Fall 2016

The Reliability and Validity of the Task Analysis Recording Procedure (TARP)

Russell Brown
University of Southern Maine

THE RELIABILITY AND VALIDITY OF THE TASK ANALYSIS RECORDING PROCEDURE (TARP)

By

Russell Brown

B.A. University of Southern Maine, 1990

M.S. University of Southern Maine, 2006

A DISSERTATION

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Psychology

(in School Psychology)

The University of Southern Maine

August, 2016

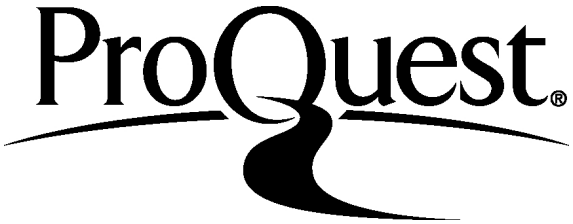
Advisory Committee:

Mark W. Steege, Professor of School Psychology, Advisor

Rachel Brown, Associate Professor of School Psychology

Harold Longenecker, Ph.D., Licensed Psychologist

Erin Sullivan, Psy.D., Board Certified Behavior Analyst (BCBA)



LIBRARY RIGHTS STATEMENT

In presenting the Dissertation, THE RELIABILITY AND VALIDITY OF THE TASK ANALYSIS RECORDING PROCEDURE (TARP), in partial fulfillment of the requirement for the Psy.D. in School Psychology at the University of Southern Maine, I agree that the Library shall make it freely available for review. I further agree that permission for copying, as provided for by the Copyright Law of the United States (Title 17, U.S. Code), of this Dissertation for scholarly purposes may be granted. It is understood that any copying or publications of this Dissertation for financial gain shall not be allowed without my written permission.

I hereby grant permission to the University of Southern Maine Library to use my Dissertation for Scholarly purposes.

_____

Russell Brown

May 2016

THE RELIABILITY AND VALIDITY OF THE TASK ANALYSIS RECORDING PROCEDURE (TARP)

Russell Brown, M.S.

Dissertation Chairperson: Dr. Mark Steege

An Abstract of the Dissertation Presented

in Partial Fulfillment of the Requirements for

the Degree of Doctor of Psychology

(in School Psychology)

University of Southern Maine

May, 2015

Task analysis data collection typically focuses on the acquisition of skills by recording the percentage of steps in the response chain completed independently and correctly. While useful as a measure of skill acquisition, percentage correct does not promote a step based analysis of factors that may promote or interfere with skill acquisition, including necessary prompts and the occurrence of challenging behavior. This study evaluated the reliability and validity of the Task Analysis Recording Procedure (TARP) in recording physical stereotypy, a behavior often emitted by participants with autism or other developmental disabilities, by comparing TARP obtained physical stereotypy data to that obtained via six second momentary time sampling. A multiple probe design was utilized to facilitate the comparison. The results show a robust correspondence between

recordings of physical stereotypy conducted by teachers using the TARP and secondary observers utilizing a six second momentary time sampling procedure. This study demonstrates that the TARP procedure is an acceptable means of recording physical stereotypy in applied settings. Moreover, these results demonstrate a teacher-friendly method of recording both the acquisition of skills and the decrease of interfering stereotypy within the context of functional life skills programming. Implications of these findings and suggestions for further research are discussed.

ACKNOWLEDGEMENTS

I would like to thank Dr. Mark Steege and Dr. Rachel Brown for the time and effort that they have put in to teaching and for keeping the program afloat during ridiculous times at the U.

I would like to thank folks on my dissertation committee, Dr. Mark Steege, Dr. Rachel Brown, Dr. Longenecker and Dr. Sullivan for the support and advice offered that helped me through the process.

I would like to personally thank Dr. Harold Longenecker for keeping me off the table and moving forward.

I would like to thank my MMCC/USM cohort, William and Michelle, for dragging me along. It is easier to tolerate challenging circumstance when you are in the midst of laughing.

Special thanks to the MMCC-Charles Street crew for support of this dissertation but more importantly for the work you do on a daily basis. You folks exemplify what is best about education.

Most importantly, I would like to thank my wife, Julie, and my son Ronan for being the cool people that they are. You both deserve an honorary degree and/or a medal of honor for the support you have provided. You are a fun duo to return home to every night—you definitely signal an improving condition.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW

Applied Behavior Analysis (ABA) is a science of human behavior focusing on improving socially significant behaviors. Social significance refers to behaviors that are meaningful, useful, and practical (Baer, Wolf, & Risley, 1968). For learners with disabilities, emphasis is placed on teaching socially significant skills that lead to increased independence and participation within school, residential, community and employment settings. Such skills are often referred to as functional life skills (FLS).

Systematic instructional strategies are typically used to teach functional life skills (Browder & Spooner, 2011; Steege & Watson, 2009). One such method for teaching complex behavior chains involves the use of task analysis. Task analysis involves breaking a complex skill into component parts of tasks or behavior chains. Task analysis can be highly effective in promoting skills necessary to increase independence (i.e., activities of daily living) and is a core feature of functional life skills programs for learners with disabilities (Browder & Spooner, 2011; Haring & Kennedy, 1988; Storey & Miner, 2011).

A key component of effective teaching utilizing task analysis is the utilization of effective prompting and fading strategies. When using a task analysis it is critical that the learner does not become reliant on artificial prompts (i.e., a prompt provided by an instructor). Prompt dependence can prevent the learner from achieving independence. Rather, the learner should become reliant on natural stimuli and reinforcement so that the next step in the chain will occasion the next response, and so on. Further, the learner should experience as few errors as possible during instruction so that correct learning of the target skill will occur most efficiently (Grow et al., 2009).

In planning for the fading of prompts, practitioners have several procedures available to transfer stimulus control from artificial prompts to the naturally occurring stimuli. These include most-to-least prompts, graduated guidance, least-to-most prompts and time delay (Cooper et al., 2007). Most to least prompts involve the teacher physically guiding the learner through a task and systematically fading prompts to less restrictive prompts in the hierarchy as the learner increasingly attends to relevant stimuli. Least to most prompt fading entails the teacher giving the learner opportunity to complete a step with the least restrictive prompt available and then providing more restrictive prompts in the hierarchy based on lack of performance (i.e., only if needed). Currently, there is a lack of research which supports a particular package of prompting and fading strategies indicating that either least to most or most to least is better (McKay, Weiss, Dickson, & Ahearn, 2014; Seaver & Bourret, 2014).

Several researchers have offered variations to the traditional models of prompt fading packages in an effort to increase the efficiency of skill acquisition and to minimize the occurrence of errors. Recently, McKay et al. (2014) compared a most to least prompting procedure with a two second delay to a simplified hand over-hand guidance procedure with a 2 second delay. Results indicated that there was no significant difference between the efficiency of the two procedures. Yet, given the less complicated method utilized in the hand-over-hand procedure the authors argued that it might be the better alternative due to the ease of implementation.

Steege, Wacker, & McMahon (1987) offered an alternative to a least to most prompt fading hierarchy. In the study the authors compared the effectiveness and efficiency of traditional least to most prompting as compared to “prescriptive prompting.”

Prescriptive prompting utilized ongoing performance assessment of the level of prompting necessary to occasion a response and used that information to inform subsequent trials. The authors found that while both methods were effective in promoting task acquisition, the prescriptive prompting method proved to be more efficient.

Task analysis data collection typically focuses on the acquisition of skills by recording the percentage of steps in the response chain completed independently and correctly. While an accurate measure of skill acquisition, percentage correct does not promote analysis of the variables, including the prompting method utilized, that may have interfered with the individual's acquisition of steps in the behavior chain (Haring & Kennedy, 1988).

When working with learners with developmental disabilities (DD) and/or autism spectrum disorders (ASD) to teach skills, teachers are frequently confronted by the learners' presentation of challenging behavior. These behaviors may include refusal, stereotypy, aggression or self-injury, among others. While numerous strategies for reducing problem behavior have been reported (e.g., punishment, extinction, among others) current practice emphasizes the teaching of functionally-equivalent and incompatible socially meaningful replacement behaviors as a way of both decreasing challenging behaviors and increasing appropriate behaviors (Horner et al., 1990; Koegel, Koegel & Dunlap, 1996; Steege & Watson, 2009).

