2012

A component analysis of a self-monitoring intervention for increasing appropriate behavior and decreasing automatically-reinforced problem behavior

Erin S. Leif
Western New England University

Running head: Self-monitoring

A Component Analysis of a Self-Monitoring Intervention for Increasing Appropriate Behavior and Decreasing Automatically-Reinforced Problem Behavior

Erin S. Leif

Western New England University &

The New England Center for Children

A dissertation submitted to the Department of Psychology in the School of Arts and Sciences at

Western New England University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Behavior Analysis

Supervised by:

Eileen M. Roscoe, PhD, BCBA-D

Acknowledgements

The completion of this dissertation has been one of the most significant academic and personal challenges I have ever had to face. Without the support, guidance, and patience of the following people, this study would not have been completed. To them, I owe my deepest gratitude.

To Eileen Roscoe: thank you for persevering with me as my advisor and mentor throughout the time it took me to complete this research and write this dissertation. You have provided valuable and constructive suggestions during all stages of this project. Your willingness to give your time so generously has been very much appreciated. To the members of my dissertation committee, Amanda Karsten, Gregory Hanley, and William Ahearn: you have generously given your time and expertise to better my work. Thank you for your contribution and your continued support. To Laura Wilhelm and Caleb Davis: thank you for offering your time and expertise to help with data collection. To the faculty of the psychology department at Western New England University: you have all provided me with a tremendous graduate education, and have truly taught me how to discover, translate, and apply knowledge toward solving human behavior problems of societal importance. To the New England Center for Children: you have provided me with an excellent atmosphere to conduct research, and I am grateful for your financial support. To Mom, Dad, Chad, and Jackie, my cheerleaders: thank you for always coming to my defense.

Abstract

Previous studies suggest that self-monitoring may be an effective treatment procedure. However, self-monitoring is typically included as one component of multicomponent interventions. Thus, it is unclear which component or combination of components is critical for the success of the intervention. We sought to extend previous self-monitoring research in a variety of ways. First, we evaluated the effects of video and in-vivo training as a procedure for teaching self-monitoring. Next, we conducted a component analysis of a self-monitoring intervention by sequentially adding successive components to determine their independent effects for increasing appropriate task engagement and decreasing automatically-reinforced stereotypy. We included self-monitoring accuracy as a dependent variable during and following self-monitoring training to determine the role of accurate self-monitoring on the other target behaviors. All participants learned to accurately and independently self-monitor their appropriate task engagement and off-task behavior during training. Results of the treatment component analysis indicated that self-monitoring plus differential reinforcement and differential reinforcement of alternative behavior (DRA) alone were equally effective components for two participants, suggesting that DRA was the critical component of the intervention. For the third participant, self-monitoring plus differential reinforcement was somewhat more effective than DRA alone. The implications of these findings for the continued use of self-monitoring interventions are discussed.

A Component Analysis of a Self-Monitoring Intervention for Increasing Appropriate Behavior and Decreasing Automatically Reinforced Problem Behavior

Intellectual disability (ID) is characterized by deficits in important life skills, such as communication, self-care, social interaction, vocational activities, and recreational activities (http://www.aaidd.org/content). Although individuals with ID can acquire and demonstrate a wide range of skills, caregiver support (e.g., prompts and programmed consequences) is often required. Therefore, independent performance of skills may not be achieved or may not maintain when external support is faded. The reliance on care providers for treatment outcomes may contribute to poor long-term outcomes for individuals with ID, including continued placement in restrictive environments (Howlin, Goode, Hutton, & Rutter, 2004). For this reason, it is important to identify interventions for promoting independent performance.

Self-monitoring has been proposed as a viable strategy for promoting independence in individuals with ID. As an independent variable, self-monitoring involves presenting materials for recording one's own behavior. Examples of these materials include paper with boxes to check (e.g., Koegel & Koegel, 1990), tokens (e.g., Newman, Tuntigian, Ryan, & Reinecke, 1997), or a tally counter (e.g., Koegel, Koegel, Hurley, & Frea, 1992). As a dependent variable, self-monitoring involves recording one's own behavior, regardless of whether the response is correct or incorrect. If it is correct, it involves the occurrence of a differential observing response (that is, a different response in the presence of each stimulus; Dube & McIlvane, 1999) that indicates discrimination of one's own behavior. Following the self-monitoring response, reinforcement may be delivered for appropriate behavior or for accurate self-monitoring.

Individuals with ID have been successfully taught to self-monitor, and self-monitoring interventions have been found effective in producing socially important outcomes (e.g., Koegel & Koegel, 1990; Stahmer & Schreibman, 1992). For example, self monitoring has been found effective in increasing a wide range of behaviors including time on-task (e.g., Blick & Test, 1987), vocational engagement and productivity (e.g., Ackerman & Shapiro, 1984), leisure item engagement (e.g., Stahmer & Schreibman, 1992), and appropriate social interactions (e.g., Sainato, Goldstein, & Strain, 1992), and for decreasing problem behavior, including stereotypy (e.g., Shabani, Wilder, & Flood, 2001), disruption (e.g., Shear & Shapiro, 1993), and self-injurious behavior (SIB; Tiger, Fisher, & Bouxsein, 2009).

Several authors have asserted that the inclusion of self-monitoring in behavioral interventions may have distinct advantages (e.g., Agran et al., 2005; Harrower & Dunlap, 2001; Koegel, Koegel, & McNerney, 2001). For example, Gilbert, Agran, Hughes, and Wehmeyer (2001) commented that teaching individuals with ID to self-monitor might facilitate behavior change by signaling the contingency in effect for responding. In other words, self-monitoring may occasion behavior that has previously been reinforced and reduce behavior that has previously resulted in extinction. Hughes et al. (2002) argued that self-monitoring might be effective because the procedure allows individuals to manage their own behavior rather than rely on others (e.g., parents, teachers, therapists) to direct and monitor their performance. Other authors have asserted that self-monitoring is easy to use (Agran et al., 2005), is cost- and time-effective (Koegel et al., 2001), and is preferred by the direct consumers of treatment (Ganz & Sigafoos, 2005).

