occurred. Finally, review feedback or answer until correct (AUC) allows the learner to continue to attempt answering the question until they answer correctly. This is either done immediately or all incorrectly answered questions are re-presented at the end of instruction. With AUC the learner isn't always informed whether or not their answer is correct, other than the questions being re-presented till it is mastered. Jaehnig and Miller's (2007) review concluded that SKR is the least effective of all forms of feedback, whereas KCR was more effective than both the absence of feedback and SKR. EKR was overall effective; however, a limitation of EKR was that it took more time for the instructor and learner. Delayed feedback was found to be beneficial, although no better than immediate feedback. AUC was slightly more beneficial than no feedback; however, all other types of feedback examined were found to be more effective than AUC (Jaehnig & Miller, 2007).

Roels, van Roosmalen and van Soom (2010) evaluated the effect of feedback to teach Bachelor of Science students course material about genetics using two computer-assisted-

instruction (CAI) packages, one of which feedback was generated by the program (i.e., the feedback was triggered by the program) and in the other feedback was student-generated (i.e., the student had the ability to trigger feedback or not). Two types of feedback were delivered in both packages: SKR and EKR. Students received SKR feedback in the program-assessed package if the student answered correctly and EKR if the answer was incorrect. In the student-generated feedback package students received SKR feedback if they answered correctly or incorrectly; however, the students receiving this package had the option to select "I don't know the answer" for a question and would be given EKR feedback for that specific question. Therefore, the student was only given further explanations or examples of the correct answer after selecting "I don't know the answer". Interestingly, students rarely chose the "I don't know" option, resulting in an imbalance of EKR between groups. The results of the study were that both packages were effective at teaching the material; however, the program-assessed feedback package delivering SKR for correct responses and EKR for incorrect responses was more effective in the overall learning of the material. The authors suggested that it was more beneficial for the program to assess whether elaborative feedback should be delivered over the request of the student.

Additionally, a meta-analysis by van der Kleij, Feskens, and Eggen (2015) was conducted to review different types of feedback and their effects on learning across various categories of feedback. The types of feedback evaluated included simple feedback, KCR, EKR, and the immediate versus delayed feedback. The results of the meta-analysis were that EKR yielded a higher effect size than any other type of feedback included in the analysis.

Based on previous research EKR demonstrated being the most effective at increasing learning in a variety of research and disciplines. Therefore, the following research aimed to

compared the EKR with SKR. SKR was chosen as it is the simplest form of feedback; however, was shown to be less effective in comparison to other types of feedback.

In summary, CAPSI is a well established computer managed program that has been shown to be as effective as a SIM alone. Additionally, FAs are one of the most effective evidence-based assessments to evaluate variables maintaining problem behaviours in both adults and children with developmental disabilities (Hanely et al., 2003). It is therefore important to determine the most effective methods of teaching individuals to administer FAs. The purpose of the research presented here was to evaluate whether different types of feedback (specifically, SKR or EKR) administered textually through the CAPSI program would differentially effective learning of both declarative and procedural knowledge of implementing an FA. The study received ethical approval from the University of Manitoba Psychology/Sociology Research Ethics board.

Methods

Participants

The experiment consisted of eight participants, including one participant who did not complete the follow-up phase. Table 1 depicts a complete list of participant characteristics. Participants were randomly assigned to two groups to alternate the order of which intervention they would receive first. University of Manitoba students were recruited to participate in the research study through posters posted throughout the University of Manitoba Fort Garry campus after obtaining approval of each department building. In addition to an honorarium of $40 for partaking in the study delivered in two equal parts at the beginning of the first two sessions, participants benefited from participating in the study by learning behavioural principles and procedures associated with conducting an FA. The recruitment package included information

about the project to inform students of time requirements and consent information (e.g., study purpose, compensation, risks, and benefits) before deciding whether to participate in the study.

Prior to baseline (i.e., assessment before intervention) participants were given a short questionnaire to measure their previous knowledge of behavioural principles and FAs to control for potential confounding variables, such as prior knowledge of FAs. Only one participant (P05) had reported having learned about FA previously, but had never observed or conducted one herself. All other participants reported they had never conducted, observed, or previously learned about FA. There was no exclusion criteria based on the questionnaire responses.

Materials

Participants were given a pre and post-test questionnaire to complete (see Appendix A, and B) and a revised version of Saltel and Yu's (2015) SIM as the learning material in the study, with the addition of a unit on DTT (see Appendix C). The SIM was shown to be effective in teaching undergraduate students to conduct FAs (Saltel, 2016).

As explained in the Introduction, DTT is a teaching method that consists of rapid consecutive teaching trials that can require knowledge of prompting procedures. A DTT unit was added to the manual for the present study to expand on concepts that are necessary for implementing the demand condition of the FA. It was speculated that knowledge about DTT may strengthen procedural accuracy with implementing the demand conditions which requires the assessor to provide simple instructions quickly and deliver prompts to the client. Prompts such as, gestures, vocal cues, and hand-over-hand guidance were delivered by the assessor to ensure the individual responded correctly to the instructions. The FA SIM was converted into 6 units for CAPSI unit tests (see Table 2).

Participants were provided all necessary materials to conduct an FA including a

datasheet, writing material, leisure materials (e.g., puzzles, bouncy ball, flashcards, and toy cars), and a timer. Participants were also provided with a desktop computer with access to www.capsiresearch.org, and a personal laptop was used to video record all in-vivo sessions. Participants were also given 3 written tests that they completed in paper form.

Setting

All research sessions were conducted in a room designated as the CAPSI laboratory in the Duff Roblin building at the University of Manitoba Fort Garry Campus. All phases were completed in the lab with a research assistant present in the room for the duration of the session. Research has demonstrated that attrition rates decreased by 73% with the addition of supervision during CAPSI studies, instead of participants completing CAPSI unit tests alone (Kehler, 2016). The lab contained two computers for the student to complete the online units and a large table for studying the material and to conduct the in-vivo procedural knowledge sessions.

Research design

A single-subject modified multi-treatment reversal design (ABCABC) (with counterbalancing across participants as explained below) was used: Baseline, Intervention A, Intervention B, Intervention A, Intervention B, Baseline. The participants were randomly divided into two groups: Group A (P03, P04, P06, P07) and Group B (P05, P08, P09, P10). See Table 2 for the counterbalancing order of each group. Participant P1 was a pilot participants and P2 consented to participation; however, did not attend any research sessions.

Procedures

The experiment consisted of four phases: baseline, intervention (Intervention A and Intervention B), post-test, and follow-up. Participants were allowed to work at their own pace;

however, they had to complete the experiment within three months of confirming consent or they would be terminated from the study. Participants were given their honorariums over the first two sessions of the study. Four participants chose to complete the study in two sessions, three participants completed the study over three sessions, and one participant did not complete the follow-up session. Participants took approximately five to six hours to complete the study regardless of the number of sessions it took to complete the study.

The study consisted of three written tests, one completed during each phase (i.e., baseline, post-test, follow-up). Each written test contained 12 questions, which were selected randomly while counterbalancing for an equal number of questions from each unit of the manual and balancing for level of difficulty using Revised Bloom's Taxonomy (RBT) (Anderson et al., 2000). RBT is a rating system to operationalize the level of difficulty of knowledge required to answer written questions, rating 1 as the least difficult, and 6 as the most difficult. The six cognitive process dimensions or levels of Bloom's Revised taxonomy are 1) knowledge (memorization) 2) comprehension, 3) application 4) analyis, 5) synthesis and 6) evaluation. The current study had questions rated from levels 1-4. All participants received the same written tests per phase and each test contained different study questions. Written tests were scored for the percentage of correct responses based on the answer key.

Baseline

Baseline consisted of two steps (see Table 3). Step 1 evaluated baseline declarative knowledge of FA. Participants were instructed to complete the baseline written test.

Step 2 evaluated baseline procedural knowledge with an applied test. The participant was read a short summary about FAs that included minimal information to best assess baseline knowledge of conducting an FA without prior knowledge. Participants were asked to implement

a five-minute FA with a confederate playing the role of an individual with a developmental disability displaying head hitting as the target problem behaviour to be assessed. The confederate followed a script on flash cards per condition and phase that was randomized. Each participant was provided everything they needed to conduct an FA.

Participants were scored on percent accuracy for multiple dependent variables for each unique FA condition based on a procedural checklist from the SIM (see Appendix D). There was no inclusion or exclusion criteria based on baseline scores. In addition, an overall percent mean accuracy score was calculated for each participant per applied FA condition session. Sessions were recorded and coded later by either the principal investigator or a research assistant. The five-minute sessions were coded using a 10-second partial interval recording method. For example, each 10-seconds the video was paused and the coder would score whether the correct dependent variables occurred anytime within that interval (see Appendix E).

Confederates for the FA followed a script containing 20 responses, randomly rotating through cue cards containing three types of behaviours (appropriate behaviour, inappropriate behaviour, and target behaviour). For example, one rotation for the demand condition might be 6 instance of target behaviour, 11 instances of appropriate other behaviour (e.g., trying to get attention from participant – poking arm or trying to make eye contact, playing with hair, clapping), and 3 instances of inappropriate non target behaviour (e.g., banging the table, banging the wall, scratching self, throwing items – items on table or taking off clothing to throw, yell). Each FA condition and phase of the study had a different script to follow.

Intervention

SIM. Intervention consisted of two steps (see Table 3). Step 3 – participants were given the FA SIM (Saltel & Yu, 2015). Participants were given as much time as they required to read

each unit and complete the study questions. Each SIM unit consisted of 7-14 questions (i.e., fill in the blanks, short essay, and multiple choice).

CAPSI. Once the participant completed a unit of the SIM, they were instructed to complete Step 4 – completing the corresponding unit test using CAPSI. The participants had to complete the SIM unit before writing the corresponding unit test. The CAPSI component consisted of a total of six-unit tests. Unit tests were comprised of 4 questions that were randomly selected from the SIM study questions. Each participant was provided a username and password to log into CAPSI through Internet Explorer. Since CAPSI was designed as a system to aid in learning material, participants were able to have access to the SIM during the CAPSI unit tests; however, participants were informed they would not have access to the SIM during the written post-test that was administered following completion of the SIM and unit tests. The participants had 15 minutes per unit test to complete the questions. A shorter amount of time for completing the unit tests was hypothesized to decrease students' referring to their notes during the unit tests. Following completion of a unit test, it was scored and one of two feedback types (i.e., intervention A or B) was delivered by the principal investigator through CAPSI immediately. The principle investigator prompted the participant to review their feedback and whether they passed the unit test or had to restudy the material. If the participant received a restudy, CAPSI and the principal investigator prompted them to restudy the material and re-write the test when they were ready. Due to the unit test questions being randomly selected, a re-test may or may not have consisted of the same questions as the original test. A participant could rewrite a unit test as many times as needed to demonstrate mastery of the material. Furthermore, if a participant mastered the unit material rapidly enough, he or she could complete multiple unit tests in one day.

Feedback. The independent variables of the intervention were feedback type – SKR and EKR. These feedback types were delivered via CAPSI in response to the participants' unit test answers. The principal investigator scored each question providing feedback on the correctness of the answer based on the SIM.

The interventions were alternated for each CAPSI unit test. For instance, if the participant received Intervention A for units 1, 2, and 4, they received Intervention B for units 3, 5 and 6. Thus, the interventions were counterbalanced across participants.

Intervention A consisted of all the above-mentioned steps while receiving SKR feedback delivered via CAPSI for completing unit tests. SKR feedback only involved delivering feedback on whether the participant answered the question correctly or incorrectly. For example, if the questions asked "Functional analysis is a type of ____" and the participant answered "behavioural assessment", the participant would receive a response of "Correct." On the other hand, if the participant answered incorrectly such as responding "intervention" to the same question, the student would be given the feedback of "Incorrect" and would then have to write the unit test again with a new random sample of questions after a period of restudy time. This intervention was alternated with intervention B which was identical to intervention A except participants received EKR feedback instead of SKR feedback. EKR involved feedback on whether the answer was correct or incorrect, plus as an explanation of which aspects of the answer were correct or incorrect, followed by an example of the correct answer. For example, if the same question was asked as in the previous example, "Functional analysis is a type of ___," and the student answered correctly, they would be delivered the feedback "Good job, you are correct. Functional analysis is a type of behavioural assessment. For example, it is used as an assessment of head banging". Consequently, if the participant answered the same question

incorrectly, they might receive feedback that says “Nice try, you are incorrect. Functional analysis is a type of behavioural assessment. It does not include intervention but is used as an assessment to determine functions of behaviour for intervention planning such as for head banging”. The EKR feedback could differ per participant based on which aspects of the participant’s response were correct or incorrect. In both interventions the participant was given a restudy if he or she answered any of the questions incorrectly or mastery if all delivered questions were answered correctly.

Post-Test

To assess for declarative knowledge, the post-test phase consisted of two steps (see Table 3). Step 5 – once the participant completed all unit tests a written test was administered to the participant. The written test was similar to the baseline written test and consisted of 12 different questions that included an equal number of questions from both intervention A and B (i.e., feedback types). Written test questions were coded based on the corresponding intervention to compare scores based on which type of feedback was delivered. For example, the post-test included six questions on which the participant received feedback via intervention A (e.g., SKR) and six questions which the participant received feedback via intervention B (e.g., EKR). Feedback was not provided for post-test answers. The percentage of correct responses from the post-test was compared to baseline scores to grade percentage of improvement from pre-to-post training.

Step 6 – following the written test, procedural knowledge was reassessed by having the participants conduct the applied test, conducting an FA again with a confederate demonstrating the same behaviour as in baseline. The percentage of accuracy from baseline to post-test were compared to evaluate learning across FA conditions and intervention type.

Follow-up

The follow-up phases consisted of two steps (see Table 3). Step 7 – participants completed a final written test in the same format as the previous tests including 12 questions that did not appear in pre- or post-tests. Participants were given as much time as they wanted to review the manual before writing the test.

During Step 8 – to assess procedural knowledge generalization and maintenance implications, participants completed a final applied test. For generalization purposes, the procedure was conducted with a different confederate who engaged in a different target behaviour than during baseline and intervention. Due to participants' availability and scheduling conflicts, not all follow-ups occurred exactly 7 days later (range 6-14 days).

Procedural Integrity and Interobserver Agreement

A second observer scored procedural integrity checks for a third of the research sessions. The observer followed a checklist of steps and a script to record whether the researcher implementing the procedure was following it correctly (see Appendix F). Procedural integrity (PI) was 100% for implementing research sessions. PI for feedback was evaluated over CAPSI for two different components: (1) Evaluating whether the correct type of feedback was delivered (i.e., SKR or EKR) as assigned; and (2) Evaluating whether the correct feedback type was delivered correctly, following the definition provided earlier. A third of CAPSI feedback was evaluated and PI was calculated by percentage of correct responding. The correct feedback type was provided during the CAPSI phase with a score of 100%. The type of feedback was delivered correctly with a score of 90%.

PI of confederates following the script during FAs was calculated for percentage accuracy of engaging in the correct amount and type of behaviours out of 20 (e.g., six

appropriate behaviours, 10 other inappropriate behaviour, and four target behaviours). Procedural integrity calculated for the confederates following the FA script was 90.86%.