The use of interval recording procedures is common in empirical studies for recording a wide range of challenging behaviors (Thompson & Borreo, 2011; Wacker, Berg, Harding, & Cooper-Brown, 2011). Two methods of behavior measurement utilized

in applied research are partial interval recording procedures and momentary time sampling (MTS). Meany-Daboul, Roscoe, Bourret and Ahearn (2007) found that the better method for estimating behavior amenable to duration measures was MTS, with partial interval recording preferred for frequency events. While both methods may be considered a "gold standard" method for recording challenging behavior (Steege, Davin, & Hathaway, 2001), what is common to both in applied settings is that they are difficult procedures to implement reliably unless one person is solely dedicated to doing so. In applied settings, one person is often charged with collecting instructional data, implementing correct prompting procedures, and managing interfering behaviors. Within such a setting, collecting either 6 or 10 second partial interval or momentary time sample data, in addition to other duties, is onerous and, accordingly, might not yield trustworthy information.

As noted above, recording the percentage of steps completed is an accurate method of measuring participant progress, but it does not yield information pertaining to challenging behavior. Similarly, direct observation (e.g., 6 second interval or momentary time samples) measures challenging behaviors, but it does not show when and where challenges occur during the instructional process. Further, these procedures are unrealistic within many applied settings because one person cannot accurately record both the correct number of steps performed and simultaneously use an interval procedure to record challenging behavior.

To address data collection problems within task analysis Steege and Watson (2009) proposed the use of the Task Analysis Recording Procedure (TARP). The TARP is both a teaching and behavior recording mechanism. The TARP generates percentage

correct data that are most often utilized in association with task analysis, provides a mechanism for recording utilized prompts, and affords the ability to record interfering behaviors that occur within the context of the specific steps of the instructional task. Steege and Watson (2009) demonstrated the utility of the TARP for both documenting the increase of functional life skills (i.e., the percent of steps performed correctly) and the decrease in challenging behaviors (i.e., the percent of times during the task analysis that challenging behavior occurred). While the TARP has been demonstrated to be useful and practical in documenting behavior change within applied settings, the accuracy and reliability of a step-based method of recording challenging behavior has not been demonstrated.

The research questions addressed by this study are:

1. Does the use of task analyses and systematic instruction procedures to teach functional life skills with participants with DD or ASD increase levels of independent skill acquisition?
2. Does the TARP adequately measure the acquisition of skills and the concomitant decrease in challenging behaviors?
3. Is the TARP an accurate and valid tool for step-based recording of challenging behavior during task analysis instructional programming?

CHAPTER 2: METHOD

Participants and Setting

There were both student and teacher participants in this study.

Students. Four elementary school aged participants attending a school-based day treatment center for children with autism and other developmental disabilities, located in the Northeast US, participated in the study. To be eligible to participate in the study all participants needed a current diagnosis of Autism Spectrum Disorder (ASD), adaptive skill deficits of more than two standard deviations below the mean as indicated on a norm referenced measure of adaptive behavior (i.e., Vineland Adaptive Behavior Scales-2nd Edition) and a current Functional Behavioral Assessment (FBA) documenting the occurrence of high rates of physical stereotypy.

Mary* was 12 years old and diagnosed with an Intellectual Disability and Down Syndrome in addition to ASD. The operational definition of physical stereotypy for Mary was: occurrences of hand flapping, arm waving, banging hands or objects together, rubbing hands or arms up and down body, rubbing objects, or rubbing body on floors, walls or furniture. The skill targeted for instruction for Mary was making a preferred drink (i.e., lemonade).

Tim was an eight years old. The operational definition of physical stereotypy for Tim was: occurrences or episodes of repetitive physical movement/motor activity including bouncing items in hand, using back of hand to tap surfaces, sifting or shaking

* All participant names are pseudonyms.

materials/objects. The skill targeted for instruction for Tim was making a preferred drink (i.e., fruit punch).

Bob was 11 years old. Bob's operational definition of physical stereotypy was: repetitive, non-adaptive motor behavior such as arm flapping or waving; tapping objects or hard surfaces with fingers; spinning in circles, pacing back and forth, jumping up and down; or touching, grabbing or rubbing genitals with hands, elbows, with or against objects. The skill targeted for instruction for Bob was recycling. Bob was taught how to sort bottles and cans.

Amy was 12 years old. Amy's operational definition of physical stereotypy was: occurrences of visually tracking hands, repetitive head shaking, body posturing, arching back, covering eyes, and playing with hair. The skill targeted for instruction for Amy was shoe tying.

All participants had individual staff support from one or more paraprofessional teaching aides. The paraprofessionals provided instruction and served as the primary recorders of physical stereotypy on the TARP. A total of seven paraprofessionals participated in the study and provided instruction for the four participants. Sessions were run two times daily, with each participant's morning paraprofessional running morning sessions and the afternoon paraprofessional running the afternoon session. One participant (Amy) had the same paraprofessional across morning and afternoon sessions.

Teachers. The paraprofessionals who worked with the students were the teacher participants. They completed a post-intervention survey about how well they liked the

TARP procedure. The paraprofessionals who completed surveys signed IRB-approved consent forms prior to participation.

Two Board Certified Behavior Analysts (BCBAs), two Masters Level Special Education Teachers, and the primary investigator collected six second momentary data and interobserver agreement (IOA) data.

All sessions were conducted in the cafeteria of the school. The cafeteria was approximately 20 x 25 feet large and contained five tables for dining, bins for garbage and used kitchen utensils, and a water cooler. An additional table was set up in the cafeteria for use in the study. The table used for all participants was a 2.5 x 4 foot table located along the wall of the cafeteria. Other students and staff were routinely present in the cafeteria during sessions.

Independent Variables and Materials

A task analysis recording procedure (TARP) data collection sheet was created for each participant based upon the skill targeted for acquisition. For Mary and Tim the recording sheets were identical because they were both instructed on the same target skills. For Bob and Amy, TARP data collection sheets were developed for their respective skills. Each TARP displayed all steps of the task analysis along with corresponding columns designed to document the steps of the task analysis that were completed independently as well as the occurrence or non-occurrence of physical stereotypy. A column for recording the specific level of prompt was included but not utilized due to a decision to utilize a hand over hand with a two second delay prompting procedure. Each participant's operational definition was included on the TARP along

with treatment integrity targets and a calculation table based upon a 13 step task analysis. An enlarged and laminated version of each TARP and a dry erase marker were provided to the paraprofessionals in order to facilitate ease of data collection during instruction.

Corresponding six second MTS sheets were created for use by the secondary observers for each participant and included the respective participant's operational definition of physical stereotypy. A Timex® Ironman watch was utilized to monitor time during six second analysis and baseline conditions. A Sony® Handycam HDR-CX 405 was used to record sessions. Ongoing data management was completed utilizing Microsoft Excel® on a HP® laptop.

For Amy and Tim necessary materials included a plastic bottle of water and drink packets. Amy used Great Value® brand lemonade and Tim used Crystal Light® fruit punch or lemonade packets.

For Bob, necessary materials included a box filled with 10 empty beverage containers (a mixture of bottles and cans) and a recycling container. The recycling container was a container the size of a garbage can that had a cover with two holes on the top. One hole was labeled "bottles" and the other hole labeled "cans".

For Amy, the only materials necessary were the table, a chair and one of her shoes that was removed prior to the start of each session.

The prompting procedure utilized by each paraprofessional was hand-over-hand with a two second delay. The paraprofessionals were instructed to silently count "one thousand one, one thousand two" and provide hand over hand prompting if the participant had not initiated the required step in the task analysis. Additionally, they were instructed

to ignore the two second delay and immediately utilize hand-over-hand if the participant was making an error. The paraprofessionals were not asked to record whether the prompting procedure was utilized as it was assumed to have been used if the step was not completed independently.

Dependent Variables and Data Collection

The dependent measures used to evaluate the TARP included (a) the participants' independent completion of steps on the task analyses, (b) the participants' display of physical stereotypy recorded via the TARP and six second MTS, and (c) acceptability ratings of the TARP method completed by the paraprofessionals who used it.

Each participant's paraprofessional(s) served as the primary observers and recorders of behavior utilizing the TARP. The paraprofessionals recorded whether the step of the task analysis was completed independently in the designated independence column by marking a (+). If a prompt was required, the paraprofessional recorded a (-) in the independence column. Upon completion of the TARP the paraprofessional calculated the level of independence by using the chart on the TARP and recorded it on the document.

To record whether physical stereotypy occurred during a specific step of the task analysis, the paraprofessional recorded a (+) if the behavior was observed or left it blank or marked a (-) if it was not observed. The paraprofessional then counted the number of steps in which physical stereotypy was observed and used the calculation chart on the TARP in order to determine the overall percentage of steps in which physical stereotypy was observed.