Although numerous authors have discussed the potential utility of self-monitoring, there is little empirical support for these proposed advantages. In addition, in the self-monitoring research literature, there are several limitations that prevent conclusions that may be drawn regarding the independent efficacy of self-monitoring. One problem is that few authors provide information regarding the accuracy of self-monitoring during training and subsequent treatment. Although self-monitoring has been proposed as an important goal for individuals with ID, little information is available on the degree to which these individuals can correctly monitor their own behavior or what effects that monitoring has on the future occurrence of that behavior. In addition, some studies have reported positive treatment effects without arranging specific contingencies for accurate self-monitoring (e.g., Koegel & Frea, 1993; Shabani, Wilder, & Flood, 2001). Most self-monitoring studies do not provide data regarding accuracy of self-monitoring (e.g., Ganz & Sigafoos, 2005). When accuracy data are provided, authors commonly report only a mean value of self-monitoring accuracy (e.g., Koegel & Koegel, 1990; Shabani, Wilder, & Flood, 2001; Stahmer & Schreibman, 1992). For instance, Koegel and Koegel (1990) described a comprehensive training procedure for teaching participants with autism spectrum disorders (ASD) to self-monitor their stereotypy. Training included a combination of modeling, prompting, prompt fading, and reinforcement. Although this study was noteworthy in that their training procedure was comprehensive and clearly described, the authors did not provide data on the participants' acquisition of the self-monitoring response during training; they presented self-monitoring accuracy data only during the post-training self-monitoring intervention. Although self-monitoring accuracy data for the non-occurrence of stereotypy was high (M = 93%; range, 90% to 98%), accuracy data for the occurrence of stereotypy was low (M = 39%; range, 18% to 72%). Thus, the participants learned to accurately self-monitor the absence of their stereotypy, but not the occurrence of stereotypy. Although participants did not accurately self-monitor the occurrence of their stereotypy, the intervention resulted in large and sustained reductions in stereotypy. These findings suggest that some other component of the intervention was responsible for the observed effects. Although the authors noted that accuracy improved over the course of the study, accuracy data were not graphically displayed, preventing visual analysis of trends in this dependent variable.

Another self-monitoring study that reported accuracy data during intervention was conducted by Newman, Buffington, and Hemmes (1996), who evaluated self-monitoring for increasing appropriate conversation skills in individuals with ASD. During the first six sessions of the self-monitoring intervention, the therapist vocally prompted participants to take a token (exchangeable for preferred edibles and activities) following each appropriate conversational response. During subsequent sessions, prompts were removed and no programmed consequences were in effect for accurate self-monitoring. Although the self-monitoring intervention resulted in increases in appropriate conversational skills, accurate self-monitoring occurred at moderate and variable levels (M = 60.3%; range, 20% to 100%). Because self-monitoring accuracy was not correlated with improved performance, these findings suggest that another component was responsible for treatment effects. One possibility is that the presence of the tokens occasioned appropriate behavior. Additionally, the intermittent reinforcement schedule in effect during the self-monitoring condition may have produced inadvertent reinforcement (i.e., some appropriate behavior may have immediately preceded reinforcer delivery), facilitating maintenance.

Another limitation of research on self-monitoring is that multiple treatment components are in effect, limiting conclusions that can be drawn regarding the effects of only presenting self-monitoring materials. For example, self-monitoring interventions often include differential reinforcement of other behavior (DRO; e.g., Shabani, Wilder, & Flood, 2001), differential reinforcement of alternative behavior (DRA; e.g., Koegel & Frea, 1993), or differential reinforcement of accurate self-monitoring (DR Accurate; e.g., Frea & Hughes, 1997; Koegel &

Koegel, 1990). After demonstrating that a self-monitoring intervention was effective, Storey and Gaylord-Ross (1987) evaluated the independent and combined effects of treatment components for increasing positive social interactions of four high school students with ID. Specifically, the authors alternated three conditions, self-monitoring combined with DRA, noncontingent reinforcement (NCR) alone, and self-monitoring alone. Results showed that self-monitoring combined with DRA and self-monitoring alone were associated with high levels of positive statements. Therefore, these findings suggest that self-monitoring alone was effective

Although this study was noteworthy because the authors conducted a component analysis of self-monitoring, there were some limitations that deserve comment. First, the authors reported within-subject data for only a single participant. Second, the self-monitoring alone condition was only conducted one time and it followed the self-monitoring with DRA condition. Because the self-monitoring condition was not replicated and it followed a condition combined with DRA, it is possible that the effects observed were due to history effects. Therefore, it is unknown whether similar effects would have been obtained if the self-monitoring alone was conducted a second time and not immediately following a condition with DRA.

The most comprehensive component analysis of self-monitoring to date was conducted by Fritz, Iwata, Camp, Rolider, and Neidert (2012), who evaluated the effects of self-monitoring and other treatment components for three adults with ID who exhibited automatically reinforced stereotypy (vocalizations or head weaving). Following a no-interaction baseline condition, a therapist conducted self-monitoring training to teach participants to record a non-occurrence of their stereotypy. Training continued until participants demonstrated 90% accurate recording following the first or second vocal prompt. Next, the authors sequentially introduced various intervention components, including self-monitoring plus DRA (accurate), self-monitoring plus

DRA (accurate) plus DRO (stereotypy), and DRO alone. During the self-monitoring plus DRA (accurate) condition, the therapist presented the self-monitoring materials and delivered a preferred edible and praise contingent for accurate self-recording. During the self-monitoring plus DRA (accurate) plus DRO (stereotypy) condition, the therapist presented self-monitoring materials and delivered praise and a preferred item contingent on accurate self-recording *and* the absence of stereotypy. During the DRO only condition, the therapist did not present self-monitoring materials and delivered praise and a preferred item contingent on the absence of stereotypy. The component analysis yielded different findings for each participant. For one participant, self-monitoring plus DRA (accurate) resulted in reductions in stereotypy and high levels of self-monitoring accuracy. Because the self-monitoring materials included instructions to refrain from stereotypy (i.e., the words "sit still" were written on the self-monitoring data sheet), the authors conducted a control condition for only this participant. During this condition, the therapist presented an alternative activity (copying words), instructed the participant to refrain from emitting stereotypy, and did not present self-monitoring materials. Immediate and sustained decreases in stereotypy were observed, suggesting that the self-monitoring materials were effective due to instructional control and/or response competition. For the second participant, the self-monitoring plus DRA (accurate) condition resulted in low levels of stereotypy only when implemented a second time following the self-monitoring plus DRA (accurate) plus DRO condition; it did not produce low levels when it was first implemented following the no-interaction baseline. These findings suggest that a recent history of DRO may have been necessary for self-monitoring plus DRA (accurate) to be effective. For the third participant, self-monitoring plus DRA (accurate) did not suppress stereotypy before or after self-monitoring plus DRA (accurate) plus DRO. However, DRO alone was effective in reducing stereotypy to low levels. Results of this study suggest that self-monitoring plus DRA (accurate) may be unnecessary or ineffective when implemented independent of a DRO contingency. In addition, self-monitoring training took several hours for one participant.

In summary, the component analyses conducted by Storey and Gaylord-Ross (1987) and Fritz et al. (2012) yielded differential findings. Storey and Gaylord-Ross found that self-monitoring plus DRA and self-monitoring alone were equally effective in increasing appropriate behavior. However, some methodological limitations prevented a clear interpretation of these findings. By contrast, Fritz et al. found that self-monitoring was either ineffective or unnecessary for reducing stereotypy when implemented independent of a DRO contingency. In light of these conflicting findings, additional component analyses of self-monitoring interventions are warranted.