To measure reliability of coding for the written questions level of difficulty, a second observer rated the level of difficulty on all three written tests using Revised Blooms Taxonomy (RBT) (Anderson & Bloom, 2001). Reliability of level of difficulty scores was 92%. Additionally, two-thirds of written tests were scored for interobserver agreement (IOA) on percentage of correct responses. The IOA on coding the written tests was 99%.

IOA was collected on the different components (e.g., responding to target behaviour, and responding to other behaviours) of each FA condition, for each condition a third of the scores were coded for IOA purposes. IOA for participants of FA conditions scored the following: Alone condition 95% (range 90 – 100%), Attention condition 93% (range 84 – 97%), Control condition 95% (range 91- 98%), and Demand condition 74% (range 71% - 92%). The demand condition consisted of a larger number of components and more complexity to conduct the session, possibly contributing to the lower level of agreement in comparison to the other conditions.

Social Validity

At completion of the follow-up session, participants were given a self-report questionnaire (see Appendix B). The researcher kept all questionnaires in a folder and were not viewed until the study ended to keep the questionnaires as anonymous as possible. The questionnaire consisted of seven statements rated on a five-point Likert scale (strongly disagree, disagree, neutral, agree, and strongly agree).

Data Analysis

The data for written tests was analyzed descriptively and applied tests were analyzed using visual inspection, a common method used in single subject research (Kahng et al., 2010;

Kazdin, 2011). Data was visually inspected for changes in the levels, trends, and variability in the data across interventions and between participants.

Results

Declarative Knowledge – Written Test

Results of the written tests suggest that the intervention consisting of the SIM, CAPSI unit tests, and feedback were effective at increasing declarative knowledge of FAs. Table 4 illustrates percent accuracy for written test scores for individual participant scores, as well as mean scores per both Groups. Increases in percent accuracy is demonstrated across all participants from baseline (second column) to post-test (third column) and follow-up (sixth column) for Group A (top part of table) and Group B (bottom part of table). All participants except P03 scored slightly higher on the follow-up written test scores (sixth column) compared to post-test scores (third column).

Elaborative vs. Simple Feedback

Group A received EKR as the first type of feedback for units 1, 2, 4 and SKR for units 3, 5, 6. In the top of Table 4, Group A's mean percentage scores were overall higher than baseline (24.72%; second column) regardless of intervention type for both post-test (78.35%; third column) and follow-up (81.39%; sixth column). Specifically, Group A's mean test scores were higher for questions that corresponded with EKR feedback for post-test (82.29%; fourth column), and for follow-up (86.11%; seventh column), compared to mean test scores corresponding with SKR feedback for post-test (74.42%; fifth column) and for follow-up (75.54%; last column). From baseline, Group A's mean change in scores improved for post-test (53.63%; third column) and follow-up (56.67%; sixth column).

The bottom part of Table 4 illustrates results from Group B. Group B received SKR as the first type of feedback for units 1, 2, and 4 and EKR feedback for units 3, 5, 6. Group B's mean percentage scores were overall higher than baseline (18.02%; second column) regardless of intervention type for both post-test (66.10%; third column) and follow-up (74.43; sixth column). Group B's written post-test mean scores for units corresponding to SKR feedback were higher (73.58%; fifth column) compared to units that corresponded to EKR feedback post-test scores (58.63%; fourth column). Written follow-up mean scores corresponding to EKR feedback were higher (78.03%; seventh column) with less variability compared to EKR post-test scores; whereas, mean scores for questions corresponding to units with SKR feedback were similar for both post-test (73.58%; fifth column) and follow up (70.83%; last column). From baseline, Group B's mean change in scores improved for both post-test (48.08%; third column) and at follow-up (56.41%; sixth column).

Procedural Knowledge – Applied FA Tests

Procedural knowledge when administering the FA procedure during baseline, post-test, and follow-up was assessed based on the components of each separate FA condition (i.e., Alone, Attention, Play, and Demand) as well as feedback type. Visual analysis of data for the total percent accuracy scores were calculated for the applied FA test per assessment condition (i.e., alone, attention, play, demand).

Alone Condition

Group A received EKR feedback during the CAPSI component for the Alone condition, while Group B received SKR feedback.

Effects of SKR Feedback during Alone Condition for participants in Group B. In Figure 1 on the left hand side, demonstrates procedural accuracy results of SKR feedback for

participants P05, P08, P09, and P10. P05, P08, and P09 (first three graphs) demonstrated similar patterns of responding during baseline, post-test, and follow-up data. All demonstrated high levels of responding across phases, no variability, and stable responding. Total percent accuracy of implementing the procedure ranged 97%-100%. Data suggests generalization at follow-up with continued high levels of accuracy of 100%.

P10's data (fourth graph) demonstrated low levels of correct responding in baseline and increased to high levels of responding in post-test and a slight decrease at follow-up. The data suggests P10's procedural accuracy improved to conduct the Alone condition with a high level of accuracy post intervention.

Effects of EKR Feedback during Alone Condition for participants in Group A. Figure 1, on the right hand side, demonstrates procedural accuracy results of EKR feedback for participants P03, P04, P06, and P07. P03's (first graph) overall percent accuracy when conducting the Alone Condition demonstrates low levels of correct responding at baseline, followed by an increase of with higher levels of correct responding near 100% at post-test and follow-up. The results suggest that P03's procedural accuracy improved to conduct the Alone Condition with a high level of accuracy post intervention, maintained the skills at follow-up and generalized to conduct the condition with a different confederate exhibiting a different target behaviour.

P04's (second graph) data suggests low moderate levels of accuracy of responding at baseline, followed by a decrease in accuracy at post-test to low levels, with an increase to near baseline levels at follow-up and variability in responding. The results suggest that P04 did not improve procedural accuracy to conduct the Alone condition.

P06's (third graph) data suggests moderate to high levels of accuracy at baseline, with a an increase to high levels of 100% accuracy of conducting the Alone condition post-training. Follow-up data suggests the skills were maintained and generalized to conduct the condition with a different confederate exhibiting a difference target behaviour.

P07 demonstrated 100% accuracy at baseline and at post-test (fourth graph). P07 did not complete the follow-up assessment. Responding appeared stable and would predict a high level of accuracy if a follow-up had occurred.

Attention Condition

Group A received Simple feedback during the CAPSI component for the Attention condition, while Group B received EKR feedback.

Effects of SKR Feedback during the Attention Condition for participants in Group A. In Figure 2, on the left hand side, demonstrates results of procedural accuracy for SKR feedback for participants P03, P04, P06 and P07. All participants demonstrated improvement of procedural accuracy with an increase in correct responses implementing the Attention Condition from baseline to follow-up. The results of participants P03 (first graph) and P06 (third graph) show low levels of correct responding at baseline, followed by an increase to moderate to high levels in post-test and follow- up, with a stable increase in responding. P07 (fourth graph) results were similar with low levels at baseline, and an increase at post-test to high levels of accuracy, there was no follow-up score to analyze. These results suggest that SKR feedback was effective at improving procedural accuracy post intervention for P03, P06 and P07.

The results of P04 (second graph) did not follow the same pattern of responding as the other participants in the group. P04's levels of responding at baseline were low-moderate, then decreased at post-test, with an increase back to low-moderate levels at follow-up. These results

suggest little improvement in procedural accuracy from baseline to post intervention. Suggesting SKR feedback was not effective to increase accuracy of responding.

Effects of EKR Feedback during the Attention Condition for participants in Group B. In Figure 2, on the right hand side, demonstrates procedural accuracy results of EKR feedback for participants P05, P08, P09, P10. Two participants P05 (first graph) and P08 (second graph) showed no difference in procedural accuracy across the phases, presenting stable responding. P05's percent accuracy levels were high across all phases, allowing for less room for improvement. Similarly, P08 had moderate-high levels of responding across all three phases. P08's percent accuracy did not improve post EKR feedback.

P09 (third graph) and P10 (fourth graph) results present similarly, both participants had low levels of responding during baseline, followed by an increase in level of responding to moderate levels during post-test and follow-up for P10, and high levels for P09. Their results showed EKR feedback increased responding for P09 and P10.

Play Condition

Group A received EKR feedback during the CAPSI component for the Play condition, while Group B received SKR feedback.

Effects of SKR Feedback during the Play Condition for participants in Group B. In Figure 3, on the left hand side, demonstrates procedural accuracy results of SKR feedback for participants P05, P08, P09, and P10. Data across Group B for the Play condition presented with all different patterns of responding. P05's (first graph) levels of percent accuracy were high and similar across baseline, post-test, and follow-up. Suggesting SKR feedback did not improve procedural accuracy of responding from baseline to post intervention.

P08's (second graph) levels of responding increased gradually from moderate-to-high from baseline to post-test and follow-up. Due to the small increase in procedural accuracy post intervention, it is unclear whether SKR feedback was effective at improving levels of correct responding.

P09's (third graph) levels of responding were low in baseline, increased to high levels at post-test and decreased to moderate at follow-up. Results demonstrate variability in responding. Suggesting there was an improvement in procedural accuracy post-intervention; however, decreased during the maintenance phase.

P10's (fourth graph) levels of responding were moderate-high during baseline and post-test, with an increase to high levels of 100% accuracy during follow-up. Results suggest an improvement in accuracy at follow-up; however, it is unclear whether the intervention is responsible for this increase since there was no improvements following intervention (i.e., post-test).

Effects of EKR feedback during the Play Condition for participants in Group A. In Figure 3, on the right hand side, demonstrates procedural accuracy results of EKR feedback for participants P03, P04, P06, and P07. Similarly, as Group B, Group A results patterns of responding varied across participants. P03's (first graph) data demonstrates low levels of responding in baseline, with an increase to high levels during post-test, followed by a decrease to moderate levels during follow-up. Demonstrating variability in responding. Suggesting there was an improvement with EKR feedback in procedural accuracy at post-test; however, decreased during the maintenance phase.

P04's (second graph) data demonstrates moderate levels of responding across all three phases, and a slight increase in responding at post-test and follow-up. Due to the small increase

in procedural accuracy from baseline to post-test, it is unclear whether EKR feedback was responsible for improving levels of correct responding.

P06's (third graph) data demonstrates moderate-to-high levels of responding at baseline, with a slight increase to high levels during post-test, decreasing to baseline levels during follow-up. Results suggest the small increase in improvement post intervention was not maintained.

P07's (fourth graph) data presents low levels of responding during baseline with an increase in levels at post-test illustrating high levels of 100% accuracy. No follow-up data was gathered to determine maintenance of procedural accuracy. P07 demonstrated improvement in procedural accuracy post EKR feedback.

Demand Condition

Group A received SKR feedback during the CAPSI component for the Demand condition, while Group B received EKR feedback.

Effects of SKR Feedback during the Demand Condition for participants Group A. In Figure 4, on the left hand side, demonstrates procedural accuracy results of SKR feedback for participants P03, P04, P06, P07. All participants in Group A demonstrated some improvement in procedural accuracy from baseline to post-test. P03's (first graph) data depicts low-moderate levels of responding during baseline with an increase to high levels of responding, followed by a slight decrease in levels of responding during follow-up. Results suggest procedural accuracy improved to high percentage of accuracy post intervention.

P04's (second graph) data depicts low levels of responding during baseline with an increase post-test to moderate levels of responding. Moderate levels of responding were not maintained, demonstrated by a decrease in levels at follow-up. Data presents variable

responding. Results suggested improvement in procedural accuracy occurred post intervention but was not maintained to the same level once intervention was removed.

P06's (third graph) data demonstrates moderate levels of responding during baseline and post-test, followed by an increase to moderate-high levels at follow-up. Data demonstrates performance did not reach high levels of accuracy.

P07's (fourth graph) data demonstrates low levels of responding during baseline followed by an increase to moderate-high levels at post-test. No follow-up data was gathered to determine maintenance of procedural accuracy. Results suggest there was improvement post intervention.

Effects of EKR feedback during the Demand Condition for participants Group B. In Figure 4, on the right hand side, demonstrates procedural accuracy results of EKR feedback for participants P05, P08, P09, and P10. Participants P05 and P09 in Group B demonstrated improvement in procedural accuracy post-intervention. P05's (first graph) data depicts low levels of responding during baseline, followed by an increase to high levels of accuracy at post-test, with a decrease once intervention was removed during follow-up to moderate levels. Results suggest there was an improvement post intervention in procedural accuracy; however, was not maintained at the same post-test

P08's (second graph) data demonstrates moderate levels of responding with a decrease in responding during post-test and follow-up, with levels decreasing to moderate-low. Suggesting no improvement post intervention.

P09's (third graph) data demonstrated an increase in accuracy post intervention. Low-moderate levels of responding are observed at baseline, followed by an increase to high levels at post-test, and a slight decrease to moderate to high levels at follow-up. Suggesting improvement in procedural accuracy post intervention.

P10's (fourth graph) data demonstrates low-moderate levels of responding during all three phases, with a slight decrease in responding at post-test. Results suggest procedural accuracy did not occur post intervention.

Results by Response Type

Data was also inspected based on the type of response required to engage in the procedure accurately. Each condition consisted of different correct responses; these responses were assigned to one of three categories that varied slightly based on the condition of the FA. The purpose was to analyze whether there were differences in improvement of the components of the FA, and if these differences were found across both intervention groups.

Alone Condition

Figure 5 illustrates the three categories of responses for the Alone condition were (1) Response in Absence of Behaviour, (2) Response to Target Behaviour, and (3) Response to Other Behaviours. For all three categories of responses, patterns of responding were similar across all participants with exception to P04's (second graph, right hand side) data shows higher levels of responding for Response to Target Behaviour at follow-up compared to low levels of responding for Response in Absence of Behaviour and Response to Other Behaviour categories.

Attention Condition

Figure 6 illustrates the three categories of responses for the Attention condition were organized the same as the Alone condition. Responding across the three categories was variable for the SKR feedback group (left hand side, graphs 1-4). For Group B who received EKR feedback, both P10 (right hand side, fourth graph) and P08 (right hand side, second graph) showed no increase in responding for Response to Target Behaviour, while P10's data showed an increase in responding from baseline to post-test and follow-up for Response in Absence of

Behaviour and Response to Other Behaviour. P08's data demonstrates high levels of responding for Response to Other Behaviour and Response in Absence of Behaviour across all three phases, and low levels with no increase in responding for Response to Target behaviour.

Play Condition

Figure 7 illustrates the three categories of responses for the Play condition were (1) Attention Provided Every 30 Seconds, (2) Response to Target Behaviour, and (3) Response to Other Behaviours. P05 (left hand side, first graph) P08 (left hand side, second graph) and P10 (left hand side, fourth graph) from the SKR feedback group showed the lowest levels of responding in the Attention Provided Every 30 Seconds category. For the EKR feedback group (right hand side), the highest level of stable responding was to the Response to Other Behaviour, with variability across the lowest level across participants.

Demand condition

Figure 8 illustrates the three categories of responses for the Demand condition were (1) Response to Demand, (2) Response to Target Behaviour, and (3) Response to Other Behaviours. The group that received EKR feedback (right hand side) demonstrated the highest levels of responding and variability for Responding to Target Behaviour. Responding across the three categories was variable for the SKR feedback group (left hand side).

Within Subject Results.

Patterns of responding were analyzed within each participant to assess differences in learning between the different FA conditions. For P03 no patterns of responding were observed for conditions that received the same type of feedback, for example, P03 received EKR for the Alone and Play condition, and SKR for the Attention and Demand condition (see Figure 9; left hand side). P03's levels of responding increased from low to moderate-high from baseline to

post-test across all four conditions. Two conditions – the alone and the attention condition – show levels of responding that continued to increase at follow-up. Whereas, in the play condition shows a large decrease in correct responding during follow-up, and a slight decrease in levels of responding in the demand condition. P03 had the most stable and correct percent accuracy during the alone condition, and the lowest percent accuracy at follow-up during the control condition.