Secondary observers watched video recordings of instructional sessions and recorded the occurrence or non-occurrence of physical stereotypy using six second MTS. Six second recording started upon the paraprofessional's delivery of the instructional prompt (i.e.,“Mary, make your drink”, “Bob, sort the returnables”). The same video recordings were also utilized to collect IOA data on the TARP (independence and step-wise recording of physical stereotypy) and to collect 6 second MTS IOA.

Procedure

Procedural Safeguards and Informed Consent. All study methods and procedures were approved through the University of Southern Maine (USM) Institutional Review Board (IRB) prior to implementation. Informed consent to participate in the study was obtained from a parent of each participant. Due to the participant’s diminished capacity to provide assent, a waiver for assent and its documentation was granted through the IRB. In order to participate in the study consent was also obtained from paraprofessionals.

Interventionist Training. Prior to the onset of the investigation, the following interventionist training modules were completed: (a) All interventionist were trained on the components of the TARP; (b) All interventionist were taught the hand-over-hand prompting procedure and participated in role play activities; (c) Secondary data recorders were trained in six second MTS procedures and achieved greater than 90% IOA in practice sessions; (d) Interventionist members reviewed operational definitions of physical stereotypy for the participant that they supported; secondary observers reviewed behavioral definitions for all participants prior to data collection activities; (e) All

interventionist were provided in vivo performance feedback from the primary investigator; (f) Primary observers requiring further support were provided video recordings of sessions to score until meeting 90% accuracy.

Baseline. During the baseline condition each participant was brought to the table in the cafeteria with necessary materials present. Each participant was then given a specific instructional prompt to complete the task (i.e. “Mary, make your drink). During baseline, a timer was set for two minutes and 6 second MTS data were collected for each student. The paraprofessionals were not required to collect physical stereotypy data during baseline; rather they recorded any steps completed independently. No prompts or consequences were delivered during baseline. Baseline data were collected via intermittent probes for 3 out of the 4 participants in order to reduce the potential for frustration and decrease the amount of time in a no instruction condition. The exception was Tim, who was moved from baseline into the intervention phase despite showing decreases in stereotypy and increases in independence. The decision to do this was based upon Tim’s emerging use of vocal language. Tim was verbally requesting “help” and it was determined it would be detrimental to his overall language development not to honor his requests.

Intervention. At the onset of the study, tasks were identified for acquisition for each participant through consultation with MMCC staff and participants’ parents. Task analyses were then conducted by breaking down the respective tasks into finite components. Individual TARP forms were developed for each participant and included: (a) the individual steps in the behavior chain, (b) a column for recording independence on a step-wise basis, (b) prompting methodology, (d) a column for recording physical

stereotypy on a step-wise basis, (e) the operational definition of physical stereotypy for the particular student, (f) a chart for easy conversion of percentages of steps completed independently and steps with physical stereotypy and (g) a section to record components of treatment integrity to be completed by secondary observers. Please see Appendix A for copies of TARPs for each participant.

During intervention sessions, paraprofessionals brought the participant to the table in the cafeteria where the necessary materials were located. The paraprofessional then provided an instructional prompt to begin the task (i.e. “Mary, make your lemonade”). From that point the only language provided by the paraprofessionals was praise for steps completed independently.

Paraprofessionals utilized a two second delay before implementing the hand-over-hand prompting procedure for each step, if necessary. Errors were interrupted immediately with a hand over hand prompt. Paraprofessionals recorded independence and steps with physical stereotypy on a laminated sheet located on the table and then transferred that information to the recording sheet upon conclusion of the instructional session. All sessions were video recorded for post-session six second MTS analysis and IOA data collection.

Treatment Integrity

Treatment integrity data were collected for 100% of intervention sessions. A checklist was located on the TARP and completed by a secondary observer once the paraprofessional finished recording data on the TARP. Treatment integrity was defined as: (a) the paraprofessional had all necessary materials, (b) appropriate reinforcement was

provided for steps completed independently, and (c) prescribed response prompts were followed (i.e., 2 second delay with hand-over-hand). Overall treatment integrity across 59 intervention sessions was 94.1% and deemed acceptable (Table 2.1). Please see Table 2.1 for individual participant treatment integrity data.

Table 2.1

Treatment Integrity Data

Participant	Number of Intervention Sessions	Treatment Integrity Percentage
Mary	23	93.8
Tim	18	91.9
Bob	9	100
Amy	9	96.2
Total	59	94.1

Interobserver Agreement

Interobserver agreement (IOA) was calculated via video recordings of sessions. The following data were compared: (a) independence on the TARP (step by step), (b) physical stereotypy on the TARP (step by step), and (c) six second MTS (interval by interval). All IOA scores were obtained by dividing the number of agreements by the number of agreements by disagreements, multiplied by 100.

For steps completed independently, paraprofessional recordings of independence were compared to those of secondary observers who scored steps completed

independently via video recordings. 52% of sessions were scored for IOA with the overall average IOA being 95%, which was deemed acceptable (Table 2.2).

Table 2.2

Percentages of inter-observer agreement between primary and secondary recorders of TARP data and between observers collecting MTS data

Participant	Overall IOA/percentage of sessions	TARP Independence	TARP PST	6 Second MTS
Mary	IOA	93	85	92
	Percent of Sessions/IOA	82	78	63
Tim	IOA	93	88	92
	Percent of Sessions/IOA	57	67	48
Bob	IOA	95	96	98
	Percent of Sessions/IOA	43	43	43
Amy	IOA	97	97	97
	Percent of Sessions/IOA	43	43	43
Total	IOA Average	95	92	95
	Percent of Sessions/IOA	52	62	52

For the percent of steps observed with physical stereotypy (PST) on the TARP, paraprofessional ratings of PST were compared to ratings of a secondary observer who scored the TARP via video recording. IOA data for TARP PST were obtained for 62% of sessions. The average IOA was 92%, which was deemed acceptable. For participant specific TARP PST data please see Table 2.2.

For 6 second MTS IOA data, two secondary observers watched video recordings and simultaneously scored the occurrence or non-occurrence of physical stereotypy. A

total of 52% of sessions were compared for 6 second MTS IOA. Raters obtained an overall IOA average of 95%, which was deemed acceptable. For participant specific 6 second MTS data please see Table 2.

Social Validity

To ascertain whether the TARP is a socially valid and "teacher friendly" mechanism for collecting task analysis data and promoting skill acquisition, a survey was administered to the paraprofessionals who worked with the students. The paraprofessionals were asked about the degree to which (a) the skill taught to their student was important to their student, (b) the skill taught would be valuable in another setting, (c) the student can complete the skill independently, (d) the TARP yielded information useful to informing instructional practices, and (e) the TARP was an efficient and reliable tool for data collection. A comment box was also available to solicit further feedback. A five point Likert Scale was developed to facilitate responses. A copy of the survey administered to the paraprofessionals is included in Appendix B.

Experimental Design

A multiple probe design, which is a variation of a multiple baseline design, was utilized in order to assess the reliability and validity of the TARP. The multiple probe design was modified in that probe conditions only occurred during baseline and not throughout the entirety of the study. The modification was made in order to avoid prolonged exposure to the baseline condition, and conversely, increase exposure to instruction. While in baseline the paraprofessional only provided the prompt to begin the task without providing any instruction. A timer was set for two minutes and secondary

observers collected 6 second MTS data on PST. The rationale for the modification was to reduce the amount of wasted instructional time for students (i.e., asking them to complete an unknown task without offering instruction) and to prevent participant frustration. Once in the intervention phase, each participant was exposed to two instructional sessions daily.

Data were analyzed in three ways. The percentages of non-overlapping data points between baseline and intervention conditions were calculated to demonstrate the effect of the independent variables on skill acquisition and the co-varying reduction in physical stereotypy. Second, the accuracy of the TARP in recording physical stereotypy as compared to 6 second MTS was assessed via visual analysis of correspondence between graphed data. Finally, social validity data were collected through the use of a questionnaire and analyzed via visual analysis.

CHAPTER 3: RESULTS

Across the four participants in the current study, results indicated that the TARP was useful and accurate in measuring skill acquisition and concomitant reductions in physical stereotypy. In addition, results suggest that adequate correspondence was achieved on the TARP via paraprofessional-recorded physical stereotypy as compared to six-second MTS data recorded by secondary observers. Data depicting these results are organized by individual participants with respect to displaying increases in task independence along with co-varying reductions in physical stereotypy (both TARP and MTS measures). Data are also displayed on an individual participant basis depicting correspondence between TARP measures of physical stereotypy as compared to those obtained via six second MTS.