In addition to the need for systematic analyses of self-monitoring interventions' efficacy, there is also a need for evaluations of their social validity. Although studies have noted the potential advantages of self-monitoring interventions, few studies have systematically evaluated the treatment acceptability of self-monitoring among the recipients. Some studies used indirect methods to assess participant preference for self-monitoring interventions (e.g., Copeland, Hughes, Agran, Wehmeyer, and Fowler, 2002; Hildebrand, Martin, Furer, & Hazen, 1990; Hughes et al., 2002; Kaplan, Hemmes, Motz, & Rodriguez, 1996; Rae, Martin, & Smyk, 1990). For example, Copeland et al. (2002) assessed participants' perceptions of treatment and the outcomes through interviews. In general, the participants reported that they thought the intervention helped them meet their goals. Although the authors assessed social validity with treatment recipients, the use of interviews has been shown to yield limited reliability. In addition, it can only be used with individuals with ID who have sufficient vocal repertoires, limiting its generality.

After finding that a self-monitoring intervention effectively increased participants' task engagement, Kapadia and Fantuzzo (1988) assessed its acceptability by measuring participants' independent use of the self-monitoring materials and found that all participants independently self-monitored. Based on this outcome, the authors concluded that the participants found the intervention to be acceptable. However, it is possible that the participants independently used the materials because they had acquired stimulus control following exposure to self-monitoring plus DRA (i.e., the self-monitoring materials may have signaled the availability of reinforcement). Thus, the participants' use of self- monitoring materials may have been a result of their recent history of reinforcement rather than treatment acceptability. In addition, the participants may not have preferred the self-monitoring component of the intervention, but may have preferred the DRA component (during this phase, participants were permitted to select and self-deliver a reinforcer following each session, although no prompts or consequences were provided by the therapist). Because the authors did not report additional treatment preference data (e.g., answers to a questionnaire or differential selection of different intervention components), one could not conclude whether the participants preferred the self-monitoring intervention.

Given these limitations, we sought to extend previous self-monitoring research in a variety of ways. First, we included self-monitoring accuracy as a dependent variable during and following self-monitoring training to ensure that participants had acquired this skill and that it maintained, allowing us to determine the role of accurate self-monitoring on the other target behaviors. Second, we evaluated the effects of video and in-vivo training as a procedure for teaching self-monitoring. Third, we systematically replicated and extended Fritz et al. (2012) by sequentially adding successive components of a self-monitoring intervention to determine their independent effects for increasing appropriate task engagement and decreasing automatically reinforced stereotypy. Fourth, we assessed one participant's treatment acceptability of multiple self-monitoring components by conducting a systematic treatment preference assessment.

General Method

Participants and Setting

Three individuals participated in the current study. Chris was a 21-year-old male diagnosed with Fragile X Syndrome. He was non-vocal and communicated through the use of pictures, simple gestures, and manual signs. Chris lived at home and attended a day program for individuals with ID. Chris was ambulatory and could independently complete basic self-help skills, such as using the bathroom, dressing, and bathing. Caregivers reported that Chris frequently exhibited off-task or competing behavior (e.g., stereotypy or engagement with non-work materials) when instructed to work or play independently. Chris' stereotypy included wrist twisting, hand shaking, body rocking, and head shaking. Chris' individualized education plan (IEP) included a specific objective to increase appropriate vocational skills that may facilitate his entry into an employment program. For this reason, we included vocational tasks in his treatment analysis.

Bob was a 39-year-old male with a diagnosis of an ASD. Bob was non-vocal and primarily communicated through the use of a communication book consisting of written words. He lived in a group home for adults with developmental disabilities and completed several hours of community-based employment per day. Bob was ambulatory and could independently complete most basic self-help skills such as using the bathroom, dressing, and bathing. Caregivers reported that Bob frequently exhibited off-task behavior when instructed to complete tasks and required additional prompts to complete familiar tasks. We included leisure, domestic, and vocational tasks in his treatment analysis.

Scott was a 14-year old male with a diagnosis of an ASD. Scott primarily communicated through the use of vocal approximations and gestures. Scott lived in a residential facility for individuals with an ASD and attended the same day program as Chris. Scott was ambulatory and could independently complete basic self-help skills, such as using the bathroom, dressing, and bathing. Caregivers reported that Scott frequently engaged in off-task or competing behavior (e.g., motor stereotypy) when instructed to work or play independently. Motor stereotypy consisted of hand flapping, body rocking, ear holding, nose picking, and finger mouthing. Scott's IEP included objectives for increasing independent leisure item engagement and the completion of simple prevocational tasks. For this reason, we included both leisure and vocational tasks in the treatment analysis.

All sessions occurred in a quiet area of participants' classrooms located within the school building (Chris and Scott) or residence (Bob). Each area contained a table, chairs, and relevant task materials. Sessions were conducted once per day for approximately 30 min, three to five times per week. All sessions were videotaped, and observers subsequently collected data from the previously recorded video footage.

Materials

We identified six tasks for use in the task assessment that (a) the participant was frequently instructed to complete, (b) were available in the participant's school or home, (c) could be placed on the table in front of the participant, and (d) resulted in a series of discrete permanent products. Table 1 lists the tasks that were included and their operational definitions.

Response Measurement and Reliability

The dependent variables for Chris, Bob, and Scott were appropriate task engagement, productivity, and self-monitoring accuracy. In addition, we measured motor stereotypy for Chris and Scott. Observers collected data using pencil and paper recording or the Direct Assessment Tracking Application (DATA) on a handheld iTouch device. *Appropriate task engagement* was defined individually for each task (as noted in Table 1). We measured appropriate task engagement using 10-s momentary time sampling (MTS) and summarized data as percentage of intervals. *Productivity* was defined individually for each task (as noted in Table 1). Observers measured productivity using frequency recording and summarized data as responses per minute (rpm). To calculate *self-monitoring accuracy*, observers compared the experimenter's record with the participant's data sheet following each self-monitoring opportunity to determine whether there was an agreement (both recorded an occurrence or nonoccurrence of appropriate behavior), and a percentage accuracy measure was reported. *Motor stereotypy* was measured using 10-s MTS and we summarized the data as percentage of intervals (see Table 2 for response definitions).

A second independent observer collected data on all dependent variables for at least 25% of sessions in each condition. Interobserver agreement was collected using paper and pencil recording or a handheld iTouch device. Interobserver agreement was calculated by dividing the number of agreements by the number of agreements and disagreements and multiplying the quotient by 100 to yield the percentage agreement. Interobserver agreement was reported for each participant as the overall mean agreement across each analysis for each dependent variable (as noted in Table 3).

Task Assessment

Procedure

A task assessment was conducted to identify tasks for inclusion in the treatment assessment. For each participant, we included six tasks that could be independently engaged with for an extended duration (at least 5 min). Because we were interested in measuring appropriate task engagement (e.g., participation in an activity in a manner consistent with its intended function) rather than simple hand-to-item contact, we included tasks that resulted in observable and measurable permanent products.