P04's data demonstrates no difference across conditions based on the type of feedback provided (see Figure 9; right hand side). Similar patterns of responding in the alone and attention conditions were observed, with low-moderate levels of responding during baseline, followed by a decrease near zero at post-test and an increase to around baseline levels at follow-up. P04's percent accuracy for the attention, control, and demand was slightly higher at follow-up compared to baseline levels; whereas levels of responding as slightly decreased at follow-up in the alone condition. The greatest increase in responding from baseline to post-test was during the demand condition and the most stable responding with a slight increase during the play condition.

P06 improved from baseline to follow-up for the alone, attention, and demand conditions. The data demonstrate a similar pattern of responding during for the attention and demand condition that received simple feedback which demonstrated an increase with stable responding (see Figure 10; left hand side). The alone and control conditions do not have a similar patterns, the alone condition has stable responding during post-test and follow-up with 100% accuracy, and the control condition levels of responding decreased to near baseline levels during follow-up. P06 had the most stable and correct percent accuracy during the alone condition, with 100% accuracy as well during follow-up during the attention condition. Demand condition has the least amount of improvement.

P07 patterns of responding only have two data points to analyze (see Figure 10; right hand side). The attention, control, and demand conditions demonstrate similar increases in responding. The attention and control similar levels of responding from moderate to high from baseline to post-test, with lower percent accuracy during the demand condition at post-test, and 100% accuracy during baseline and post-test for the alone condition. Demand condition presents the least amount of improvement in learning.

P05 had similar patterns of responding across three conditions, with high percent accuracy during the alone, attention, and play conditions at baseline with less room for improvement (see Figure 11; left hand side). SKR during the play condition did not increase correct responding the highest levels in improvement were found in the demand condition.

P08's patterns of responding are similar across all conditions, with responding similar responding levels (see Figure 11; right hand side). Total percent accuracy did not vary much in the alone and attention condition, with the most improvement during the play condition, and no improvement during the demand condition.

P09 data present similar patterns of responding during the attention and demand conditions which received elaborative feedback (see Figure 12; left hand side). P09 had 100% accuracy during the alone condition across phases and improved the most during the attention condition from baseline to follow-up. P09 maintained percent accuracy the last during the control condition.

P10's patterns of responding were similar between the alone and attention conditions (see Figure 12; right hand). P10 data does not demonstrate learning during the demand condition, and the most improvement during the alone condition.

Social Validity

Results from the social validity questionnaire (see Appendix B) were evaluated based, ranking which components of the training package they preferred, specifically asking which type of feedback they preferred (see Table 7). Both SKR and EKR were rated equally for preference on average (3.83), and participants strongly agreed that what they learned was valuable (4.33). Additionally, rehearsing the procedure was rated on average as the most important by participants (4.50).

Discussion

Declarative Knowledge: Written Test

As summarized in Table 4, all participants in both groups improved on written tests, suggesting that the SIM and CAPSI unit test with feedback were effective at increasing declarative knowledge of FAs. Overall, however, the results did not suggest that one type of feedback was appreciably more effective at increasing learning compared the other. However, suggestive differences were found.

Group A, which received EKR feedback for the initial units, on average scored higher on written tests compared to Group B which received SKR feedback first. The test scores on unit questions that received EKR feedback (82.29, 86.11) were slightly higher compared to the unit questions that received SKR feedback (74.42, 76.54) (see Table 4). It may be that receiving EKR feedback initially provided a model for subsequent responding on how to answer the CAPSI and written test questions, in comparison to Group B, which initially received SKR feedback with no model. Additionally, the total mean scores of Group A for post-test (78.35) and follow-up (81.39) were higher than those of Group B for post-test (66.10) and follow-up (74.43). This may have been due to the order of the feedback or individual differences. Conversely, Group B participant responses on unit questions which EKR feedback was delivered after SKR feedback

were lower than that of responses in which SKR feedback was delivered on unit questions during the post-test; however, questions that corresponded with EKR scores increased during the follow-up test (78.03) compared to post-test (58.63). This suggests EKR feedback still had potential benefits at follow-up for Group B and may have been more effective at maintaining the knowledge as well as presenting less variability among participants scores compared to questions that corresponded with SKR feedback. EKR feedback for Group A also demonstrated less variability across participants scores compared to questions that corresponded with SKR feedback at post-test and follow-up. Interestingly, Attali and van der Kleij (2017) found that EKR feedback was more effective compared to SKR feedback when EKR feedback was administered for the first incorrect response but made no difference if administered after the first correct response. Although this finding is different from that of the present study, it supports the importance of providing EKR feedback. The benefit may be that providing EKR feedback initially over CAPSI may have primed participants on the level of responding required to demonstrate mastery of the material compared to SKR feedback since it included more details about the answers including an example.

Based on the results of the present study, SKR feedback delivered with CAPSI is sufficient to increase learning; however, it may be worth providing EKR feedback initially to prime the participant on the level of detail required to meet accuracy standards for future responding. This conclusion differs from those of Jaehnig and Miller (2007), and van der Kleij, Feskens, and Eggen (2015) who found that EKR feedback was more effective at increasing learning in comparison to SKR feedback. These differences may be due to the material being taught or the research designs of the various studies.

Potential reasons for the results could be attributed to the function of feedback for participants. Although it was hypothesized that feedback functioned as a reinforcer or punisher, feedback as a behavioural learning principle may suggest other operant procedures. Mangiapanello and Hemmes (2015) suggested that other operant procedures may explain the function of feedback such as antecedent controls (e.g., establishing operations and discriminative stimuli). Based on the learning history of participants it is possible that the feedback delivered through CAPSI functioned as one of these other antecedent controls. For example, SKR feedback may have signaled quicker access to completion of the test, whereas EKR feedback signaled a delay in completion. Future research could further examine the function of feedback in applied research.

Procedural Knowledge: Applied FA Test

The results for procedural knowledge demonstrated that overall participants increased procedural accuracy to conduct the different FA conditions, with exception of P04 who did not show improvement on the Alone or Attention condition at post-test.

In the Alone condition all participants but one reached near 100% accuracy of conducting the Alone condition post intervention. The Alone condition of an FA requires the least amount of responses from the participants. The results suggest that SKR feedback was sufficient to increase procedural knowledge to conduct an Alone condition of an FA, while both types of feedback were effective at increasing accurate application of the condition for all but P04 (see Table 5; top two parts of table; second, third, and fourth column). Four of the participants (P05, P08, P09, P07) demonstrated at or near 100% accuracy when conducting the condition at baseline. The overall mean in percent accuracy was higher for the SKR feedback group; however, the baseline performance for three of the four participants was very high leaving little room for improvement.

Overall the mean scores for conditions that corresponded with SKR feedback were slightly higher in comparison to scores that corresponded with EKR feedback for the Alone, Play, and Demand conditions, while the Attention condition that corresponded with EKR feedback yielded higher overall mean scores in comparison to those that received SKR feedback (see Table 6). This suggests that SKR feedback was slightly more effective at increasing procedural knowledge for three of the four conditions in comparison to EKR feedback. It should be noted, though, that feedback was provided only on CAPSI unit tests and no feedback was directly was provided on procedural applied tests. Thus, the feedback variable may or may not be responsible for changes in procedural performance. Previous research that looked at behavioural assessment training provided feedback on procedural performance to train individuals to reach high levels of accuracy of implementing procedures (Saltel, 2016; Stokes & Lusielli ,2008; Trahan & Worsdell, 2011). Future research should compare SKR and EKR feedback provided on application of the procedures to evaluate whether the types of feedback delivered for procedural knowledge impact the effectiveness of learning to conduct an FA.

The patterns of responding among participants during applied tests were most similar for the Alone and Attention conditions regardless of intervention type. This may be due to both conditions having a similar number of components and complexity of responding required to conduct the condition. This finding is similar to other studies that found that participants scored better on the Alone and Attention conditioned and have difficulty mastering the Play and Demand conditions which consist of many components to implementing the procedure (Almenary et al., 2015; Lambert et al., 2014; Pence et al., 2014; Saltel, 2016). Results from Saltel (2016) study demonstrated that of the four conditions that were taught (i.e., Alone/Ignore, Attention, Play/Control, and Demand/Escape), the Play and Demand condition proved to be the

most difficult for students to learn, resulting in lower rates of procedural accuracy when conducted with a confederate. Saltel (2016) suggests this is due to the fact that both of these conditions have multiple components to implementing the procedures compared to Alone and Attention conditions, which follow just one rule.

Contrary to previous studies, participants' percent accuracy in the Play condition of the present study were at a similar level (i.e., moderate-high) and similar to the level in the Attention condition. The percent accuracy of the Demand condition was the lowest. The present study attempted to address this issue with the addition of the DTT unit, although it was not sufficient to increase the percentage of procedural accuracy. It is hypothesized that the number of components of this condition is the reason for the lower accuracy. Based on the post-test questionnaire, participants stated that practicing the procedures was the most helpful. The Demand condition may benefit from extra practice compared to the other conditions.

Limitations

Limitations for the present study included that only three participants in Group A completed all three phases of the study. P07 did not complete the follow-up, which could have either positively or negatively affected the mean scores for Group A across both declarative and procedural knowledge. Group A also had P04 as a participant, whose scores tended to be lower and whose data patterns differed from the other participants (see Figure 1-4) thereby lowering mean scores and increasing the standard deviation in whichever intervention group the participant was in at the time. However, this limitation would have impacted both interventions (i.e., SKR and EKR) equally since the interventions were counterbalanced. The small sample size is another limitation as it precluded the use of inferential statistics; however, visual analysis did reveal individual differences that would likely not be found using statistics. Another potential

limitation was that written test scores were higher on the follow-up test compared to the post-test scores. While the follow-up took place one week after the intervention, it would be expected to produce slightly lower test scores as compared to the post-test. This disparity may be due to using RBT to code questions for level of difficulty. Previous research has found that the reliability of coding using RBT is not strong since questions can fit multiple levels, making it difficult for two independent coders to agree on levels (Crone-Todd & Pear, 2001; Ertmer, Sadaf, & Ertmer, 2011). It is recommended that further studies should randomize which of the three tests participants receive across conditions to increase reliability of written test scores. Additionally, each written test contained 12 study questions – two questions per unit. Adding more questions to baseline and post-test written tests may have provided a better representation of learning the written material and making comparisons between units. Similarly, adding more questions to the CAPSI unit tests would have allowed for more opportunities to provide feedback per unit.

Lastly, the present study used confederates to test the procedural accuracy of implementing the FA. This is a potential limitation for generalization of implementing the procedure with individuals with intellectual and developmental disabilities demonstrating real challenging behaviours. Future research could replicate the present study using "real-clients" to evaluate whether feedback type differentially impacts generalization of implementing the FA.

Future Research

Based on the results of this study, SKR feedback delivered via CAPSI is sufficient to increase both declarative and procedural knowledge; however, it may be worth providing EKR feedback initially to prime the participant on the level of responding required to demonstrate mastery of the material. Future research could include a comparison of SKR and EKR feedback

delivered after the application of FA procedures, rather than only following unit tests, to evaluate whether the types of feedback delivered for procedural knowledge affect learning to conduct an FA. Additionally, it may be beneficial to evaluate types of feedback delivered using other forms of media offered through CAPSI such as providing feedback through video modelling. Video modelling has been shown to be effective as a component of BST, and feedback and modelling have been shown to be the two most useful components (Ward-Horner and Sturmey, 2012). Additionally, in the present study feedback was only provided through CAPSI, and not additionally on written tests. Providing feedback on the post-test could be evaluated to assess benefits to maintaining skills.

Summary and Conclusion

In summary, both SKR and EKR feedback had benefits to increasing learning to conduct different FA conditions. When presented first for CAPSI unit tests, EKR feedback yielded higher overall written test scores in comparison to the scores of those who received SKR initially. However, SKR feedback may be sufficient feedback to increase learning for computer-aided instruction. The benefits of SKR feedback are that it requires less time than EKR feedback from instructors to provide, and participants did not have a preference between either type of feedback. Groups that received SKR feedback performed FA conditions at a similar or higher level of accuracy in comparison to EKR feedback, suggesting both type of feedback have their benefits.

Online learning systems are becoming increasingly important across educational settings due to the COVID-19 pandemic. Moreover, online systems reduce barriers to education in both rural and urban communities that include limited providers, transportation, mobility, and time (Perle et al., 2014). This study contributes to research related to e-learning, in the effort to

provide the most effective online training system to healthcare providers who otherwise would not have access to this type of training. CAPSI is a well-established online learning tool that can continue to incorporate evidence-based procedures into an already effective system. The feedback component should continue to be researched to maximize the potential of the system.

Feedback is a powerful aspect of our everyday lives; we receive it in the workplace, in education, and in other everyday situations. This makes feedback a topic that is important to explore and continue to learn more about. Continuing to learn about which types of feedback are most effective will help shape the way materials are taught through e-learning and with concomitant savings in time and money. Most importantly, the present study adds to research on improving the educational system to maximize students' learning of declarative knowledge and procedural skills.

References

Almenary, F. M., Wallace, M., Symon, J. B. G., & Barry, L. M. (2015). Using international videoconferencing to provide staff training on functional behavioral assessment. *Behavioral Interventions, 30,* 73-86.

Anderson, Lorin W., and Benjamin Samuel Bloom. *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman, 2001.

Arnal, L. (2013). Teaching individuals to conduct a preference assessment procedure using computer- aided personalized system of instruction. (Doctoral dissertation). Retrieved from MSpace. (http://hdl.handle.net/1993/22083)

Attali, Y., & van der Kleij, F. (2017). Effects of feedback elaboration and feedback timing during computer-based practice in mathematics problem solving. *Computers & Education*, *110*, 154-169.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis1. *Journal of Applied Behavior Analysis*, 1(1), 91–97. doi:10.1901/jaba.1968.1–91

Beavers, G. A., Iwata, B. A., & Lerman, D. C. (2013). Thirty years of research on the functional analysis of problem behavior. *Journal of Applied Behavior Analysis*, *46*(1), 1-21.

Bloom, B. S. (1956). Taxonomy of educational objectives. Vol. 1: Cognitive domain. *New York: McKay*, 20-24.

Boris, A. L. (2016) Training ABA service providers to conduct the assessment of basic learning abilities – revised using a self-instructional manual and video modeling. (doctoral dissertation). Retrieved from MSpace (http://hdl.handle.net/1993/31643)

Boris, A., Thomson, K., Murphy, C., Zaragoza Scherman, A., Dodson, L., Martin, G., Fazzio,

D., Yu, C.T. (2011). An evaluation of a self-instructional manual for conducting discrete-trials teaching with children with autism. This article was accepted by *Developmental Disabilities Bulletin*, but the journal ceased to exist (in 2015) before the article was published. (Copies can be obtained from Garry Martin. Email: garry. martin@umanitoba.ca)

Crone-Todd, D. E., & Pear, J. J. (2001). Application of Bloom's taxonomy to PSI. *The Behavior Analyst Today*, *2*(3), 204.

de Oliveira, M., Jaksic, H., Lee. M., Wightman, J., Wang, C., Pedreira, K., Martin, T., Yu, C.T., Vause T., Feldman, M., Pear, J. (2016). The effects of student peer review on efficacy of computer-aided system of instruction to teach discrete trial teaching. *Journal on Developmental Disabilities, 22*(2), 80-90.