Figure 1 depicts the increases observed in skill acquisition and co-varying decreases in physical stereotypy for Mary. For Mary, data collected during baseline revealed very low levels of independence with the task, ranging from 0-8% of steps completed independently. Concurrently, she displayed significantly high rates of physical stereotypy as measured by 6 second MTS (95-100% of intervals). A visual analysis of improvements in independence and reduction of interfering physical stereotypy suggested strong co-variation. This is further supported by 100% percentage of non-overlapping data points for both independence and physical stereotypy.

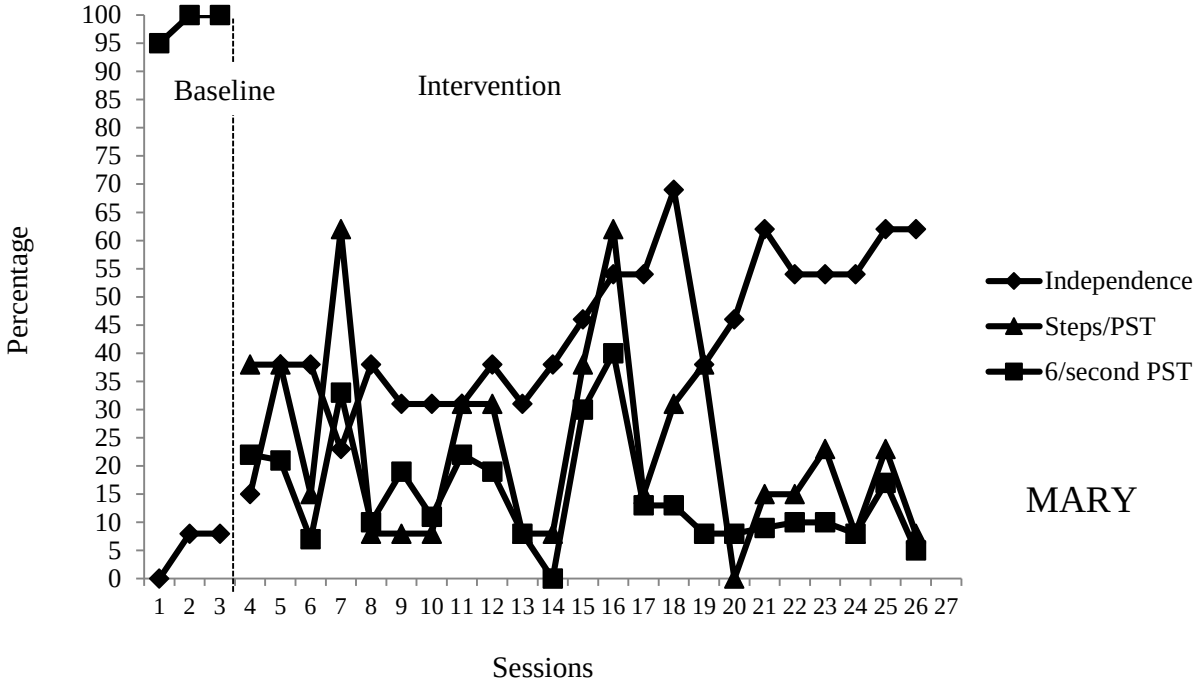


Figure 1: Independence and PST Data for Mary

Visual analysis of physical stereotypy data revealed acceptable correspondence between data collected by the paraprofessionals on the TARP and observers utilizing 6 second MTS. Please refer to Figure 2 for the data depicting correspondence data. Close inspection of the two data paths reveals that they are similar. Nonetheless, there were a few notable exceptions. For example, due to the lack of agreement between observers, the first physical stereotypy data point recorded in the intervention phase by a paraprofessional was deemed inaccurate. In order to boost recording accuracy, an additional training session was completed with the paraprofessional consisting of watching and scoring the video recording of the session. Upon video review, the paraprofessional scored the percentage of steps with stereotypy at 38% rather than the 92% initially scored. As another example, during session 20 the paraprofessional scored PST as occurring in 38% of the steps, while the secondary observer recorded PST as

occurring in 8% of intervals. On this occasion it appeared as though the 6 second procedure failed to capture occurrences of stereotypy due to the nature of the timing. An additional exception was noted in session 21 which showed zero occurrence of physical stereotypy from the perspective of the paraprofessional recording on the TARP. On this occasion it appeared that the paraprofessional was unable to see the occurrences of physical stereotypy due to the nature of the prompting procedure and body positioning.

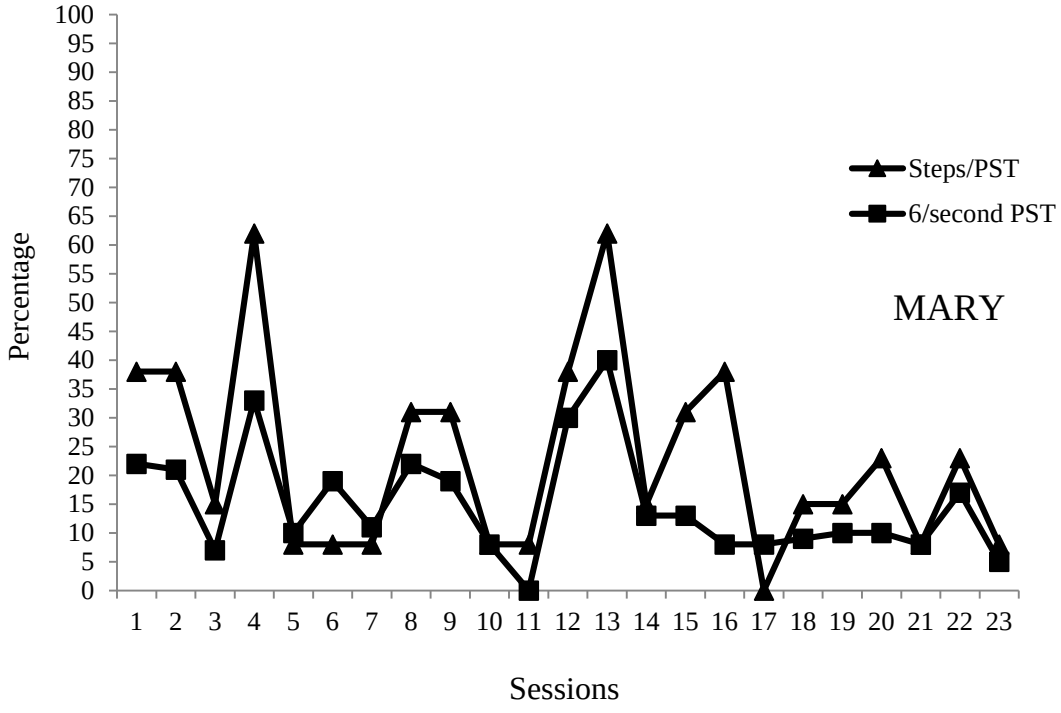


Figure 2: PST Correspondence for Mary

Figure 3 depicts the increases observed in skill acquisition and co-varying decreases in physical stereotypy for Tim. With respect to independence, Tim's percentage of non-overlapping data was 83%. The percentage of non-overlapping data with respect to reductions in physical stereotypy was 100%. Taken together, Tim's response to the intervention was robust.