Task Assessment - Phase 1. The experimenter singly presented each task during 10 successive trials to assess whether the participant could complete the task independently. The same task was presented across 10 consecutive trials. During each trial, the experimenter presented the task materials (e.g., one bead and a string) and the instruction "Do X." If the participant completed the task correctly and independently, the experimenter delivered brief praise and a small edible. If the participant completed the task incorrectly or did not respond within 5 s, then the experimenter removed the materials and represented them for the next trial. Tasks that were independently completed for 90% of trials were included in Task Assessment – Phase 2.

Task Assessment - Phase 2. The experimenter presented each task singly during 5-min sessions and assessed each task three times (task presentation was alternated across sessions). At the start of each session, the experimenter presented the task materials and the instruction "You can do X if you want." The experimenter then turned away from the participant, diverted her attention, and pretended to act busy for the remainder of the session. No programmed consequences were delivered for appropriate task engagement, productivity, or problem behavior. A task associated with relatively low levels of appropriate task engagement and productivity was selected for inclusion in the treatment analysis.

Task Assessment Results

Six tasks associated with at least 90% independent completion were identified for each participant during Phase 1 of the task assessment (Figure 1) and were included in Phase 2. Results of Phase 2 of the task assessment are depicted in Figure 2. For Chris, three tasks (i.e., stuffing envelopes, beading, and sorting money) were associated with low levels of appropriate task engagement (i.e., less than 50%), low rates of productivity (less than 5 RPM), and moderate levels of motor stereotypy. We selected the beading task because it was associated with lower levels of appropriate task engagement than stuffing envelopes, and lower levels of productivity than sorting money. For Bob, weight lifting and stamping were associated with low levels of appropriate task engagement (less than 50%) and productivity (less than 5 RPM). Weight lifting was included in his treatment evaluation because it was associated with the lowest levels of appropriate task engagement and productivity. For Scott, four tasks (i.e., beading, building Legos, and completing math facts) were associated with low levels of appropriate task engagement (less than 50%) and moderate levels of motor stereotypy. However, filing cards, building Legos, and completing math facts were associated with moderate levels of productivity. We selected beading for use in Scott's treatment evaluation because it was associated with low levels of appropriate task engagement and productivity, and the highest levels of stereotypy.

Functional Analysis

Procedure

Prior to the treatment assessment, a functional analysis based on the procedures described by Iwata, Dorsey, Slifer, Bauman, and Richman (1982/1994) was conducted for Chris and Scott to identify the variable(s) maintaining their problem behavior. The functional analysis included three test conditions and one control condition. Sessions were 5 min, and we rapidly alternated conditions using a multielement design.

Attention. The purpose of this condition was to determine whether problem behavior was maintained by social positive reinforcement in the form of attention. The experimenter sat at a table with the participant, stated, “Sit here. I need to do some work,” and then pretended to read over some paperwork while remaining at the table. Contingent on the occurrence of problem behavior, the experimenter immediately delivered attention by stating, “Stop that” or, “That looks funny,” paired with brief physical contact. For non-discrete occurrences of stereotypy that extended in duration, the experimenter delivered attention every 5 s until the participant stopped engaging in the response for at least 1 s. The experimenter ignored the participant’s non-target behavior.

Escape. The purpose of this condition was to determine whether behavior was maintained by social negative reinforcement in the form of escape from demands. Based on caregiver report, five tasks that the participant completed independently but did not readily comply with were presented within each session. The experimenter sat at the table with the participant and continuously presented tasks using a three-step prompting hierarchy (i.e., vocal, model, then physical). If the participant did not comply with a demand within 5 s of the initial vocal prompt, then the experimenter represented the vocal prompt with a model prompt. If the participant did not comply with the demand within 5 s of the model prompt, then the experimenter manually guided the participant to emit the correct response. The experimenter delivered praise contingent on correct responses that occurred before the physical prompt. Contingent on the occurrence of problem behavior, the experimenter removed all materials and turned away from the participant for 15 s. At the end of the 15-s escape interval, the experimenter presented the next demand. If problem behavior occurred during the escape interval, the experimenter did not delay initiation of the next demand presentation. The experimenter ignored the participant's non-target behavior.

Ignore. The purpose of this condition was to test whether behavior was maintained by automatic reinforcement. The experimenter sat at a table with the participant, stated, "Sit here. I need to do some work" and then pretended to read over some paperwork while remaining at the table. The experimenter ignored all participant behavior.

Control. The experimenter sat at a table with the participant and presented a variety of leisure items. The experimenter delivered attention ("Nice playing!") paired with brief physical contact every 15 s. If the participant emitted problem behavior immediately prior to the next scheduled delivery of praise, the experimenter delayed attention delivery by 5 s. No demands were presented to the participant. If the participant initiated interaction with the experimenter, she responded and played with the participant for 5 to 10 s. The experimenter ignored the participant's problem behavior

Extended Ignore. We conducted consecutive ignore sessions (as described by Vollmer, Marcus, Ringdahl, & Roane, 1995) to rule out maintenance by social reinforcement. If problem behavior maintained during this phase, this provided additional support that the behavior was maintained by automatic reinforcement.

Functional Analysis Results

Results of the functional analysis are depicted in Figure 3. Chris (top panel) exhibited moderate, undifferentiated levels of motor stereotypy across all conditions during the multielement functional analysis. During an extended series of no-interaction sessions, Chris' motor stereotypy maintained with no downward trend, suggesting that his stereotypy was maintained by automatic reinforcement.

Scott (bottom panel) exhibited differentially higher levels of motor stereotypy during the no-interaction condition. However, he also exhibited motor stereotypy at variable levels in the other conditions. During an extended no-interaction phase, Scott's motor stereotypy also maintained with no downward trend, suggesting that his stereotypy was maintained by automatic reinforcement.

Treatment Component Analysis

Procedure

During the treatment component analysis, we assessed the effects of self-monitoring intervention components by sequentially adding them across successive phases until clinically acceptable increases in appropriate behavior were observed. Next, we withdrew and reintroduced the effective treatment phase by alternating it with a previously identified ineffective treatment phase, using a reversal design to demonstrate experimental control. Treatment analysis phases included (a) self-monitoring materials, (b) self-monitoring materials plus DRA for accurate self-monitoring, and (c) self-monitoring materials plus DRA for accurate self-monitoring *and* appropriate task engagement. A self-monitoring materials condition was conducted prior to self-monitoring training to assess whether the materials alone exerted stimulus control prior to training. In addition, we evaluated the effects of DRA (without self-monitoring) for appropriate task engagement. Sessions were 5 min.

During all self-monitoring conditions (including training), the experimenter presented the self-monitoring materials at the start of the session. Self-monitoring materials included a data sheet (see Appendix A) and a pencil. The data sheet included two columns and column headers.

The column header on the left depicted a photo of the participant engaging in the task appropriately and the text “YES.” The column header on the right depicted a photo of the participant not engaging with the task and the text “NO.” There were five self-recording boxes under each column header, and each column of boxes was a different color to enhance discrimination across box types. The experimenter played a re-recording of a tone, which was delivered according to a variable-momentary (VM) 30-s schedule.