Ertmer, P. A., Sadaf, A., & Ertmer, D. J. (2011). Student-content interactions in online courses: The role of question prompts in facilitating higher-level engagement with course content. *Journal of Computing in Higher Education*, *23*(2-3), 157.

Fazzio, D., & Martin, G. (2012). *Discrete Trials Teaching with Children with Autism: A Self-Instructional Manual.*

Gallien, T., & Oomen-Early, J. (2008). Personalized versus collective instructor feedback in the online courseroom: Does type of feedback affect student satisfaction, academic performance and perceived connectedness with the instructor? *International Journal on ELearning, 7*(3), 463-476.

Grant, L. K., & Spencer, R. E. (2003). The personalized system of instruction: Review and applications to distance education. *The International Review of Research in Open and Distributed Learning*, *4*(2).

Hanley, P., Iwata, B., McCord, B. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis, 36,* 146-185.

Hansen, J. B. (1974). Effects of feedback, learner control, and cognitive abilities on state anxiety and performance in a computer-assisted instruction task. *Journal of Educational Psychology*, *66*(2), 247-254

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, *77*(1), 81-112.

Hu, L., & Pear, J. J. (2016). Effects of a self-instructional manual, Computer-Aided Personalized System of Instruction, and demonstration videos on declarative and procedural knowledge acquisition of the Assessment of Basic Learning Abilities. *Journal of Developmental Disabilities*, *22*, 64-79.

Hu, L., Pear, J. J., & Yu, C.T. (2012). Teaching university students to implement the Assessment of Basic Learning Abilities using Computer-Aided Personalized System of Instruction. *Journal of Developmental Disabilities*, *18*, 12–19.

instruction. In R. Zheng (Ed.), Digital technologies and instructional design for

Iwata, B. A., Wallace, M. D., Kahng, S., Lindberg, J. S., Roscoe, E. M., Conners, J., Hanely, G. P., Thompson, R. H., Worsdell, A. S. (2000). Skill acquisition in the implementation of functional analysis methodology. *Journal of Applied Behavior Analysis*, *33*(2), 181-194.

Jaehnig, W., Miller, M. (2007). Feedback types in programmed instruction: A systematic review. *The Psychological Record, 57*, 219-232.

Kahng, S. W., Chung, K. M., Gutshall, K., Pitts, S. C., Kao, J., & Girolami, K. (2010). Consistent visual analyses of intrasubject data. *Journal of Applied Behavior Analysis*, *43*(1), 35-45.

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

Kehler, K. (2016). The effects of supervision on reducing attrition in research participants. University of Manitoba, Unpublished manuscript.

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

Lavie, T., & Sturmey, P. (2002). Training staff to conduct a paired-stimulus preference assessment. *Journal of Applied Behavior Analysis, 35,* 209-211.

Majer, K., Hansen, D., & Dick, W. (1971). Note on effects of individualized verbal feedback on computer-assisted learning. *Psychological reports*, *28*(1), 217-218.

Mangiapanello, K. A., & Hemmes, N. S. (2015). An analysis of feedback from a behavior analytic perspective. *The Behavior Analyst*, *38*(1), 51-75.

Martin, T. L., Pear, J. J., & Martin, G. L. (2002a). Analysis of proctor marking accuracy in a computer-aided personalized system of instruction course. *Journal of Applied Behavior Analysis*, *35*(3), 309-312.

Martin, T. L., Pear, J. J., & Martin, G. L. (2002b). Feedback and its effectiveness in a computer-aided personalized system of instruction course. *Journal of Applied Behavior Analysis*, *35*(4), 427-430.

Moore, J. W., & Fisher, W. W. (2007). The effects of videotape modelling on staff acquisition of functional analysis methodology. *Journal of Applied Behavior Analysis*, *40*(1), 197-202.

Moore, J. W., Edwards, R. P., Sterling-Turner, H. E., Riley, J., DuBard, M., & McGeorge, A. (2002). Teacher acquisition of functional analysis methodology. *Journal of Applied Behavior Analysis, 35,* 73–77.

Pear, J. J., & Crone-Todd, D. E. (1999). Personalized System of Instruction in cyberspace.

Journal of Applied Behavior Analysis, 32, 205-209.

Pear, J. J., & Kinsner, W. (1988). Computer-Aided Personalized System of Instruction: An effective and economical method for short- and long-distance education. *Machine Mediated Learning, 2,* 213-237.

Pear, J. J., Schnerch, G. J., Silva, K. M., Svenningsen, L. and Lambert, J. (2011). Web-based computer-aided personalized system of instruction. *New Directions for Teaching and Learning*, *2011*, 85–94.

Pedreira, K., & Pear, J. (2015). Motivation and attitude in a computer-aided personalized system of instruction course on discrete-trials teaching. *Journal on Developmental Disabilities*, *21*(1), 45-51.

Pence, S. T., St. Peter, C. C., & Giles, A. F. (2014). Teacher acquisition of funtional analysis methods using pyramidal training. *Journal of Behavioral Education, 23,* 132-149.

Perle, J. G., Burt, J., & Higgins, W. J. (2014). Psychologist and physician interest in telehealth training and referral for mental health services: an exploratory study. *Journal of Technology in Human Services*, *32*(3), 158-185.

personalized learning. (pp. 164-190) Hershey, PA: IGI Global. doi:10.4018/978-1-

Phillips, K. J., & Mudford, O. C. (2008). Functional analysis skills training for residential caregivers. *Behavioral Interventions*, *23*(1), 1-12.

Reboy, L. M., & Semb, G. B. (1991). PSI and critical thinking: Compatibility or irreconcilable differences?. *Teaching of Psychology*, *18*(4), 212-215.

Roels, P., Van Roosmalen, G., & Van Soom, C. (2010). Adaptive feedback and student behaviour in computer-assisted instruction. *Medical education*, *44*(12), 1185-1193.

Saltel, L. (2016). Evaluation of a Self-Instructional Manual to Teach Functional Analysis to

Assess Problem Behaviours for Persons with Developmental Disabilities (unpublished dissertation). University of Manitoba, Winnipeg.

Saltel, L., & Yu, C.T. (2012). Functional Analysis Self-Instructional Manual. University of Manitoba, Winnipeg.

Simister, H. D., Tkachuk, G. A., Shay, B. L., Vincent, N., Pear, J. J., & Skrabek, R. Q. (2018). Randomized controlled trial of online acceptance and commitment therapy for fibromyalgia. *The Journal of Pain*, 19(7), 741-753. https://doi.org/10.1016/j.jpain.2018.02.004

Stokes, J. V., & Luiselli, J. K. (2008). In-home parent training of functional analysis skills. *International Journal of Behavioral Consultation and Therapy*, *4*(3), 259-263.

Svenningsen, L., & Pear, J. J. (2011). Effects of computer-aided personalized system of instruction in developing knowledge and critical thinking in blended learning courses. *The Behavior Analyst Today*, *12*(1), 33–39.

Svenningsen, L., Bottomley, S., & Pear, J. J. (2018). Personalized learning and online

Thompson, R., Iwata, B. (2007). A comparison of outcomes from descriptive and functional analysis of problem behavior. *Journal of Applied Behaviour Analysis*, *40*, 333-338.

Trahan, M. A., & Worsdell, A. S. (2011). Effectiveness of an instructional DVD in training college students to implement functional analyses. *Behavioral Interventions*, *26*(2), 85-102.

Van der Kleij, F. M., Feskens, R. C., & Eggen, T. J. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes: A meta-analysis. *Review of educational research*, *85*(4), 475-511.

Wallace, M. D., Doney, J. K., Mintz-Resudek, C. M., & Tarbox, R. S. (2004). Training educators

to implement functional analysis. *Journal of Applied Behavior Analysis,* 37, 89–92.

Ward-Horner, J., & Sturmey, P. (2012). Component analysis of behavior skills training in functional analysis. *Behavioral Interventions, 27*(2), 75-92.

Wightman, J., Boris, A., Thomson, K., Martin, G. L., Fazzio, D., & Yu, C. T. (2012). Evaluation of a self-instructional package for teaching tutors to conduct discrete-trials teaching with children with autism. *Journal on Developmental Disabilities*, *18*(3), 33-44.

Zaragoza Scherman, A., Thomson, K., Boris, A., Dodson, L., Pear, J. J., & Martin, G. (2015). Online training of discrete-trial teaching for educating children with autism spectrum disorders: A preliminary study. *Journal on Developmental Disabilities*, *21*(1), 22-32.

Table 1.

Participant Characteristics obtained from project information and consent package

Group A					
Participant	Age	Sex	Languages	Level of Education	Major
P03	18-24	M	English	High School	Microbiology
P04	18-24	F	English	High School	Film
P06	18-24	M	English, Amharic	Undergraduate	Psychology
P07	18-24	F	English	Secondary Education	Computer Science
Group B					
Participant	Age	Sex	Languages	Level of Education	Major
P05	18-24	F	English, Vietnamese	High School	Psychology
P08	18-24	F	English	University	Political Science
P09	18-24	M	English	High School	Nursing
P10	30-34	F	English, Hausa	Undergraduate	Bachelor of Social Work

Table 2.

Functional Analysis Self-Instructional Manual CAPSI Unit Tests and Intervention Type

Unit #	Topic	Group A: Feedback Type	Group B: Feedback Type
Unit 1	Introduction and Functional Analysis Overview	Elaborative	Simple
Unit 2	Alone Condition	Elaborative	Simple
Unit 3	Attention Condition	Simple	Elaborative
Unit 4	Play Condition	Elaborative	Simple
Unit 5	Discrete Trial Training	Simple	Elaborative
Unit 6	Demand Condition	Simple	Elaborative

Table 3.

Procedure by Phase

Phase	Step	Task
Baseline	1	Declarative Knowledge: Written Test
	2	Procedural Knowledge: Applied FA Test
Intervention	3	Study Self-Instructional Manual (SIM) unit
	4	a. Complete CAPSI unit test until mastery (100%) b. Feedback type delivered through CAPSI following unit test c. Repeat Steps 3 and 4 until all SIM and CAPSI unit tests (6 units) are complete
Post-Intervention	5	Declarative Knowledge: Written Test
	6	Procedural Knowledge: Applied FA Test
Follow-Up	7	Declarative Test: Written
	8	Procedural Test: Applied FA

Table 4.

Declarative Knowledge Written Test Scores - Elaborative Feedback vs. Simple Feedback

Group A

Participant	Baseline	Post-Test	Post-test % correct EKR	Post-test % correct SKR	Follow-Up	Follow-Up % correct EKR	Follow-Up % correct SKR
P03	18.00	86.08	91.67	80.50	85.83	83.33	88.33
P04	8.33	47.50	62.50	32.50	60.00	75.00	45.00
P06	29.17	95.83	91.67	100	98.33	100	96.67
P07	41.58	84.00	83.33	84.67	-	-	-
Mean Scores(*SD*)	24.72(12.42)	78.35(18.36)	82.29(11.92)	74.42(27.27)	81.39(15.96)	86.11(10.39)	75.54(22.65)
Mean Change from Baseline		53.63			57.67		

Group B

Participant	Baseline	Post-Test	Post-test % correct EKR	Post-test % correct SKR	Follow-Up	Follow-Up % correct EKR	Follow-Up % correct SKR
P05	19.42	64.58	62.50	66.67	69.38	80.43	58.33
P08	19.33	81.83	69.33	94.33	85.00	78.33	91.67
P09	29.17	84.67	77.67	91.67	92.50	93.33	91.67
P10	4.17	33.33	25.00	41.67	50.83	60.00	41.67

Mean Scores(*SD*)	18.02(8.94)	66.10(20.42)	58.63(20.14)	73.58(21.35)	74.43(15.97)	78.03(11.89)	70.83(21.65)
Mean Change from Baseline		48.08			56.41		

Note. Scores in percent (%) accuracy. Columns represent overall mean percent accuracy per test phase and percent accuracy of scores of questions that corresponded with either receiving EKR or SKR feedback specifically.

Table 5.

Procedural Knowledge of Implementing Functional Analysis – Total Percent Accuracy

Alone Condition Total Percent Accuracy			
Feedback: Simple			
Participant	Baseline	Post-Test	Follow-Up
P05	100	100	100
P08	97	100	100
P09	100	98	100
P10	34	100	96
M(SD)	82.75(28.17)	99.50(0.87)	99.00(1.73)
Feedback: Elaborative			
Participant	Baseline	Post-Test	Follow-Up
P03	3	93	99
P04	36	5	28
P06	76	100	100
P07	100	100	-
M(SD)	54(37.16)	75(40.23)	76(33.71)
Attention Condition Total Percent Accuracy			
Feedback: Simple			
Participant	Baseline	Post-Test	Follow-Up
P03	36	67	93
P04	22	3	28
P06	49	74	97
P07	49	100	-
M(SD)	39(11)	61(36)	73(32)
Feedback: Elaborative			
Participant	Baseline	Post-Test	Follow-Up
P05	93	100	95

P08	64	67	67
P09	25	100	95
P10	3	67	64
M(SD)	46(35)	84(17)	80(15)
Play Condition Total Percent Accuracy			
Feedback: Simple			
Participant	Baseline	Post-Test	Follow-Up
P05	87	89	87
P08	60	68	77
P09	30	91	50
P10	63	63	100
M(SD)	60(20)	78(12)	79(18)
Feedback: Elaborative			
Participant	Baseline	Post-Test	Follow-Up
P03	31	93	43
P04	46	51	57
P06	79	90	80
P07	37	100	-
M(SD)	48(19)	84(19)	60(60)
Demand Condition Total Percent Accuracy			
Feedback: Simple			
Participant	Baseline	Post-Test	Follow-Up
P03	39	90	75
P04	12	53	21
P06	47	53	67
P07	27	62	-
M(SD)	31(13)	65(15)	54(24)
Feedback: Elaborative			
Participant	Baseline	Post-Test	Follow-Up

P05	22	90	66
P08	50	41	34
P09	41	80	65
P10	34	25	29
M(SD)	37(10)	59(27)	49(17)

Table 6.

Elaborative Feedback Versus Simple – All Functional Analysis Condition

	M(SD)			
Feedback	Alone	Attention	Play	Demand
Simple	99(1.39)	66(34.48)	79(14.43)	60(19.99)
Elaborative	75(37.58)	82(15.72)	73(21.06)	54(23.11)

Note. Procedural accuracy total mean scores (post-test and follow-up scores combined) based solely on feedback type.

Table 7.