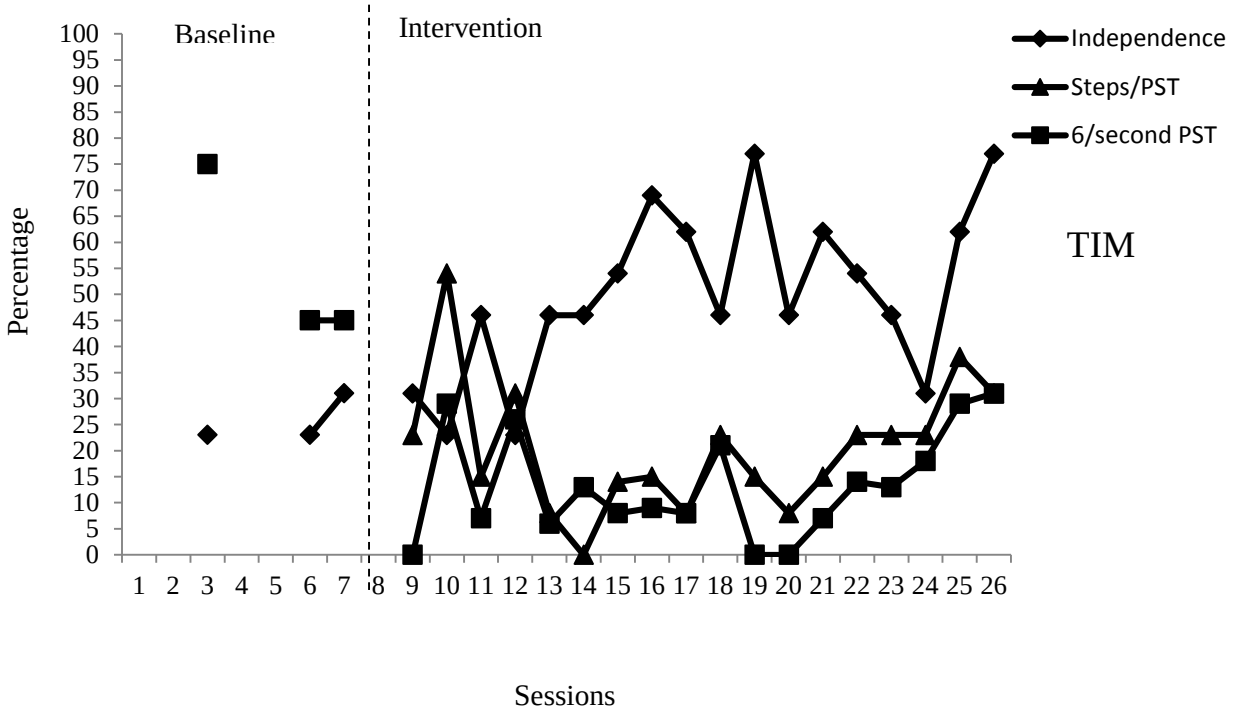


Figure 3: PST Correspondence for Tim

Visual analysis of physical stereotypy data for Tim revealed strong correspondence between data collected by the paraprofessionals on the TARP and that collected by secondary observers utilizing 6 second MTS (Figure 4). Notable about correspondence data collected about Tim was the overall similar pattern in trends. Further, after the moderate spread noted in the first two data points, the spread narrowed and remained tight throughout sessions. This may be indicative of improved fluency with the TARP and prompting procedure as the paraprofessionals gained experience with both.

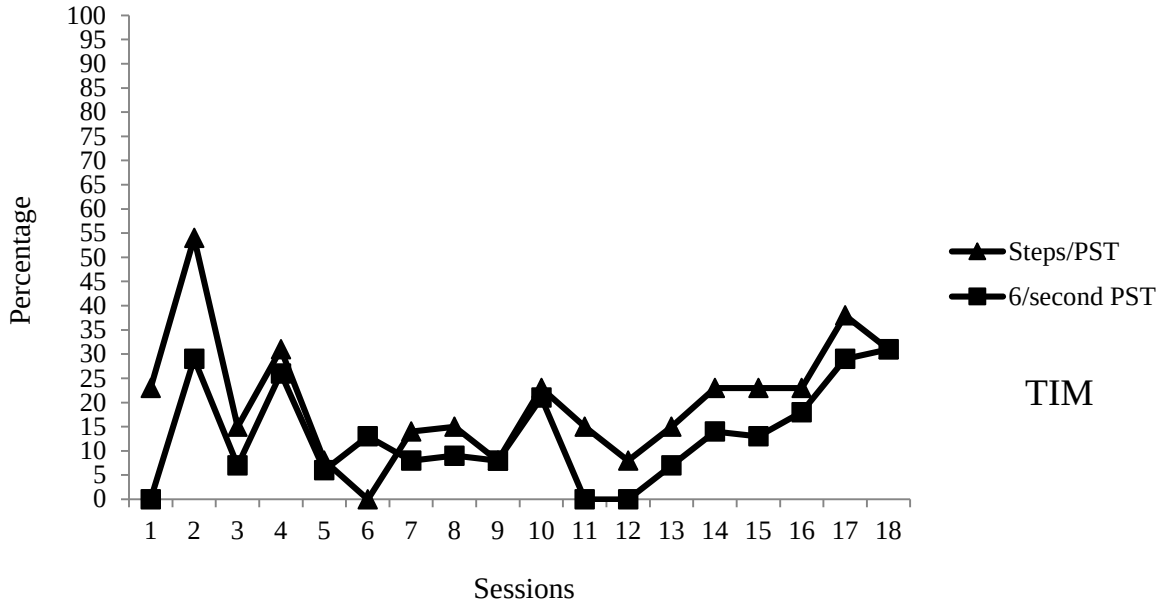


Figure 4: PST Correspondence for Tim

Figure 5 depicts the increases observed in skill acquisition and co-varying decreases in physical stereotypy for Bob. These baseline data revealed an absence of independence with the task, as he was unable to complete any steps on his own. Concurrently, he displayed significantly high rates of physical stereotypy as measured by 6 second momentary time sampling. A visual analysis of independence and reduction of interfering physical stereotypy revealed a dramatic response to intervention. This is further supported by 100% non-overlapping data points for both independence and physical stereotypy (PST).

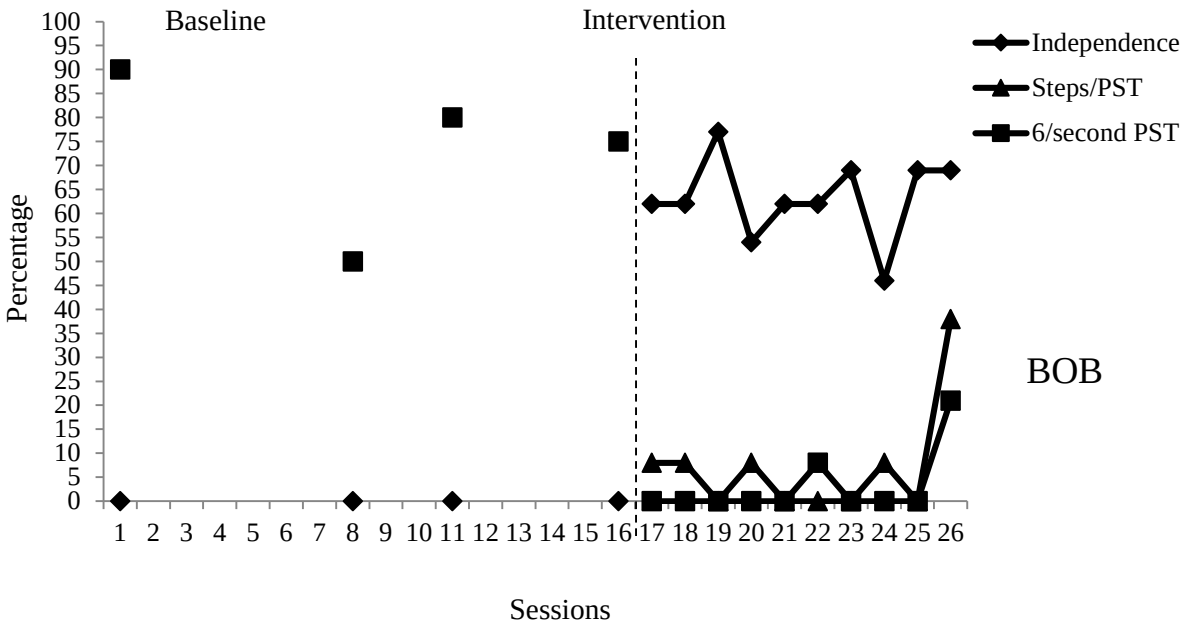


Figure 5: Independence and PST Data for Bob

Visual analysis of physical stereotypy data for Bob revealed strong correspondence between data collected by the paraprofessionals on the TARP and that collected by secondary observers utilizing 6 second MTS. Notable about correspondence data collected about Bob was the drop in PST to zero levels when using 6 second MTS. PST data collected through the TARP were one step (8%) of the task analysis. While PST may have occurred during specific steps, it did not coincide with MTS measurement. Nonetheless, these data still reflect low rates of interfering behavior. Figure 6 provides a visual depiction of Bob's PST correspondence data.