Self-monitoring materials. At the start of each session, the experimenter presented the instruction, “I have some work for you to do. You can work and check boxes if you want, but you’re not earning anything.” The experimenter did not deliver any additional instructions or prompts nor did she deliver programmed consequences for accurate self-monitoring or appropriate task engagement. The purpose of this condition was to assess the effects of only presenting self-monitoring materials on appropriate task engagement, productivity, self-monitoring accuracy, and stereotypy.

Self-monitoring training. Self-monitoring training was conducted for each participant following the first implementation of the self-monitoring materials condition. A multiple baseline across participants design was used to demonstrate experimental control.

Video self-monitoring. The experimenter conducted three types of video training sessions, occurrence, nonoccurrence, and interspersed occurrence and nonoccurrence. Before each training session, the experimenter read a script of instructions (see below) to the participant.

I have a video for us to watch. At the end of the video, you will need to check a box. If you saw yourself working in the video, you will need to check the “YES” box (the experimenter pointed to the picture of the participant working appropriately on the self-monitoring sheet). And if you saw yourself NOT working in the video, you will need to check the “NO” box (the experimenter *pointed to the picture of the participant not working on the self-monitoring sheet). If you check the correct box, you will get a (name of preferred edible).*

Training sessions consisted of ten trials, and the experimenter played a video clip during each trial. Each 10-s video clip depicted video footage that was captured during the participant's task assessment. During occurrence sessions, the experimenter played the same video clip on every trial, and the clip depicted the participant correctly completing the task. During nonoccurrence sessions for Chris and Scott, the experimenter played one of three video clips of the participant engaging in off-task behavior (i.e., one clip depicted the participant sitting still and doing nothing, one clip depicted the participant engaging in motor stereotypy, and one clip depicted the participant manipulating non-task related items). The same clip was played during all trials within a session, and the three different off-task behavior clips were rotated across sessions. For Bob, only one video clip of him engaging in off-task behavior (e.g., sitting still and doing nothing) was shown. During interspersed occurrence and nonoccurrence sessions, the experimenter randomly rotated the video clips such that on-task and off-task clips were each depicted for five trials. At the end of each clip, a timer sounded, and the experimenter vocally prompted the participant to place a check in the correct box. In addition, the experimenter pointed to the correct box. Contingent on correct self-monitoring, the experimenter delivered praise and a small edible. Vocal and point prompts were systematically faded in 2-s increments following two consecutive sessions with 80% accuracy. If the participant incorrectly self-monitored (i.e., checked the incorrect box), the experimenter erased the error, turned away from the participant for 3 s, and then prompted the correct response. The experimenter did not deliver praise or an edible following the prompted correct response, and the experimenter initiated the next trial. Self-monitoring training continued until the participant met a criterion of 80% accuracy and independence across two consecutive trials.

In vivo self-monitoring. An interspersed occurrence and nonoccurrence in-vivo training session was conducted to ensure that the participant could demonstrate self-monitoring of his behavior as it occurred in real time. Before each training session, the experimenter read a script to the participant similar to the following:

I have some work for you to do. When you hear the timer beep, you will always need to check a box. When the timer beeps and you are working (the experimenter demonstrates working with the task materials), you will need to check the "YES" box (the experimenter pointed to the picture of the participant working appropriately on the self-monitoring sheet). When the timer beeps and you are NOT working (the experimenter demonstrated sitting still and not engaging with the task materials), you will need to check the "NO" box (the experimenter pointed to the picture of the participant not working on the self-monitoring sheet). If you check the correct box, you will get a (name of preferred edible).

Training sessions consisted of ten trials. During each trial, the experimenter sounded the tone when the participant was exhibiting (a) an occurrence of appropriate task engagement or (b) a nonoccurrence (i.e., off-task behavior). When the tone sounded, the experimenter presented a vocal and point prompt (to the correct box) for the participant to self-monitor whether or not he was exhibiting appropriate task engagement or off-task behavior at that moment. The experimenter conducted five occurrence and five nonoccurrence trials, randomly rotating between them in a 1:1 ratio when possible. The experimenter delivered praise and a small edible for correct self-monitoring. Vocal and point prompts were systematically faded in 2-s increments following two consecutive sessions with 80% accuracy. If the participant incorrectly self-monitored (i.e., checked the incorrect box), the experimenter erased the error, turned away from the participant for 3 s, and then prompted the correct response. The experimenter did not deliver praise or an edible following the prompted correct response, and the experimenter initiated the next trial. The criterion for the participant to complete self-monitoring training was 80% accuracy and independence across two consecutive trials.

Self-monitoring materials + DRA (accurate). This phase was identical to the self-monitoring materials condition, except for the addition of the DRA (accurate) component. That is, the experimenter delivered praise and a preferred edible or token contingent on the participant's accurate self-monitoring at the sound of the tone. At the start of each session, the experimenter presented the instruction, "I have some work for you to do. When you hear the beep, if you check the correct box, you can have a (name of edible or token)." The purpose of this condition was to evaluate the effects of adding DRA for accurate self-monitoring to self-monitoring materials on participants' appropriate task engagement, productivity, self-monitoring accuracy, and stereotypy.

Self-monitoring materials + DRA (accurate & engagement). This phase was identical to the self-monitoring materials + DRA (accurate) phase, except for the addition of the DRA (engagement) component. That is, the experimenter delivered praise and a preferred edible or token only if the participant exhibited both accurate self-monitoring *and* appropriate task engagement at the sound of the tone. At the start of each session, the experimenter presented the instruction, "I have some work for you to do. When you hear the beep, if you are working and you check the YES box, you can have a (name of edible or token)." The purpose of this condition was to evaluate the enhancing effects of adding DRA for task engagement to the self-monitoring materials plus DRA (accurate) phase on participants' appropriate task engagement, productivity, self-monitoring accuracy, and stereotypy.

DRA (engagement). At the start of each session, the experimenter presented the instruction, "I have some work for you to do. When you hear the beep, if you are working, you can have a (name of edible or token)." The tone sounded according to the same VM 30 s schedule in effect during the preceding self-monitoring materials + DRA (accurate & engagement) condition. However, the self-monitoring materials were not presented to the participant. If the participant was exhibiting appropriate task engagement when the tone sounded the experimenter delivered a preferred edible or token. The purpose of this condition was to evaluate the independent effects of DRA on appropriate task engagement, productivity, and stereotypy.

Results

Figure 4 depicts results of the treatment component analysis for Chris. During the first implementation of the self-monitoring materials condition (prior to self-monitoring training), Chris' appropriate task engagement (top panel) initially occurred at high levels, but rapidly decreased. In addition, an increasing trend in stereotypy was observed. Chris displayed moderate levels of productivity (middle panel). However, he never exhibited accurate self-monitoring (bottom panel), suggesting that he did not have this skill in his repertoire prior to training.