Participant Post-Test Social Validity Questionnaire

Statement	Post-Training – Mean (range)
1. I found the manual helpful	4.17 (1-5)
2. I found the online tests helpful	3.83 (1-5)
3. I preferred simple feedback (e.g., correct/incorrect)	3.83 (2-5)
4. I preferred elaborative feedback (e.g., detailed)	3.83 (2-5)
5. I found rehearsing the procedure important	4.50 (4-5)
6. I would have liked to see video demonstrations	4.17 (2-5)
7. I found what I learned valuable	4.33 (2-5)

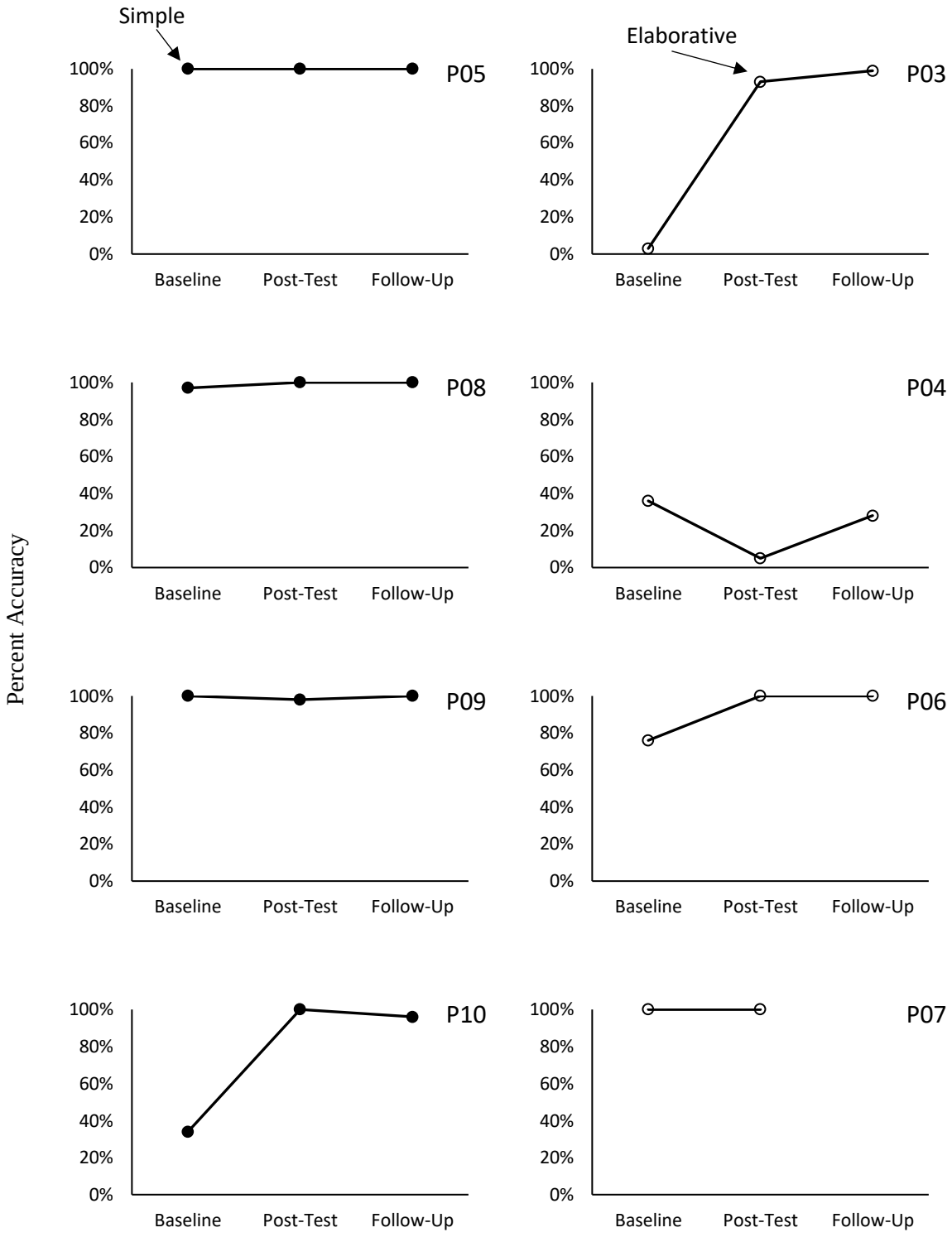

Figure 1. Procedural Knowledge- All response types: Percent Accuracy for the Alone Condition: Simple vs. Elaborative Feedback for baseline, post-test, and follow-up.

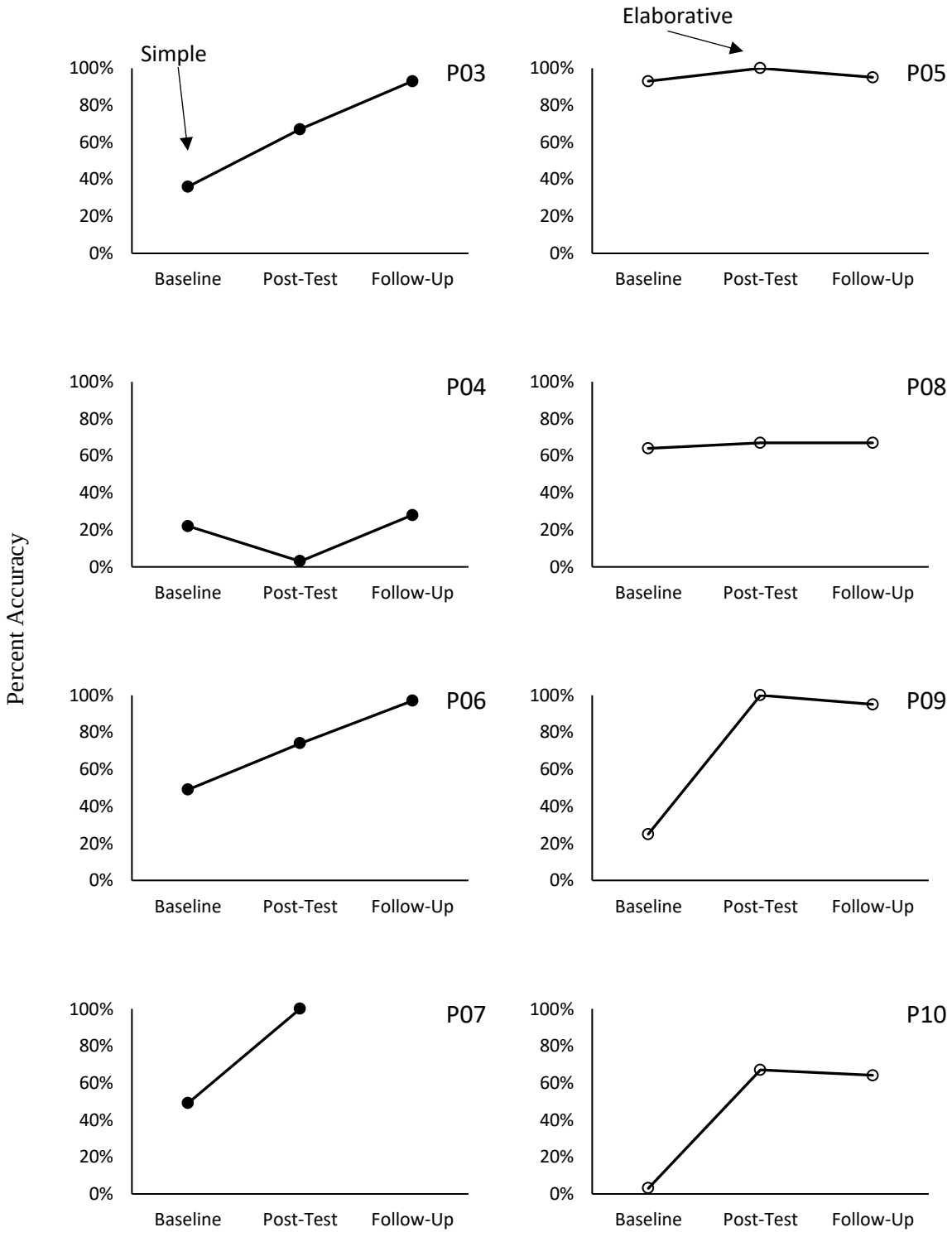

Figure 2. Procedural Knowledge- All response types: Percent Accuracy for the Attention Condition: Simple vs. Elaborative Feedback for baseline, post-test, and follow-up.

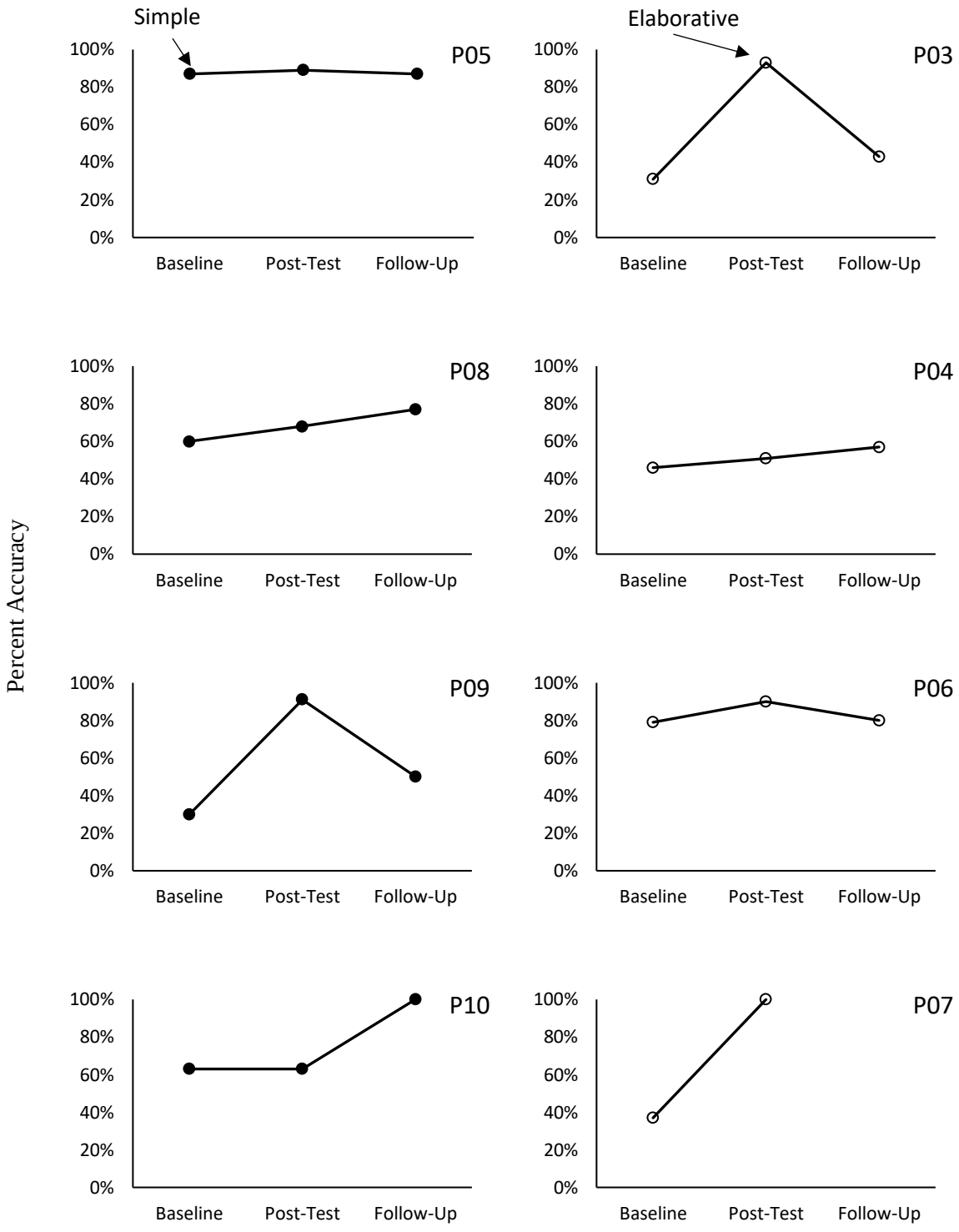

Figure 3. Procedural Knowledge - All response types: Percent Accuracy for the Play Condition: Simple vs. Elaborative Feedback for baseline, post-test, and follow-up.

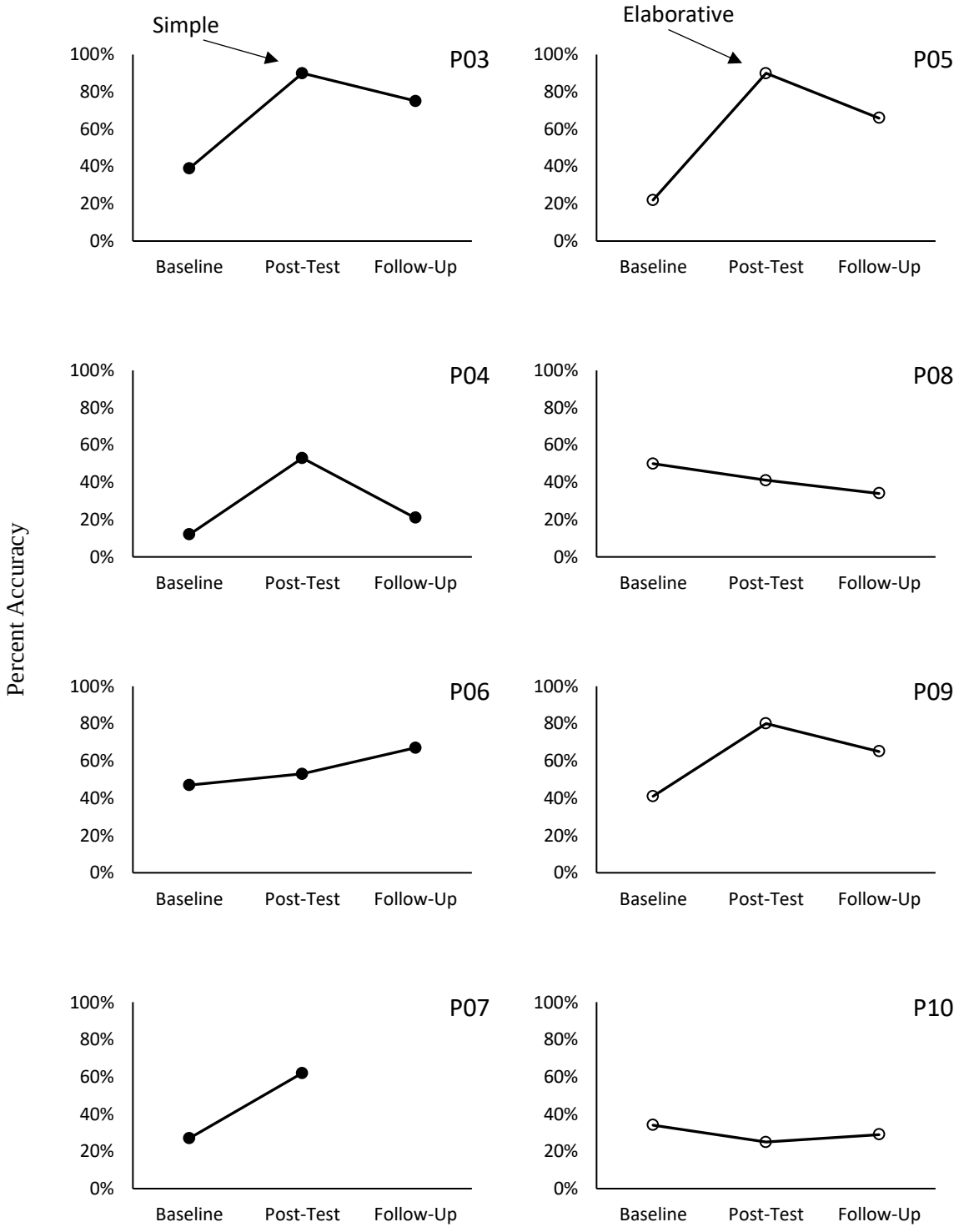

Figure 4. Procedural Knowledge - All response types: Percent Accuracy for the Demand Condition: Simple vs. Elaborative Feedback for baseline, post-test, and follow-up.