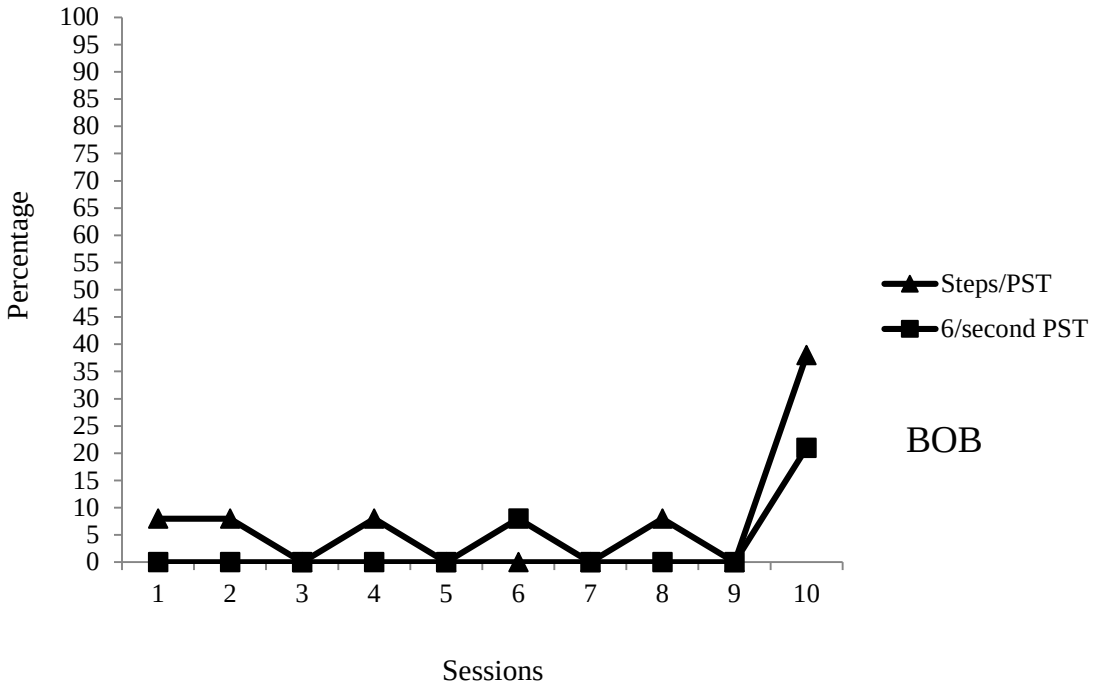


Figure 6: PST Correspondence for Bob

Figure 7 depicts the increases observed in skill acquisition and co-varying decreases in physical stereotypy for Amy. For Amy, data collected during baseline revealed low levels of independence with shoe tying. Simultaneously, she displayed significantly high rates of physical stereotypy as measured by MTS. A visual analysis of independence and reduction of interfering physical stereotypy revealed a significant response to intervention. This is further supported by a percentage of non-overlapping data points of 100% for both independence and physical stereotypy.

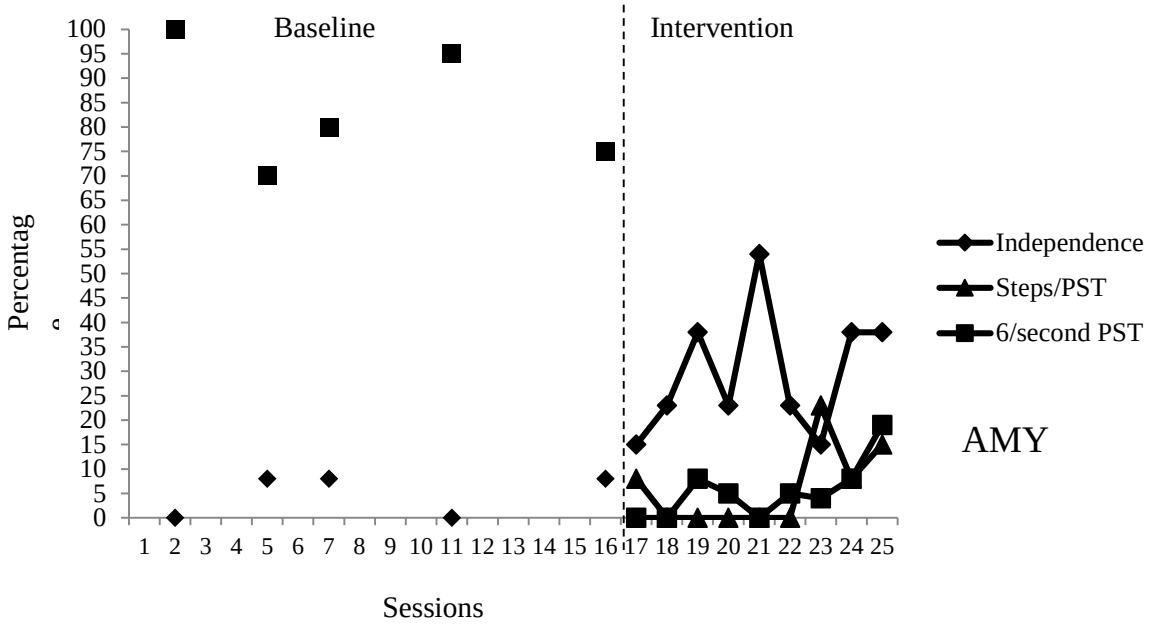


Figure 7: Independence and PST Data for Amy

Visual analysis of physical stereotypy data for Amy revealed strong correspondence between data collected by the paraprofessionals on the TARP and that collected by secondary observers utilizing 6 second MTS. An exception to this was observed in session 23. While the TARP PST procedure led to a score of 23%, MTS data scored it as occurring during 4% of intervals. Although the MTS method did detect the occurrence of PST, during this session the duration of the task was extended due to multiple errors (i.e., letting go of the shoelace). This served to drive down the percentage of PST due to the higher number intervals needed to complete the task. Figure 8 provides a visual depiction of Amy's PST correspondence data.

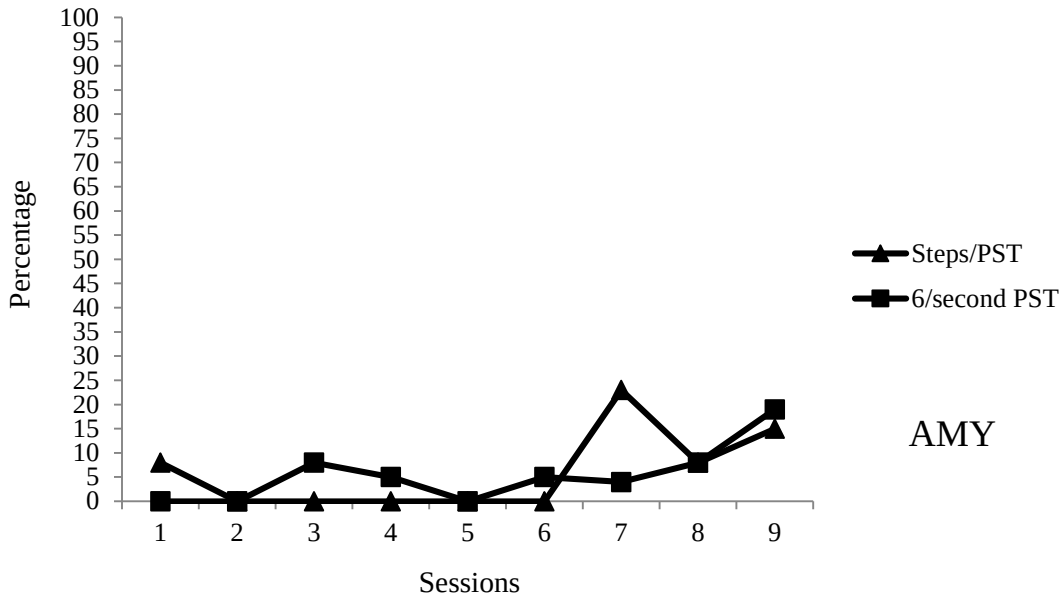


Figure 8: PST Correspondence for Amy

Figure 9 depicts time series data for all the participants. Due to the replication effects it appears that the observed change in participants' behavior was due to intervention and not to other variables. Additionally, data collected by paraprofessionals using the TARP corresponded strongly to six second momentary time sampling data collected by secondary observers. Taken together, the TARP proved to be a reliable data collection and intervention tool for all participants.