Figure 5 (top panel) depicts the results of self-monitoring training for Chris. Chris quickly acquired the self-monitoring response when video clips depicted only appropriate task engagement or off-task behavior on each trial. However, Chris' performance became more variable when the occurrence and nonoccurrence video clips were interspersed and he was required to make a conditional discrimination on each trial. We observed that Chris frequently checked boxes on the self-monitoring worksheet out of order, rather than sequentially. Therefore, we included a visual aid in the form of a cardboard cutout window that could be placed over the row of boxes that corresponded with each trial. During the next trial, the window could be slid down to the next row. Following this modification, Chris acquired the self-monitoring response during interspersed video trials. Given his previously variable performance, we conducted more sessions than required by our pre-determined mastery criterion to ensure that Chris had this response at strength. The visual aid remained in place for the remainder of the study. During in-vivo training, Chris initially exhibited only off-task behavior, preventing us from interspersing both occurrence and nonoccurrence trials. Thus, we began in-vivo training by conducting only nonoccurrence trials. After Chris demonstrated independent and accurate self-monitoring of his off-task behavior, we initiated occurrence trials by presenting an instruction to "string the bead" when the task materials were presented. Upon introduction of this instruction, Chris immediately displayed appropriate task engagement and rapidly acquired the self-monitoring response during occurrence trials. Finally, we conducted a series of interspersed in-vivo training trials. The experimenter placed the task materials on the table and did not deliver supplemental instructions because Chris independently exhibited appropriate task engagement during this phase. During this phase, Chris rapidly acquired the self-monitoring response. Self-monitoring training took approximately 4.25 hours for Chris.

During the second implementation of the self-monitoring materials condition (following self-monitoring training), Chris did not exhibit appropriate task engagement or productivity and displayed moderate levels of stereotypy. In addition, Chris did not exhibit accurate self-monitoring, suggesting that this behavior did not occur when only the materials were presented.

During the first self-monitoring materials + DRA (accurate) phase, levels of appropriate task engagement and productivity did not change from the previous phase, although stereotypy decreased slightly. However, Chris' self-monitoring accuracy increased to high levels. Chris was off-task on each self-monitoring opportunity and accurately self-monitored that he was off-task. Thus, including consequences for accurate self-monitoring increased this behavior but did not facilitate increases in the self-monitored target behavior.

During self-monitoring materials + DRA (accuracy & task engagement), Chris' appropriate task engagement and productivity initially did not change relative to the preceding phase, and stereotypy was high and variable. In addition, self-monitoring accuracy was variable and then decreased to low levels. For this reason, we introduced a pre-session exposure prompt prior to the start of the session to ensure contact with the reinforcer in effect for appropriate engagement and accurate self-monitoring. Prior to the start of the session, we vocally prompted Chris to "string the bead," sounded the timer, and instructed him to "check the YES box." We then delivered an edible for appropriate task engagement and accurate self-monitoring. From sessions 21 to 25, when the pre-session prompt was conducted, Chris' appropriate task engagement and productivity increased, stereotypy decreased, and self-monitoring accuracy increased. When the pre-session prompt was withdrawn during sessions 26 to 30, Chris' appropriate task engagement and productivity decreased, stereotypy remained low, and self-monitoring accuracy was variable. We re-introduced the pre-session prompt (session 31 to session 37), and Chris' appropriate engagement and productivity increased. However, his levels were more variable and self-monitoring accuracy was low. During this time, we observed that Chris frequently refused the edible when it was delivered contingent on appropriate task engagement and accurate self-monitoring, suggesting that it no longer functioned as a reinforcer. Because Chris typically earned tokens that could be exchanged for computer time in the classroom, we delivered tokens instead of edibles from session 38 to the end of his treatment analysis. Each token could be exchanged for 30 s of computer time, and a maximum of 10 tokens or 5 min of computer time could be earned per session. Following this modification, appropriate task engagement, productivity, and self-monitoring accuracy increased to high levels, and stereotypy decreased to low levels.

During the second implementation of self-monitoring materials + DRA (accurate), appropriate task engagement, productivity, and self-monitoring accuracy remained high and stereotypy remained low, suggesting that the self-monitoring materials + DRA (accurate) intervention may be effective following an immediate history of self-monitoring materials + DRA (accurate & task engagement). When reimplementing the self-monitoring materials condition, low levels of appropriate task engagement, productivity, and self-monitoring accuracy and moderate levels of stereotypy occurred. During the third implementation of self-monitoring materials + DRA (accurate), high levels of appropriate task engagement, productivity and self-monitoring accuracy and low levels of stereotypy were observed. Thus, these effects were similar to those observed in the second self-monitoring materials + DRA (accurate) phase, failing to replicate our initial finding that this phase may be ineffective following a self-monitoring materials phase. Therefore, the self-monitoring materials + DRA (accurate) phase produced increases in appropriate task engagement irrespective of whether it followed self-monitoring materials or self-monitoring DRA (accurate & task engagement) conditions. We hypothesized that perhaps the presence of the self-monitoring materials or the tokens, both of which were presented to the participant during the self-monitoring + DRA (accurate) phase, may have exerted stimulus control over appropriate behavior and facilitated response maintenance.

Next, we evaluated the effects of DRA (engagement) without self-monitoring by altering this condition with the self-monitoring materials condition. During both replications of DRA (engagement), appropriate task engagement and productivity persisted, but levels were not as high as those observed during the previous self-monitoring + DRA (accurate) phase. The slight decrements in performance may have been due to the withdrawal of the self-monitoring materials, which may have functioned as a discriminative stimulus. During the self-monitoring materials phase, in which the self-monitoring materials were available but the reinforcement contingency was withdrawn, levels of engagement and productivity initially maintained, but subsequently decreased to zero levels, and an increasing trend in stereotypy was observed. In addition, Chris did not self-monitor his behavior on any occasion.

Finally, we implemented self-monitoring materials + DRA (accurate) once again, and appropriate task engagement, productivity, and self-monitoring accuracy increased. Stereotypy decreased but was somewhat variable. These data indicated that both DRA (engagement) and self-monitoring + DRA (accurate) were effective, whereas self-monitoring materials alone was not. Although self-monitoring + DRA (accurate) resulted is somewhat higher levels of appropriate task engagement, it's efficacy may have been dependent on the preceding (albeit not immediate) history with DRA (accurate & engagement).

Figure 6 depicts results of the treatment component analysis for Bob. During the first implementation of the self-monitoring materials condition (prior to self-monitoring training), Bob's appropriate task engagement (top panel) and productivity (middle panel) occurred at low levels. These findings suggested that self-monitoring materials had no effect on appropriate task engagement prior to training. In addition, Bob never exhibited accurate self-monitoring (bottom panel), suggesting that he did not have this skill in his repertoire prior to training.