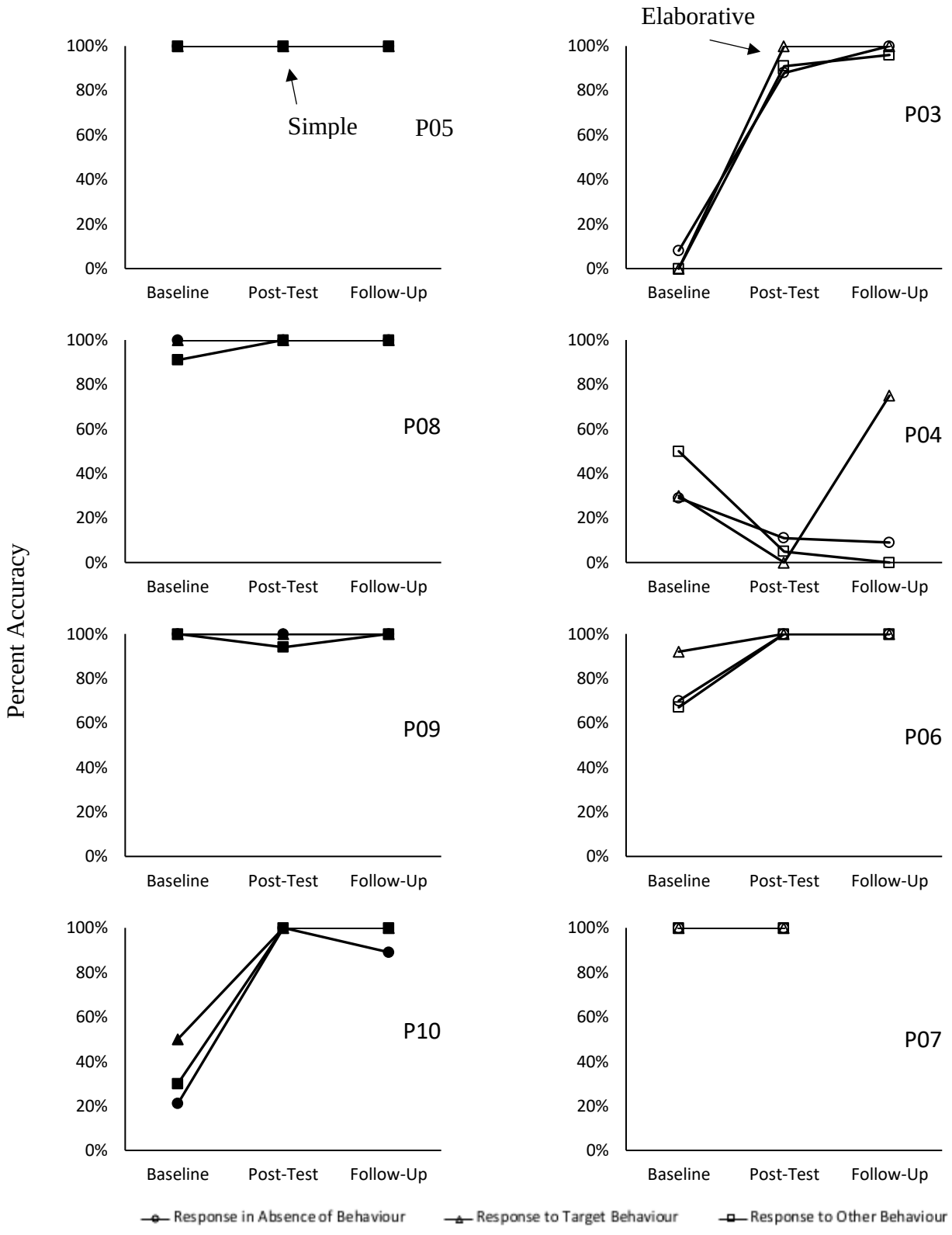

Figure 5 . Percent Accuracy per Response Type in the Alone Condition: Simple Feedback vs. Elaborative Feedback.

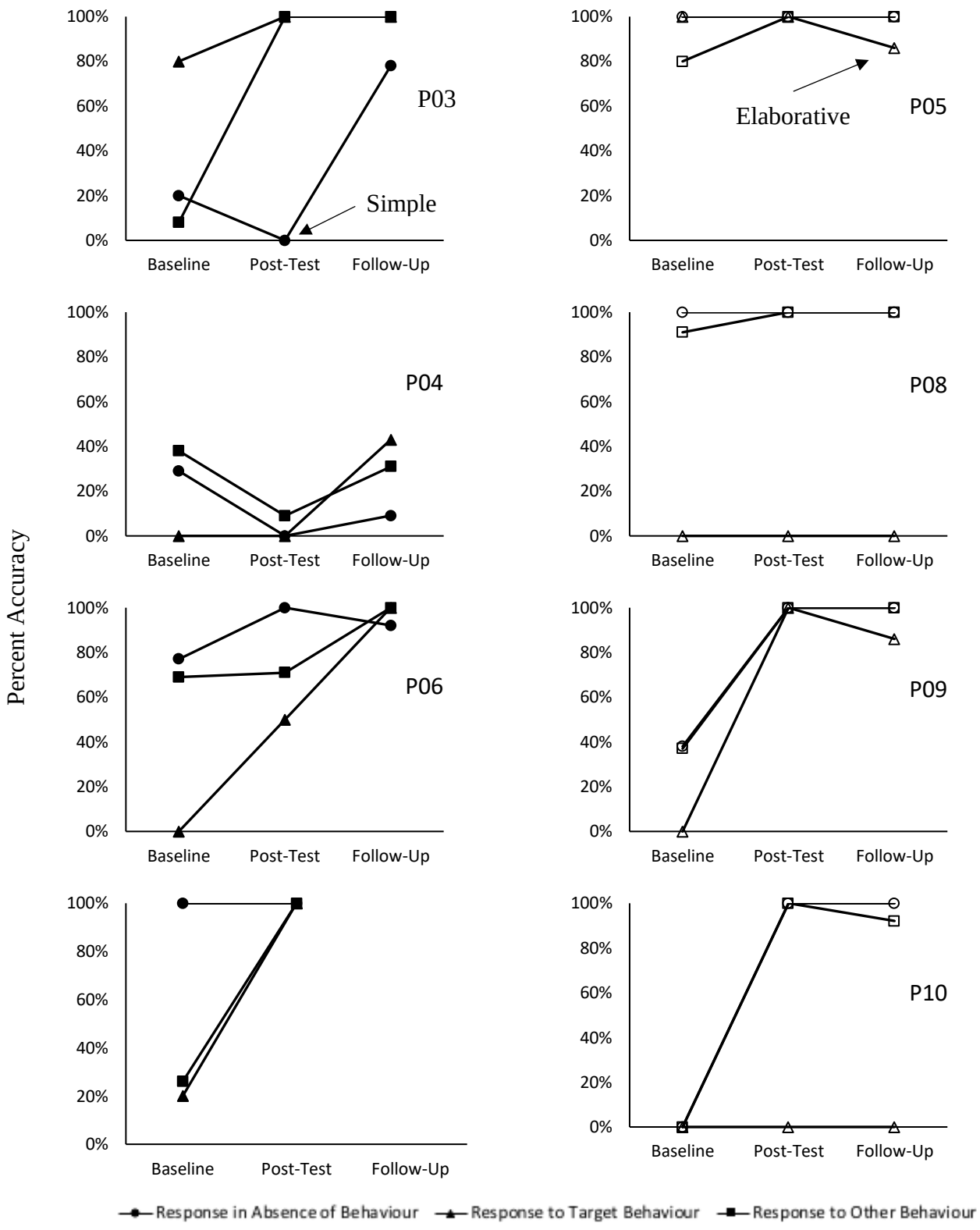

Figure 6. Percent Accuracy per Response Type in the Attention Condition: Simple Feedback vs. Elaborative Feedback.

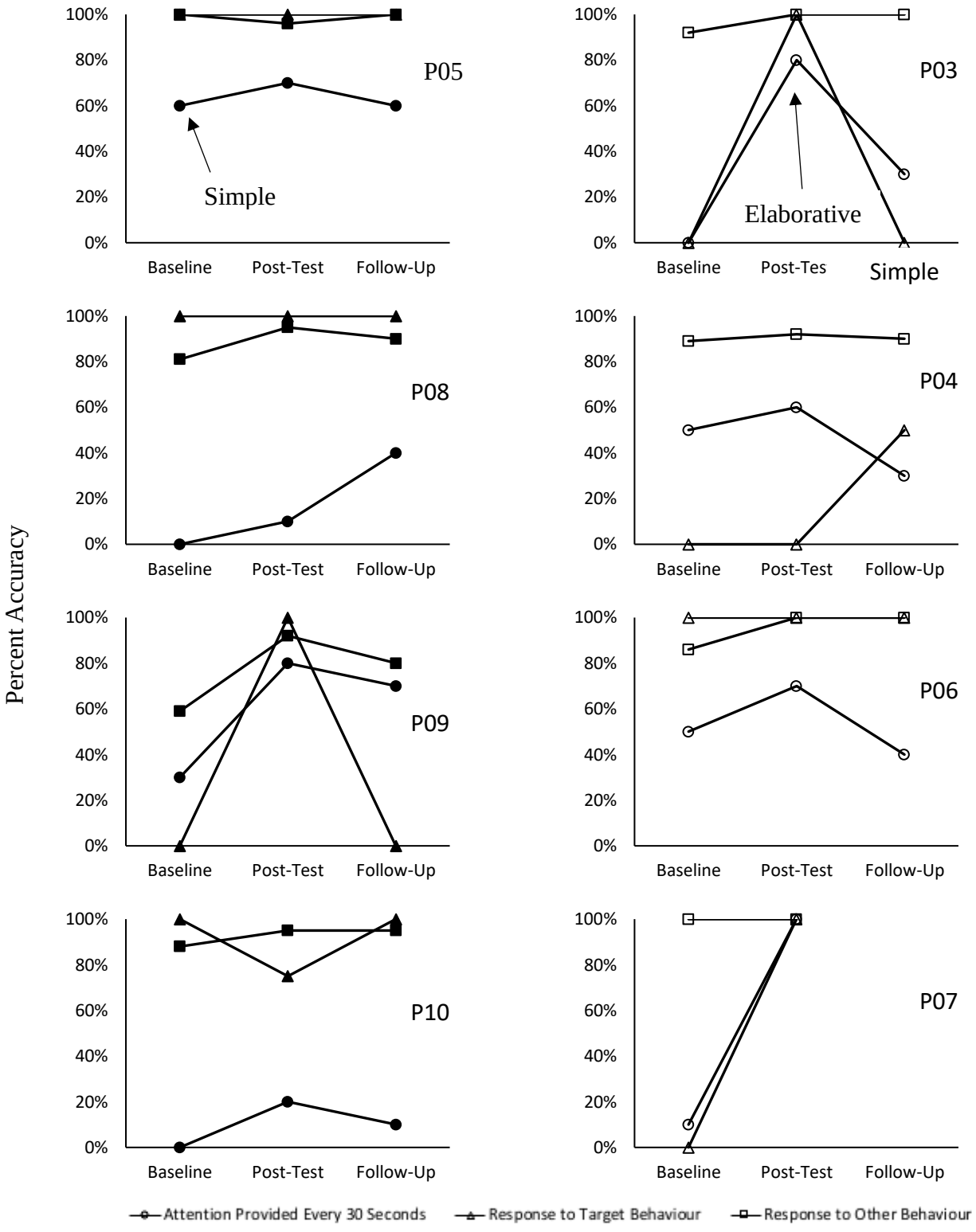

Figure 7 . Percent Accuracy per Response Type in the Play Condition: Simple Feedback vs. Elaborative Feedback.

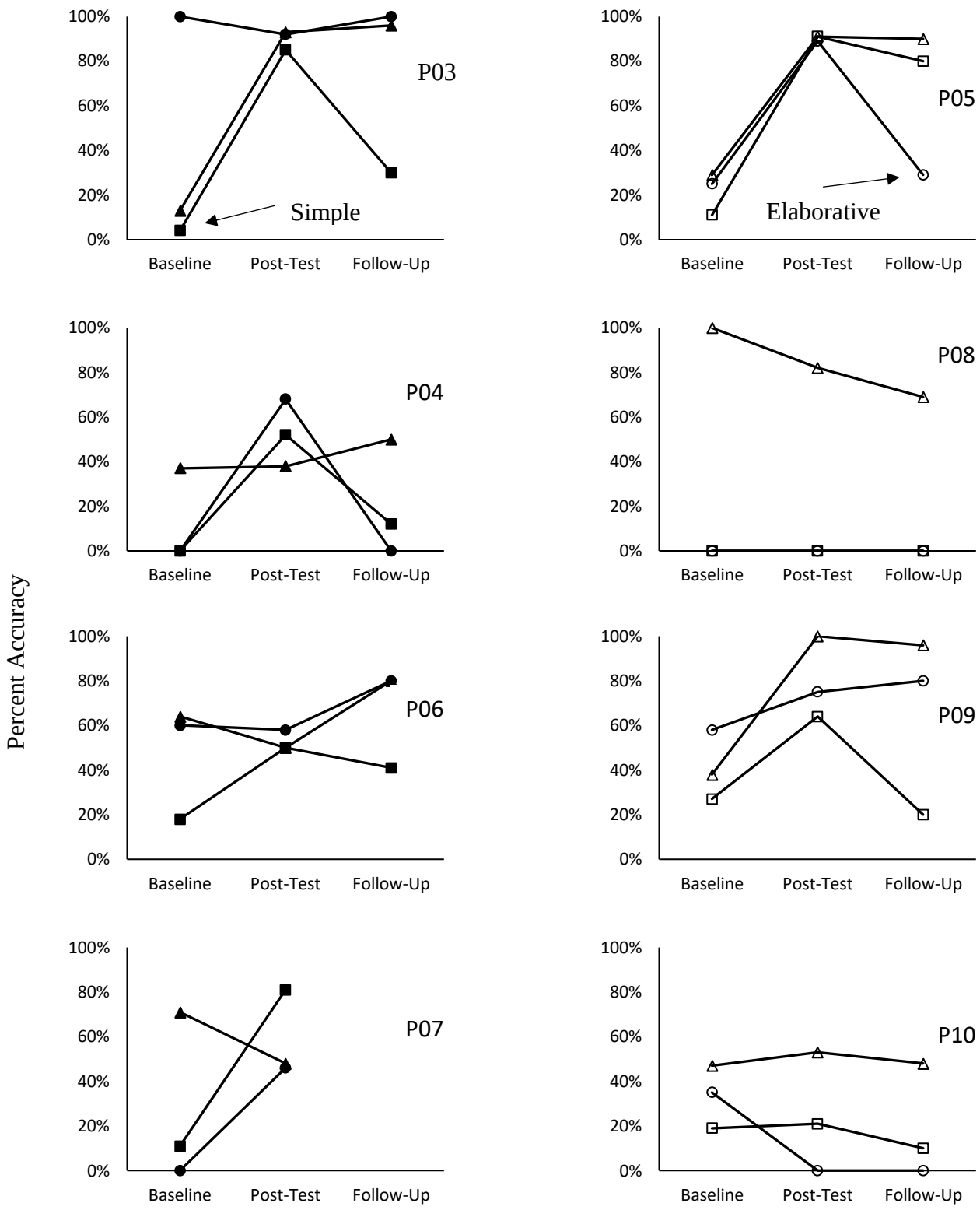

Figure 8. Percent Accuracy per Response Type in the Demand Condition: Simple Feedback vs. Elaborative Feedback.