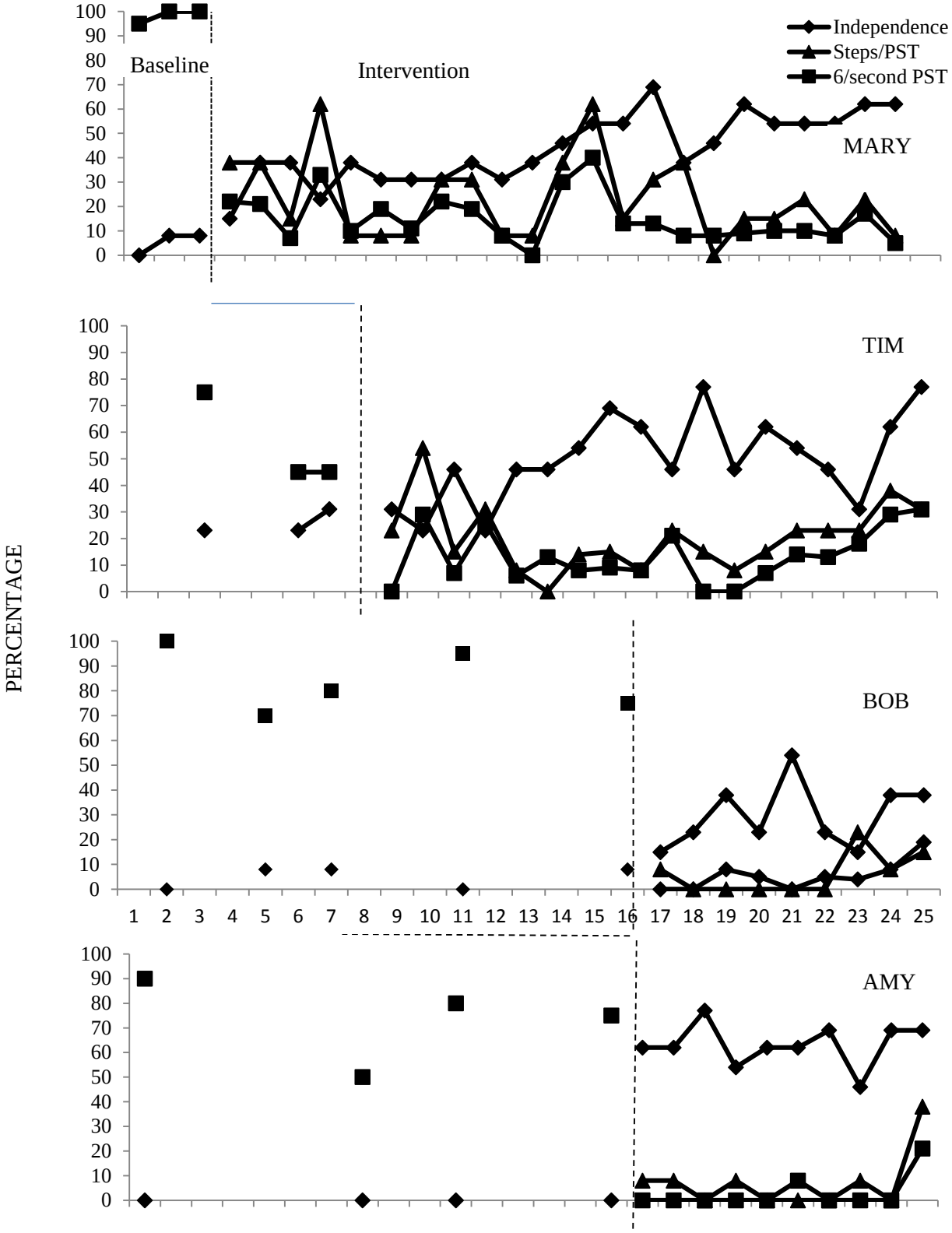


Figure 9: Graph depicting all Participants

Social Validity

At the conclusion of the study, each paraprofessional completed a questionnaire consisting of five questions. A five point Likert scale (1-5) was utilized to frame responses. A score of "5" represented "Strongly Agree" and a score of "1" represented "Strongly Disagree." Paraprofessionals were also asked to respond to the supplemental statement "The TARP is a teacher friendly data collection tool," utilizing the same Likert scale.

Table 3.1 displays the results of the social validity questionnaire. All responders either agreed or strongly agreed with the following statements:

- "The skill taught via the task analysis was important for my student."
- "The skill taught to my student would be valuable to my student in other settings."
- "The TARP yielded information useful to guiding instruction."
- "The TARP was an efficient and reliable tool for data collection."

Paraprofessionals provided variable responses to the statement "My student is able to complete the skill taught to him/her independently." This reflects the range of independent skill acquisition that the participants have demonstrated thus far.

The supplemental question was designed to probe the ease of use of the tool for the paraprofessionals. Two of the paraprofessionals indicated that the TARP was a "teacher friendly" tool by endorsing "Strongly Agree" and three others indicated

"Agree." One paraprofessional indicated a "Neutral" response and one paraprofessional endorsed "Disagree." Within the comment box, the paraprofessionals who were either neutral or disagreed with the statement commented that it was difficult to provide the prompting required while also collecting data.

Table 3.1

Social Validity Questionnaire

Questionnaire item	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
The skill taught via the task analysis was important for my student	7	0	0	0	0
The skill taught to my student would be valuable to my student in other settings (e.g. home, community).	6	1	0	0	0
My student is able to complete the skill taught to him/her independently	1	1	4	0	1
The TARP yielded information useful to guiding instruction.	6	1	0	0	0
The TARP was an efficient and reliable tool for data collection.	4	3	0	0	0
The TARP is a teacher friendly tool for data collection	2	3	1	1	0

CHAPTER 4: DISCUSSION

In applied settings, instruction is enhanced when decisions are informed by data. However, data need to be accurate and reliable in order to efficiently foster practices that promote growth and independence. Obtaining accurate and reliable data can be challenging in applied settings where resources are often limited and the needs of students are substantial. This is particularly evident in the context of task analysis instruction as often the only information obtained is the percentage of steps in the task analysis completed independently. While useful, percentage correct data often are not enough to adequately inform instruction (Haring & Kennedy, 1988). In other words, percentage data alone do not tell enough of the story. Additional information that is helpful includes at what step of the task analysis is instruction is disrupted by interfering behaviors. Steege & Watson (2009) described and illustrated the Task Analysis Recording Procedure (TARP) as a means to enhance task analysis data collection. They demonstrated that the TARP was useful and practical in documenting behavior change in terms of both skills acquisition (i.e., steps completed independently) and occurrences of interfering behavior (i.e., steps of the task analysis in which interfering behavior occurred). While numerous studies have demonstrated the reliability of skill acquisition data collection (e.g., Steege, Wacker, & McMahon, 1987), the validity of a step-based measure of stereotypy had not previously been examined.

This study contributes to the research on data-collection procedures by assessing the reliability and validity of the TARP in measuring physical stereotypy exhibited by four students with disabilities in the context of socially relevant functional life skills tasks. Results suggested that data collected by paraprofessionals using the TARP were

consistent with data obtained via six second MTS, which is considered a "gold standard" method of data collection for challenging behaviors (Steege, Davin, & Hathaway, 2001). This study also demonstrated that increases in independence co-varied with decreases in the display of physical stereotypy.

This study illustrated a method for efficiently documenting the decrease of interfering behaviors by increasing motorically incompatible replacement behaviors. Additionally, results are supportive of previous literature emphasizing the importance of teaching functionally-equivalent and socially meaningful replacement behaviors as a means of decreasing challenging behavior (Horner et al., 1990; Koegel, Koegel, & Dunlap, 1996; Steege & Watson, 2009). This study has direct implications for, and applications within, applied settings in which the concurrent measurement of skill acquisition and reductions of interfering behavior are required. Moreover, the TARP allows for an efficient and accurate measure of both sets of target behaviors. For example, when using the TARP, an instructor is able to record steps completed independently and steps in which interfering behavior occurred. Such data provide a way to measure directly response covariation between skill acquisition and interfering behaviors. Importantly, this study's findings document that such data recording can be done within the context of socially meaningful behaviors. The results also have direct implications for goal setting and progress monitoring by measuring interfering stereotypy directly within the context of functional life skills.

Baer, Wolf, and Risley (1968) described social significance in relation to behaviors that are meaningful, useful, and practical. The social validity survey data indicated that the paraprofessionals liked the TARP method and found it effective. With

social significance in mind, it is worth noting that the paraprofessionals continued to implement the TARP instructional methods and data recording procedures after the study ended. This speaks to the social validity of both the skills being taught and to the instructional strategies and recording procedures. Students who participated in the study increased their independence with skills that were useful and practical. Paraprofessionals demonstrated that the TARP was a reliable and valid method of recording challenging behavior in the context of functional life skills instruction. The TARP appears to be a tool that can be used to facilitate the acquisition of socially significant skills in applied settings.

Limitations and Future Research

A significant limitation of this study was that only physical stereotypy was addressed. The current data might not reflect TARP applications with other types of behaviors. Future research could extend this methodology to other topographies of behavior that might interfere with acquisition of functional life skills (e.g., self-injury, opposition, aggression, vocal stereotypy, among others). These studies would demonstrate the generality of the TARP in documenting response covariation between interfering behaviors and functional replacement behaviors.