Figure 5 (middle panel) depicts the results of self-monitoring training for Bob. Similar to Chris, Bob quickly acquired the self-monitoring response when video clips depicted only appropriate task engagement or off-task behavior on each trial. Bob's performance also became variable when the when the occurrence and nonoccurrence video clips were interspersed. Similar to Chris, Bob checked boxes on the self-monitoring worksheet out of order, rather than sequentially. Therefore, we introduced an identical visual aid to that used with Chris for Bob. Following the introduction of the visual aid, Bob acquired the self-monitoring response during interspersed video trials. During in-vivo training for Bob, we only presented interspersed trials. Because Bob displayed independent appropriate task engagement during interspersed trials, a supplemental instruction (like that used for Chris) was not needed. Bob demonstrated independent and accurate self-monitoring on 90% of trials following nine training sessions. Self-monitoring training took approximately 2.58 hours for Bob.

During the second implementation of the self-monitoring materials condition (following self-monitoring training), Bob did not exhibit appropriate task engagement or productivity. In addition, Bob did not self-monitor his behavior. Similar to Chris, self-monitoring training with Bob did not result in accurate self-monitoring or increases in appropriate task engagement when only self-monitoring materials were presented.

During the first self-monitoring materials + DRA (accurate) phase, levels of appropriate task engagement increased slightly but subsequently decreased, and productivity remained low. However, Bob's self-monitoring accuracy increased to high levels. Similar to Chris, Bob was off-task on each self-monitoring opportunity, and accurately self-monitored that he was off-task. Thus, including consequences for accurate self-monitoring increased this behavior but did not facilitate increases in the self-monitored target behavior.

During self-monitoring materials + DRA (accurate & task engagement), Bob's appropriate task engagement and productivity rapidly increased, and self-monitoring accuracy remained high. During the second implementation of self-monitoring materials + DRA (accurate), appropriate task engagement and productivity initially maintained at high levels, but then decreased to baseline levels. This finding differed from that obtained for Chris, for whom self-monitoring materials + DRA (accurate) was effective following exposure to self-monitoring materials + DRA (accurate & engagement). Although task engagement and productivity decreased, self-monitoring accuracy remained high. When implementing self-monitoring materials + DRA (accurate and engagement) a second time, we observed a rapid increase in appropriate task engagement and productivity, and self-monitoring accuracy remained high. When conducting self-monitoring materials + DRA (accurate) a third time, we observed decreases in appropriate task engagement and productivity. However, self-monitoring accuracy remained high and stable. Although there were increases in accurate self-monitoring, we did not observe increases in appropriate task engagement. This response increased only when the reinforcement contingency was in effect for both accurate self-monitoring *and* appropriate task engagement.

Next, we assessed the independent effects of DRA (engagement) without self-monitoring by alternating this condition with the self-monitoring materials condition. Appropriate task engagement and productivity occurred at high levels during DRA (engagement) and at low levels during the self-monitoring materials condition. In addition, with the exception of one session, Bob did not self-monitor his behavior. These findings suggested that, for Bob, DRA was the critical component of the intervention.

Finally, we implemented self-monitoring materials + DRA (accurate and engagement) once again, and appropriate task engagement, productivity, and self-monitoring accuracy increased. In summary, these data indicated that DRA (engagement) and self-monitoring + DRA (accurate & engagement) were equally effective in increasing appropriate task engagement. These findings differed from those obtained for Chris, for whom the inclusion of self-monitoring materials appeared to enhance the effectiveness of DRA.

Figure 7 depicts results of the treatment component analysis for Scott. During the first implementation of the self-monitoring materials condition (prior to self-monitoring training), Scott's appropriate task engagement and productivity occurred at low levels and his motor stereotypy occurred at low-to-moderate and variable levels. Similar to Chris and Bob, these findings suggested that self-monitoring materials had no effect on appropriate task engagement or productivity prior to training. In addition, Scott never exhibited accurate self-monitoring, suggesting that he did not have this skill in his repertoire prior to training.

Figure 5 (bottom panel) depicts the results of self-monitoring training for Scott. Similar to Chris and Bob, Scott quickly acquired the self-monitoring response when video clips depicted only appropriate task engagement or off-task behavior on each trial. However, Scott's performance became more variable when the video clips were interspersed and he was required to make a discrimination on each trial. Because Scott emitted sequencing errors, we introduced an identical visual aid to that used for Chris and Bob. Following the introduction of the visual aid, Scott acquired the self-monitoring response during interspersed video trials. During interspersed trials of in-vivo training, Scott demonstrated independent and accurate self-monitoring on 90% of trials following seven training sessions. Self-monitoring training took approximately 3.5 hours for Scott.

During the second implementation of the self-monitoring materials condition (following self-monitoring training), Scott displayed moderate-to-high levels of appropriate task engagement, high levels of productivity, low-to-moderate and variable levels of stereotypy, and moderate levels of self-monitoring accuracy. However, appropriate task engagement, productivity, and self-monitoring accuracy were on a decreasing trend towards the end of the phase.

During the first self-monitoring materials + DRA (accurate) phase, levels of appropriate task engagement and productivity were lower than that observed during the previous self-monitoring materials phase. Scott also displayed low levels of stereotypy. However, Scott's self-monitoring accuracy increased to moderate-to-high levels.

During the first implementation of self-monitoring materials + DRA (accuracy & task engagement), Scott's appropriate task engagement and productivity initially increased but then decreased to low levels, and stereotypy increased slightly. In addition, self-monitoring accuracy occurred at moderate-to-high levels. Scott began refusing the edible when it was delivered contingent on appropriate task engagement and accurate self-monitoring, suggesting that it no longer functioned as a reinforcer. To assess whether the edibles functioned as a reinforcer and to evaluate whether the self-monitoring materials may have been responsible for the low levels of stereotypy observed, we implemented DRA (engagement) with edible reinforcement. During this condition, Scott displayed low levels of appropriate task engagement and productivity and high levels of stereotypy, suggesting that edibles no longer functioned as a reinforcer. In addition, these findings suggested that the presence of the self-monitoring materials may have competed with stereotypy or exerted stimulus control for low levels of stereotypy.

Because edibles no longer functioned as a reinforcer, we stopped delivering edibles and began using tokens as a reinforcer for the remainder of his treatment analysis. Based on the observation that Scott frequently requested breaks from various classroom tasks and activities, each token resulted in a 30-s break (for a maximum of 10 tokens or a 5-min break per session). During the second implementation of self-monitoring materials + DRA (accurate), Scott displayed low levels of appropriate task engagement, productivity, and motor stereotypy, and high, stable levels of self-monitoring accuracy. Thus, including consequences for accurate self-monitoring increased this behavior but did not facilitate increases in the self-monitored target behavior.