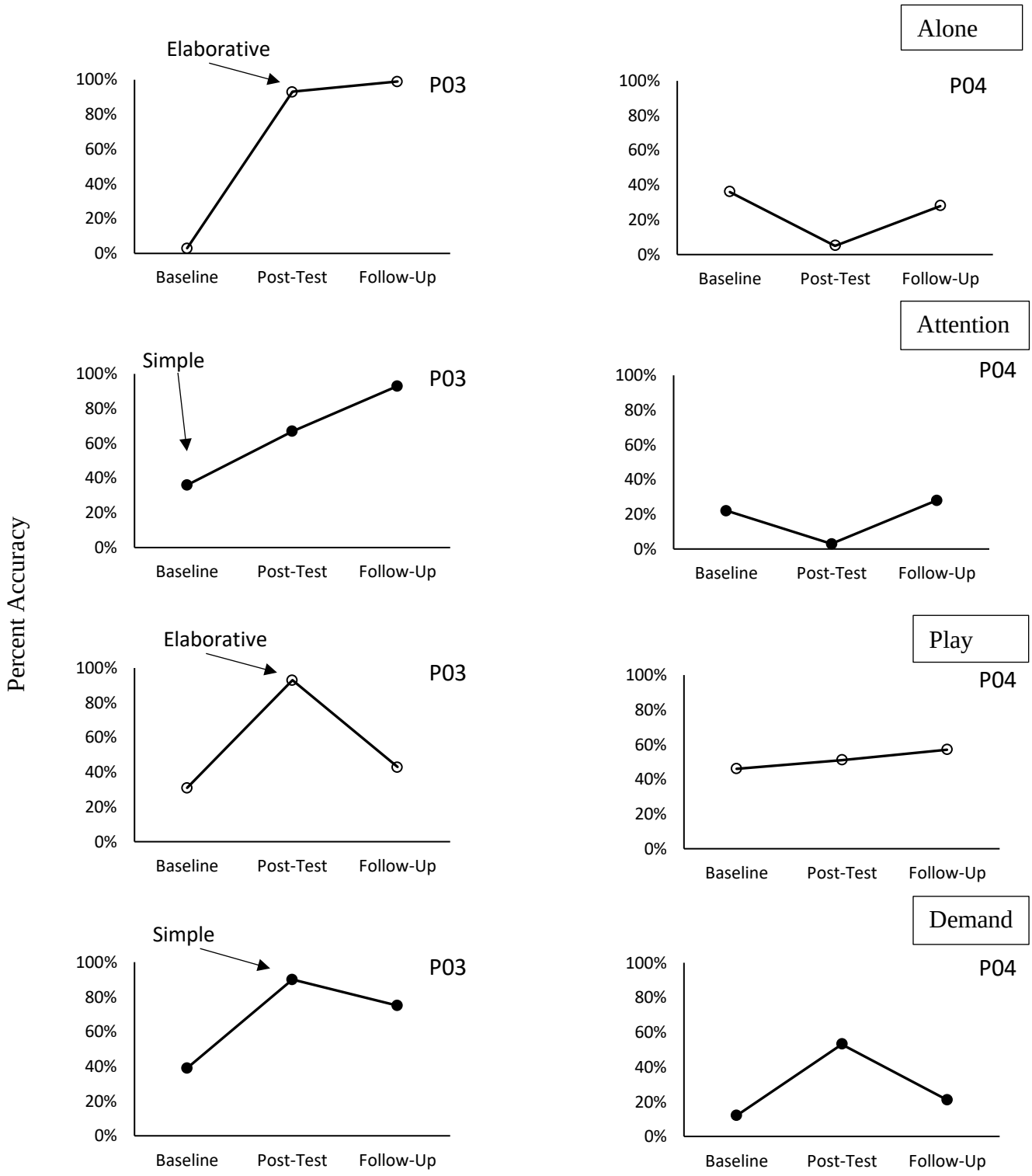

Figure 9 . Total Percent Accuracy for Participants P03 and P04: All FA conditions

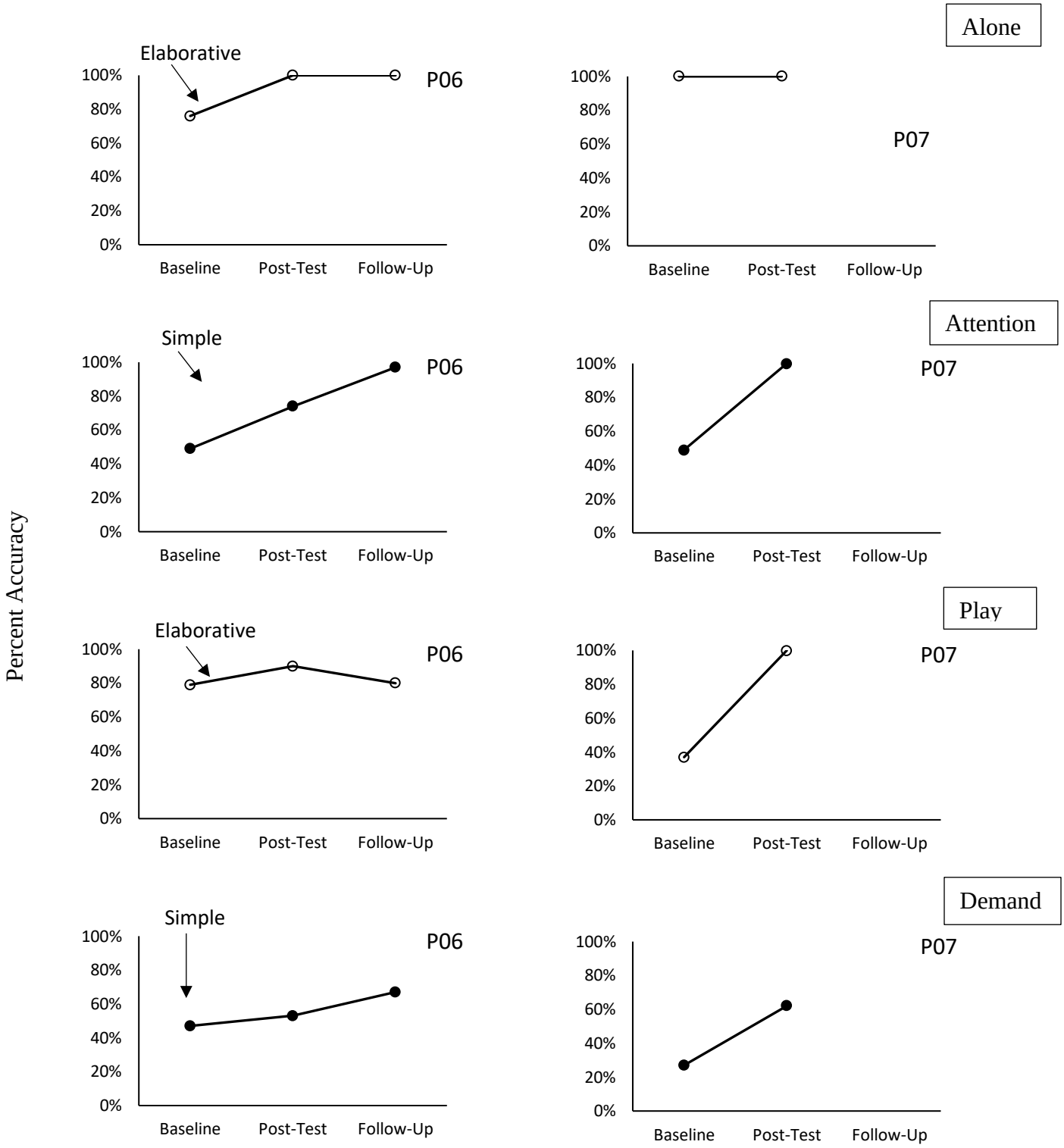

Figure 10 . Total Percent Accuracy for Participants P06 and P07 Group A: All FA conditions

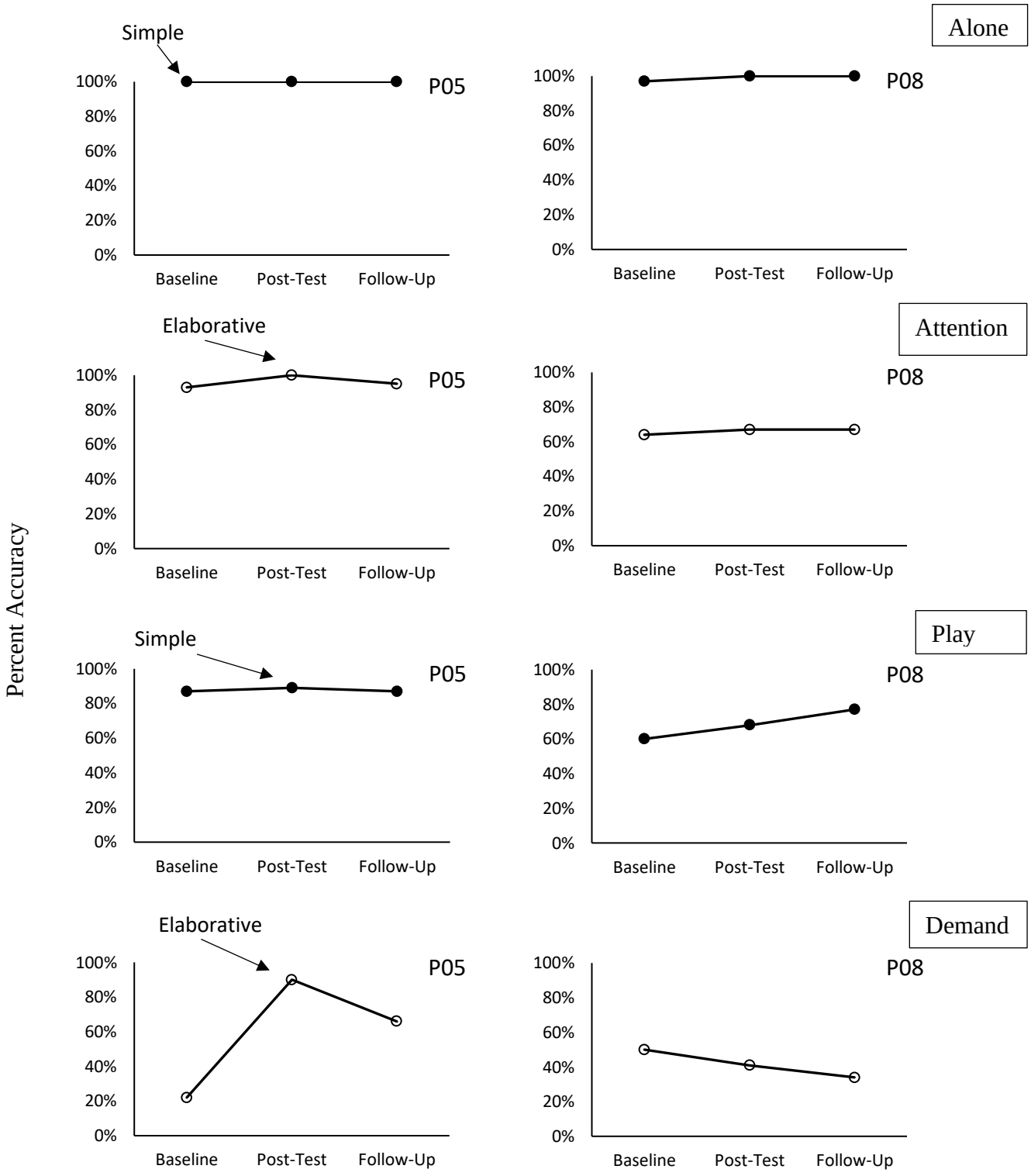

Figure 11. Total Percent Accuracy for Participants P05 and P08 Group B: All FA conditions

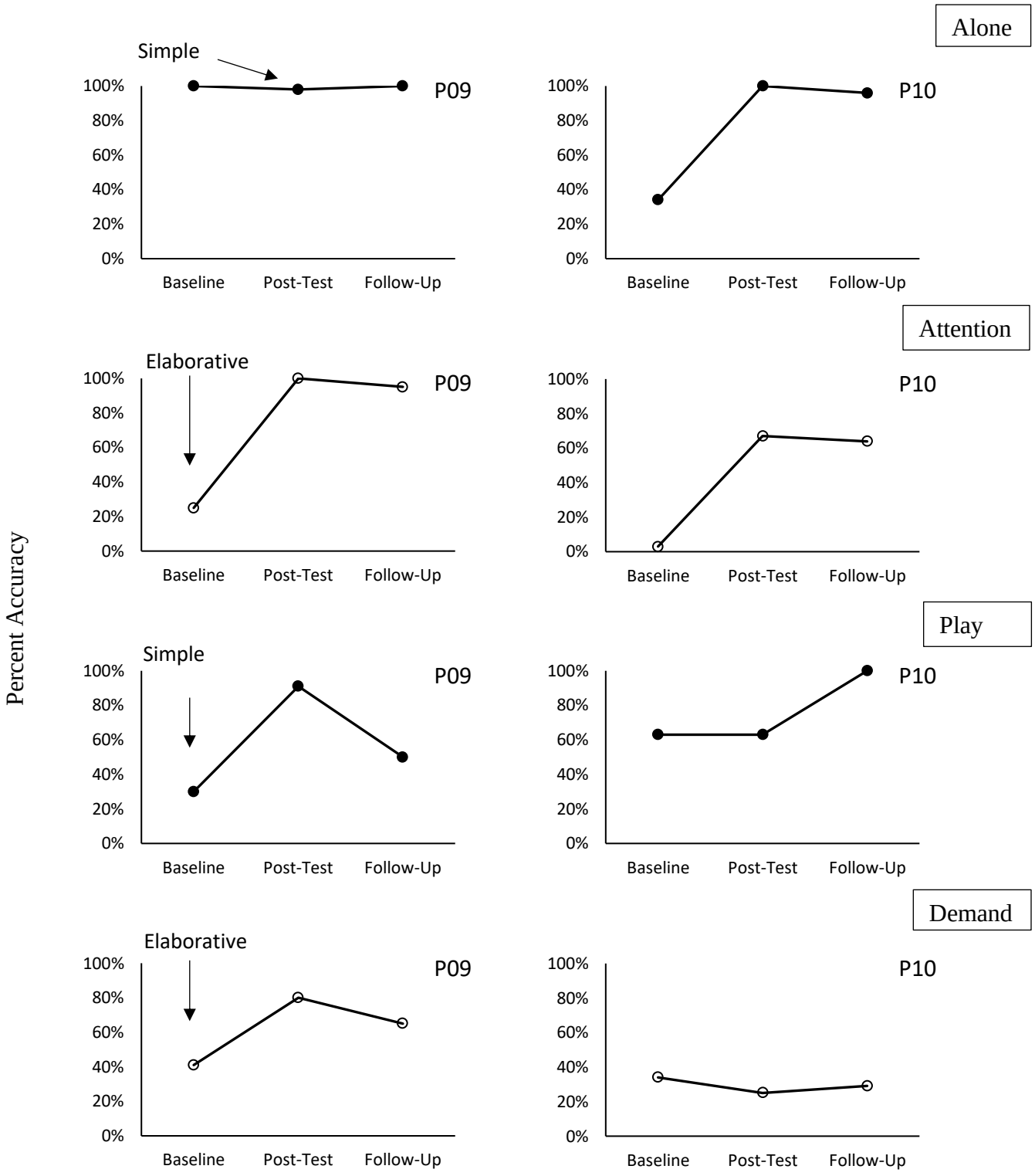

Figure 12. Total Percent Accuracy for Participants P09 and P10 Group B: All FA conditions

Appendices

Appendix A: Pre-Test Questionnaire

Check the box for each answer that states your opinion best					
	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
1. I have previously conducted a functional analysis					
2. I have previously observed a functional analysis					
3. I have previously learned about functional analysis					

Appendix B: Post-Test Questionnaire

Check the box for each answer that states your opinion best					
	Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly Agree 5
1. I found the manual helpful					
2. I found the online tests helpful					
3. I preferred text feedback					
4. I preferred video feedback					
5. I found what I learned valuable					

Appendix C: DTT Manual Chapter
Unit 5: Discrete Trial Teaching or *DTT*

Discrete Trial Teaching (DTT) is a teaching method designed to develop new behaviours and shape existing behaviours. The core features of the DTT method are the instruction, the client's response (behaviour) after receiving the instruction, and the consequence given following the response. In a typical trial, the instructor will be situated across from the client and deliver an instruction (demand) to the client. The instructor will then wait for the client's response. If the client responds, the instructor follows the response with a consequence.