Second, the current study employed a sample of paraprofessionals previously well versed in task analysis instruction. Despite previous experience, some difficulty in initial data collection was noted and required further training in order to obtain reliable data. Further, paraprofessionals expressed that it was challenging to utilize an intensive physical prompting procedure and collect data simultaneously. Paraprofessionals new to

instruction via task analysis likely would require enhanced support. Future research could examine pre-intervention TARP training procedures and modifications to promote ease of implementation.

Finally, while intervention sessions were run in a natural environment where the skills targeted for acquisition would be expected to be displayed, a threat to internal validity was present by having sessions videotaped by secondary observers. It is unclear how the presence of secondary observers and a video camera in the environment might have impacted the behavior of participants and paraprofessionals. It is conceivable that the obtained results were influenced by this environmental confound in some way. Future research could control for this in order to reduce the potential impact of secondary observers.

CHAPTER 5: SUMMARY

Students with Autism Spectrum Disorders (ASD) often display behaviors that interfere with skill acquisition. As such, instructional procedures often need to be modified in order to promote increases in skill and decreases in behaviors that impede acquisition. Decisions made around instructional modifications should be informed by data, with the quality of decisions being directly linked to the quality of the data obtained.

Task analysis data collection has been limited to percentage correct data in applied settings, which often does not promote a step wise analysis of potential inhibitors to skill acquisition. This study evaluated the reliability and validity of the Task Analysis Recording Procedure (TARP) in recording physical stereotypy, a behavior often emitted by students with ASD, by comparing TARP obtained physical stereotypy data to that obtained via six second momentary time sampling. The results indicated a robust correspondence between recordings of physical stereotypy conducted by teachers using the TARP and secondary observers utilizing a six second momentary time sampling procedure.

This study suggests that the TARP procedure is an acceptable means of recording physical stereotypy in applied settings as well as recording both the acquisition of skills and the decrease of interfering stereotypy within the context of functional life skills programming.

REFERENCES

Ahearn, W.A., Clark, K.M., MacDonald, Rebecca P.F., & Chung, B.I. (2007). Assessing and treating vocal stereotypy in children with autism. *Journal of Applied Behavior Analysis*, 40, 263-275. doi: 10.1901/jaba.2007.30-06.

Ahrens, E.N., Lerman, D.C., Kodak, T., Worsdell, A.S., & Keegan, C. (2011) Further evaluation of response interruption and redirection as treatment for stereotypy. *Journal of Applied Behavior Analysis*, 44, 95-108. doi: 10.1901/jaba.2011.44-95.

Baer, D.M., Wolf, M.M., & Risley, T.R. (1968). Some current dimensions of applied behavior Analysis. *Journal of Applied Behavior Analysis, 20,* 313-327.

Browder, D.M. & Spooner, F. (2011). *Teaching students with moderate and severe disabilities*. New York, NY: The Guilford Press.

Cooper, J.O., Heron, T.E., & Heward, W.L. (2007). *Applied behavior analysis (*2nd ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

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

Haring, T.G. & Kennedy, C.H. (1988). Units of analysis in task-analytic research. *Journal of Applied Behavior Analysis,* 21, 207-215.

Horner, D. R. & Keilitz, I. (1975). Training mentally retarded adolescents to brush their teeth. *Journal of Applied Behavior Analysis*, 8, 301-309.

Horner, R.H., Dunlap, G., Koegel, R.L., Carr, E.G., Sailor, W., Anderson, J., Albin, R.W. & O'Neill, R.E. Toward a technology of "nonaversive" behavioral support. *Journal of The Association for Persons with Severe Handicaps, 15, 125-132.*

Koegel, L.K., Koegel, R.L., & Dunlap, G. (Eds.) (1996) *Positive behavioral support: Including People with difficult behavior in the community.* Baltimore, Brookes.

McKay, J.A., Weiss, J.S., Dickson, C.A. & Ahearn, W.H. (2014). Comparison of prompting hierarchies on the acquisition of leisure and vocational skills. *Behavior Analysis in Practice,*7, 91-102. doi: 10.1007/s40617-014-0022-3.

Meany-Daboul, M.G., Roscoe, E.M., Bourret, J.C., & Ahearn, W.H. (2007). A comparison of momentary time sampling and partial interval recording for evaluating functional relations. *Journal of Applied Behavior Analysis*, 40, 501-514. doi: 10.1901/jaba.2007.40-501.

Radley, K.C., O'Handley, R.D., & Labrot, Z.C. (2015). A comparison of momentary time sampling and partial-interval recording for assessment of effects of social skills training. *Psychology in the Schools,* 52 (4), 363-378. doi: 10.1002/pits.21829.

Saini, V., Gregory, M.K., Uran, K.J., & Fantetti, M.A. (2015). Parametric analysis of response interruption and redirection as treatment for stereotypy. *Journal of Applied Behavior Analysis*, 48, 96-106. doi: 10.1002/jaba.186.

Seaver, J.L. & Bourret, J.C. (2014) An evaluation of response prompts for teaching behavior chains. *Journal of Applied Behavior Analysis,* 47, 777-792. doi: 10.1002/jaba.159.

Steege, M.W., Davin, T. & Hathaway, M. (2001). Reliability and accuracy of a performance-based behavioral recording procedure. *School Psychology Review,* 30, 252-261.

Steege, M.W., Wacker, D.P., & McMahon, C.M. (1987). Evaluation of the effectiveness and efficiency of two stimulus prompt strategies with severely handicapped students. *Journal of Applied Behavior Analysis,* 20, 293-299.

Steege, M.W. & Watson, T.S. (2009) *Conducting school-based functional behavioral assessments* (2nd ed.) . New York, NY: The Guilford Press.

Storey, K. & Miner, C. (2011). *Systematic instruction of functional skills for students and adults with disabilities.* Springfield, IL: Charles C Thomas.

APPENDIX A: SAMPLE TARP RECORDING FORM

Name: **001** Date:_____
Target Skill: Drink Teacher: _____

Steps of TA	Independence (+ or -)	Instructional Prompts	Interfering Behavior
1. Get a bottle of water		2secHOH	PST
2. Open bottle of water and put cap down		2secHOH	PST
3. Get a drink mix		2secHOH	PST
4. Open drink mix		2secHOH	PST
5. Pour drink mix into bottle		2secHOH	PST
6, Throw drink packet in trash		2secHOH	PST
7. Put cap on water bottle		2secHOH	PST
8. Shake water bottle		2secHOH	PST
9. Get a wipe		2secHOH	PST
10. Wipe table		2secHOH	PST
11. Throw away wipe		2secHOH	PST
12. Open bottle		2secHOH	PST
13. Take a drink		2secHOH	PST
.		2secHOH= 2 second delay then HOH	PST= Physical Stereotypy Number of steps with: PST=
	% Independence		% of steps with: PST=

Treatment Integrity

The teacher had all necessary materials? Y or N

Appropriate reinforcement was provided for steps in the chain completed correctly? Y or N

Prescribed response prompts were followed? Y or N

Steps	0/13	1/13	2/13	3/13	4/13	5/13	6/13	7/13	8/13	9/13	10/13	11/13	12/13	13/13
%	0	8	15	23	31	38	46	54	62	69	77	85	92	100

Physical Stereotypy: defined as occurrences of hand flapping, arm waving, banging hands or objects together, rubbing hands/arms up and down own body, rubbing objects, rubbing body on floors, walls or furniture.

APPENDIX B: SOCIAL VALIDITY QUESTIONNAIRE

Social Validity Teacher Survey

Please circle the choice that most closely represents your view.
Scale: 1= Strongly Disagree
2= Disagree
3= Neutral
4= Agree
5= Strongly Agree

1. The skill taught via the task analysis was important for my student.

1 2 3 4 5

2. The skill taught to my student would be valuable to my student in other settings (e.g. home, community).

1 2 3 4 5

3. My student is able to complete the skill taught to him/her independently.

1 2 3 4 5

4. The Task Analysis Recording Procedure (TARP) yielded information useful to guiding instruction.

1 2 3 4 5

5. The TARP was an efficient and reliable tool for data collection.

1 2 3 4 5

The TARP is a “teacher friendly” data collection tool. (Supplemental Question)

1 2 3 4 5

Comments:

BIOGRAPHY OF THE AUTHOR

Russell Brown was born in Portland, Maine. He grew up in Gray, Maine and graduated from Gray-New Gloucester High School in 1986. In 1990, he graduated from the University of Southern Maine (USM) with a B.A. in Communication. In 2006, he earned an M.S. in School Psychology at USM. In 2011, he became a Board Certified Behavior Analyst. Currently, Mr. Brown is completing his pre-doctoral internship at a special purpose school for children with disabilities. He is a candidate for the Psy. D. degree in School Psychology from the University of Southern Maine.