As previously mentioned, DTT is designed to develop new behaviours and shape existing behaviours. Therefore, the consequence that is provided to the client after their response is contingent on whether the client engaged in the appropriate behaviour following the instruction.

We described in Unit 1 how positive reinforcement could increase the frequency of a behavior or maintain the occurrence of a behavior at a high rate. The stimulus presented following the behavior is called a reinforcer.

If the instructor gives an instruction to the client, and the client engages in a correct response, the response is followed by a positive reinforcer. A correct response can be any of the three following behaviours:

Independent – where the client responds correctly to the instructor without needing any prompting from the instructor or the environment (ex. Instructor says "Clap" and the client claps).

Partially Prompted – where the client responds correctly to the instructor after receiving a prompt that only provides a minimal to moderate amount of physical, verbal, gestural, or environmental guidance (ex. Instructor says "Clap", then taps the top of the client's hands, and then the client claps).

Fully Prompted – where the client responds correctly to the instructor after receiving a prompt that provides the most amount of physical, verbal, gestural, or environmental guidance (ex. Instructor says "Clap", then uses both hands to grasp both of the client's hands, and hand-over-hand engages the client in clapping).

If the instructor gives an instruction to the client, and the client engages in a correct response, the response is followed by an error correction. An incorrect response can be any of the two following behaviours:

Challenging Behaviour – where after the instruction is given, the client engages in a problem behaviour (ex. Instructor says "Clap", and the client bites the instructor).

Other non-target Behaviour – where after the instruction is given, the client engages in a behaviour that is not the target behaviour or an approximation of the target behaviour (ex. Instructor says "Clap", and the client stands up and starts jumping on the spot).

Following an incorrect response, the instructor will error correct the client. An error correction procedure is as follows:

1) The instructor will attempt to block the incorrect response.
2) If the attempt to the block the incorrect response was unsuccessful, the instructor will ignore the incorrect response, remove any materials (if necessary), remain neutral in facial expression, and do not provide attention or eye contact to the client for 3-5 seconds.
3) The instructor will re-present the materials (if necessary), re-issue the instruction to the client, and provide the most intrusive prompting necessary immediately after the instruction is given to ensure a correct response.

Below are the steps for running a discrete trial.

Before Running a Discrete Trial

1. This type of teaching is typically done with the instructor and the client in the learning environment. Depending on the target that is intended to be taught, the learning environment may be several different areas. Most often, discrete trials are taught in the learning environment at a table or on the floor, with the instructor situated across from the individual. Subject to the nature of the target being taught, the trial may be performed away from the table or floor if appropriate. Once you have identified the appropriate learning environment, bring any required materials (if any) to the learning environment.

2. When you are in the learning environment, set-up the required materials (if any). If you are teaching at a table, provide a chair for yourself and for the client at the table. Ensure that your chair is close enough to accurately observe the clients response to your instruction and to provide prompting (if needed), blocking (if needed), error correction (if needed), and reinforcement.

   If you are teaching on the floor, select a safe and comfortable area that can accommodate the skill you are teaching (ex. Motor imitation of kicking). Ensure that you are sitting or standing close enough to the client to accurately observe the clients response to your instruction and to provide prompting (if needed), blocking (if needed), error correction (if needed), and reinforcement.

3. Clear all materials not being utilized for the DTT trials from the table (including play materials or otherwise).

4. Have a datasheet prepared containing the exemplars that you will be teaching, a pen or pencil, and a clipboard or binder if needed.

During the session

1. Bring the client to the designated teaching area and have them sit down. If they sit in the desired area, provide them with praise. Then, take your seat.

2. Issue your target instruction to the client. Below is a summary of the potential behaviours that you will engage in following your clients behaviour. For this summary, we will use the scenario of the instructor giving the instruction "clap" to the client, with the target behaviour being the client clapping:

Client's Behavior	Your Behavior
The client performs the target behavior independently or with prompting (Clapping)	Praise the correct behaviour, and record the outcome of the trial on the data sheet.
The client engages in a non-target behavior (Jumping)	Follow the error correction procedure, and record the outcome of the trial on the data sheet.
The client engaging in a challenging behaviour (Biting)	Ignore the behavior, follow the error correction procedure, and record the outcome of the trial.

1. Continue running the desired amount of trials and recording the outcome of the trials on the datasheet, until the duration of the teaching session has reached the desired amount of time or the amount of trials ran is sufficient.

After the Session

1. Allow the client to leave the teaching area.

2. Prepare the same datasheet or a new datasheet for your next DTT session, put away any materials that you no longer need, and collect any new materials that you will require.

Review Exercise for Unit 5

1. *Discrete Trial Teaching (DTT)* is a teaching method designed to_____?

2. What are the three core features of the DTT method?

3. The consequences that are provided to the clients after their responses are contingent on _____?

4. What are the three types of correct responses?

5. What are the two types of incorrect responses?

6. After an incorrect response, what procedure should the instructor follow?

7. When setting up the learning environment, what fives reasons are given for sitting close to the client?

8. During your DTT session, how do you respond if, after you've given an instruction, your client:

    a. Performs the target behaviour correctly?

    b. Performs a non-target behaviour?

    c. Engages in a challenging behaviour?

9. What is the first thing that should be done after a DTT session is finished?

Unit 5 Answer Key

1. *Discrete Trial Teaching (DTT)* is a teaching method designed to *develop new behaviour and shape existing behaviours?*

2. What are the three core features of the DTT method?
   The instruction
   The client's response (behaviour) after receiving the instruction
   The consequence given following the response

3. The consequences that are provided to the clients after their responses are contingent on *whether the client engaged in the appropriate behaviour following the instruction*?

4. What are the three types of correct responses?
   Independent
   Partially Prompted
   Fully Prompted

5. What are the two types of incorrect responses?
   Challenging Behaviour
   Other non-target behaviour

6. After an incorrect response, what procedure should the instructor follow?

   1) *The instructor will attempt to block the incorrect response*
   2) *If the attempt to the block the incorrect response was unsuccessful, the instructor will ignore the incorrect response, remove any materials (if necessary), remain neutral in facial expression, and do not provide attention or eye contact to the client 3-5 seconds*
   3) *The instructor will re-present the materials (if necessary), re-issue the instruction to the client and provide the most intrusive prompting necessary immediately after the instruction is given to ensure a correct response*

7. When setting up the learning environment, what fives reasons are given for sitting close to the client?

   1) *Accurately observe the clients response*
   2) *Provide prompting if needed*
   3) *Blocking if needed*
   4) *Error Correction if needed*
   5) *Provide Reinforcement*

8. During your DTT session, how do you respond if, after you've given an instruction, your client:
   a. Performs the target behaviour correctly?
      Praise the correct behaviour and record the outcome of the trial on the data sheet
   b. Performs a non-target behaviour?

Follow the error correction procedure and record the outcome of the trial on the data sheet

c. Engages in a challenging behaviour?
Ignore the behaviour, follow the error correction procedure, and record the outcome of the trial.

Appendix D: Participant Procedural Checklist (Saltel, 2016)

Alone/Ignore Condition

Correct Target Behaviour:

1. Not initiating any interaction with the client.

2. Not interacting with the client upon occurrence of target or non-target behaviours.

Incorrect Target Behaviour:

1. Initiating interaction with the client.

2. Interacting with the client upon occurrence of target or non-target behaviours.

Attention Condition

Correct Target Behaviour:

1. Not initiating any interaction with the client.

2. Not interacting with the client upon occurrence of non-target behaviours.

3. Interacting with the client upon occurrence of target problem behaviour in the same interval the behaviour occurred of in the following one.

Incorrect Target Behaviour:

1. Initiating interaction with the client.

2. Interacting with the client upon occurrence of non-target behaviours.

3. Not interacting with the client upon occurrence on target problem behaviour.

Play (Control) Condition

Correct Target Behaviour:

1. Initiating interaction with the client within 30 seconds from the beginning of the sessions or from previous interaction.

2. Not interacting with the client for 5 seconds upon occurrence of non-target inappropriate behaviours.

3. Not interacting with the client for 5 seconds upon occurrence of target problem behaviour.

Incorrect Target Behaviour:

1. Not initiating interaction with the client within 30 seconds from the beginning of the sessions of from previous interaction.

2. Interacting with the client within 5 seconds upon occurrence of non-target inappropriate behaviours.

3. Interacting with the client within 5 seconds upon occurrence of the target problem behaviour.

4. Placing demands

Demand Condition

Correct Target Behaviour:

1. Presenting a demand within 30 seconds from the beginning of sessions or from previous demand.

2. Not removing demand upon occurrence of error, no response, or occurrence of non-target behaviours. The trainee could either keep presenting the same demand or prompt the correct response.

3. Removing material upon occurrence of the target problem behaviour.

4. Removing demand upon occurrence of the target problem behaviour.

Incorrect Target Behaviour:

1. Not presenting a demand within 30 seconds from the beginning of sessions or from previous demand.

2. Removing demand upon occurrence of error, no response, or occurrence of non-target

behaviours.

3. Not removing material upon occurrence of the target problem behaviour.

4. Not removing demand upon occurrence of the target problem behaviour.

Appendix E: Applied Test Datasheets

Functional Analysis Coding Datasheet: Alone/Attention

Did the participant set up a datasheet: YES or NO — Video Code:

Did the participant set up the environment correctly: YES or NO — Coder:

	Interval	Confederate Target Behavior		Participant Response Correct to Target			Confederate Other behaviour		Participant Response Correct to Other			Participant Correct in Absence of Behaviour		
		YES	NO	YES	NO	NA	YES	NO	YES	NO	NA	YES	NO	NA
1.	0-9s	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
2.	10-19s	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
3.	20-29s	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
4.	30-39s	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
5.	40-49s	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
6.	50-59s	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
7.	1:00-1:09m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
8.	1:10-1:19m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
9.	1:20-1:29m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
10.	1:30-1:39m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
11.	1:40-1:49m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
12.	1:50-1:59m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
13.	2:00-2:09m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
14.	2:10-2:19m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
15.	2:20-2:29m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
16.	2:30-2:39m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
17.	2:40-2:49m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
18.	2:50-2:59m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
19.	3:00-3:09m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
20.	3:10-3:19m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
21.	3:20-3:29m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
22.	3:30-3:39m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
23.	3:40-3:49m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
24.	3:50-3:59m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
25.	4:00-4:09m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
26.	4:10-4:19m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
27.	4:20-4:29m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
28.	4:30-4:39m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
29.	4:40-4:49m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
30.	4:50-4:59m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
31.	5:00-5:09m	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Appendix E: Continued

Functional Analysis Coding Datasheet: Control

Did the participant set up a datasheet: YES or NO Video Code:

Did the participant set up the environment correctly: YES or NO Coder:

	Interval	Confederate: Target Behavior	Participant Response: ignore for at least 5 secs	Confederate: Other behaviour	Participant Response: Correct to Other	Attention provided within 30 seconds
		YES NO	YES NO NA	YES NO	YES NO NA	YES NO
1.	0-9s					NA
2.	10-19s					
3.	20-29s					
4.	30-39s					
5.	40-49s					
6.	50-59s					
7.	1:00-1:09m					
8.	1:10-1:19m					
9.	1:20-1:29m					
10.	1:30-1:39m					
11.	1:40-1:49m					
12.	1:50-1:59m					
13.	2:00-2:09m					
14.	2:10-2:19m					
15.	2:20-2:29m					
16.	2:30-2:39m					
17.	2:40-2:49m					
18.	2:50-2:59m					
19.	3:00-3:09m					
20.	3:10-3:19m					
21.	3:20-3:29m					
22.	3:30-3:39m					
23.	3:40-3:49m					
24.	3:50-3:59m					
25.	4:00-4:09m					
26.	4:10-4:19m					
27.	4:20-4:29m					
28.	4:30-4:39m					
29.	4:40-4:49m					
30.	4:50-4:59m					
31.	5:00-5:09m					

Appendix E: Continued

Functional Analysis Coding Datasheet: Demand

Did the participant set up a datasheet: YES or NO Video Code:

Did the participant set up the environment correctly: YES or NO Coder:

CT Confederate Target Behavior **CO Confederate Other behaviour** **CD Confederate Response to Demand**		**IB Ignore Target Behaviour or Other Behaviour** **RI Participant Response to Target: removed items** **RD Participant Response to Target: removed demand 30s (+/- 2 s)**	**DP Participant Response Correct to Demand: PROMPT (wait 5-7 seconds)** **P Participant Response Correct to Demand: Praise (within 2-3s)** **D Placed Demand correctly (2-3 secs after response or initial)** **DX Demand should have been placed (only score in incorrect)**
Interval	**Confederate**	Correct	Incorrect
0-9			
10 - 19			
20-29			
30-39			
40-49			
50-59			
1:00-1:09			
1:10-1:19			
1:20-1:29			
1:30-1:39			
1:40-1:49			
1:50-1:59			
2:00-2:09			
2:10-2:19			
2:20-2:29			

Appendix F: Procedural Integrity for Sessions

Procedural Integrity Checklist for Sessions

1) Present and Review Consent form
2) Give and review questionnaire prior to delivering written pre-test
3) Performed script for Appendix B/Baseline/Pre-test
4) Provide participant with written pre-test
5) Performed Script (Appendix B2) prior to FA role play with confederate
    a. Materials were readily available
6) Provide participant with Self-Instructional Manual and follow Appendix C script
7) Following participant complete of SIM unit, followed script Appendix D
8) Provided feedback and scored CAPSI unit tests
    a. Unit 1
    b. Unit 2
    c. Unit 3
    d. Unit 4
    e. Unit 5
    f. Unit 6
9) Performed script D2 and gave participant written post-test and materials for FA post-test
10) Provided Participant with honorarium
11) Discussed scheduled follow-up session