Next, we evaluated self-monitoring materials + DRA (accurate & task engagement) with token reinforcement. During this phase, appropriate task engagement and productivity increased to high levels, stereotypy remained low, and self-monitoring accuracy maintained at high, stable levels. Thus, clinically significant improvements in all dependent variables were observed. We then returned to self-monitoring materials + DRA (accurate). During this condition, appropriate task engagement, productivity, and self-monitoring accuracy initially maintained at high levels, but subsequently decreased. Stereotypy increased slightly over subsequent sessions. Next, we re-implemented self-monitoring materials + DRA (accurate and task engagement) and observed high levels of appropriate task engagement, productivity, and self-monitoring accuracy, and low levels of stereotypy, replicating the effects observed in the previous self-monitoring materials + DRA (accurate and task engagement) phase.

Next, we assessed the independent effects of DRA (engagement) without self-monitoring materials by alternating DRA (engagement) and self-monitoring materials conditions. During DRA (engagement), appropriate task engagement and productivity occurred at high levels and stereotypy occurred at low levels. During the self-monitoring materials phase, when materials were available but the reinforcement contingency was withdrawn, levels of engagement and productivity decreased and stereotypy increased. In addition, self-monitoring accuracy decreased to low levels. These findings suggested that DRA was the critical intervention component for Scott.

Treatment Acceptability Assessment

Procedure

We used procedures similar to those described by Hanley, Piazza, Fisher, Contrucci, and Maglieri (1997) to determine treatment acceptability for Bob. Specifically, we presented Bob with three response options that were each associated with a different intervention evaluated during the treatment assessment. The purpose of this assessment was to determine whether self-monitoring, irrespective of its efficacy, was selected for use by Bob when he was offered a choice between self-monitoring materials alone or in combination with DRA relative to DRA alone.

Color Preference Assessment. To ensure that the signals (i.e., different colors of construction paper) for the concurrent-chains assessment did not differentially affect responding in the absence of contingencies, we conducted a brief paired-stimulus assessment (similar to that described by Fisher et al., 1992) of different colored pieces of construction paper. Fifty-six trials were presented (each colored card was presented 14 times in the left and right positions, and each card was presented with every other card). Three colors that were selected on a moderate number of trials (not the most or the least) were selected for inclusion.

Treatment Acceptability Assessment Procedure. During all sessions, there were three tables, aligned side-by-side, and tape on the floor marking where Bob's feet were positioned when asked to make a selection. Three chairs were positioned in front of each table, and the chairs were positioned so that the back legs lined up with a line that was parallel to the table. Colored cards were affixed to the backs of the chairs. Participant selections during the initial link resulted in access to terminal links and their corresponding contingencies. A selection was defined as Bob pointing with an isolated index finger and making contact with one of the color cards on the back of the chairs or simply making contact with the initial link card with his hand. At the start of each session, the experimenter instructed Bob to select a stimulus (initial link) to contact contingencies associated with the terminal link. After Bob made a selection, he was instructed to sit in the corresponding chair, and the terminal link antecedents and consequences associated with that chair were initiated. The order of the colored cards was rotated clockwise across trials.

Baseline. During baseline, selecting any of the initial link stimuli resulted in the delivery of an edible. Bob was instructed to sit at the corresponding table, received an edible for doing so, and the trial ended.

Reinforcement. During reinforcement, a selection response in each of the initial links was associated with a different treatment. The red card was associated with self-monitoring plus DRA (accurate & task engagement), the green card was associated with DRA (engagement), and the orange card was associated with the self-monitoring materials condition. Prior to the reinforcement phase, we conducted two forced-exposure trials to expose Bob to the terminal link contingencies associated with initial link responses.

Initial Link. Before Bob was given the opportunity to make the first selection, the experimenter stated, "You can choose to sit at any one of these tables. If you choose the red card, you can work and check boxes. When you hear the timer beep, if you were working and check the correct box, you can have a (name of edible). If you choose the green card, you can work. When you hear the timer beep, if you are working, you can have a (name of edible). If you choose the orange card, you can work and check boxes, but you're not earning anything. Touch the colored card of the table you'd like to sit at." Contingent on a card selection, the experimenter prompted Bob to sit at the table associated with the selected card and initiated the terminal link.

Terminal Link. During the terminal link, the experimenter conducted the treatment condition associated with the selected card for 2 min. During all terminal link treatment conditions, the experimenter presented the same task that was used during the self-monitoring intervention (e.g., weight lifting). The experimenter conducted each of the treatment conditions exactly as it was conducted during the self-monitoring intervention analysis described above. If Bob selected the card associated with the self-monitoring materials plus DRA (accurate & task engagement) treatment option during the initial link, then the experimenter presented self-monitoring materials to him prior to the start of the terminal link session. During the 2-min session, the experimenter delivered praise and a preferred edible contingent on appropriate task engagement and accurate self-monitoring at the sound of the tone. If Bob selected the card associated with the DRA (engagement) treatment during the initial link, then the experimenter did not present self-monitoring materials to him prior to the start of the session. During the session, the experimenter delivered praise and a preferred edible contingent on appropriate task engagement. If Bob selected the card associated with the self-monitoring materials condition (the control) during the initial link, then the experimenter presented self-monitoring materials to him prior to the start of the session. During the 2-min session, the experimenter delivered no programmed consequences contingent on accurate self-monitoring or appropriate task engagement.

Treatment Acceptability Assessment Results

Figure 8 depicts the results of the treatment preference assessment for Bob. During baseline, Bob's initial link selections were undifferentiated. Bob selected the middle chair on every trial, regardless of the colored card affixed to the back of the chair. During the reinforcement condition, Bob selected each colored card and experienced each treatment at least one time following the forced-exposure trials. However, after trial 15, Bob consistently selected the initial link card associated with self-monitoring + DRA (accurate & task engagement) treatment option, suggesting that he preferred DRA combined with self-monitoring materials rather than DRA alone.

Discussion

Results of previous studies (e.g., Koegel & Koegel, 1990; Stahmer & Schreibman, 1992) indicated that self-monitoring interventions are effective for increasing appropriate behavior and decreasing problem behavior. However, few studies have evaluated the independent effects of the different components commonly included in self-monitoring interventions. Studies that have conducted component analyses of self-monitoring interventions (Fritz et al., 2012; Storey & Gaylord-Ross, 1987) have yielded mixed findings. Results of the current study indicated that teaching participants to accurately self-monitor, and subsequently providing self-monitoring materials without differential reinforcement, was ineffective for increasing participants' appropriate task engagement. However, increases in appropriate behavior were produced when participants were provided with the self-monitoring materials and reinforcement was delivered for accurate self-monitoring (Chris) or for accurate self-monitoring *and* appropriate task engagement (Bob and Scott).

For all participants, simply providing access to self-monitoring materials was ineffective for increasing appropriate task engagement both before and after self-monitoring training. During the self-monitoring materials phase following training, when materials were available and the reinforcement contingency for accuracy was withdrawn, levels of engagement and productivity did not maintain (Chris and Bob) or initially maintained but subsequently decreased (Scott). These findings suggested that although participants had this skill in their repertoire, they would not self-monitor in the absence of programmed reinforcement.

For all participants, the self-monitoring materials + DRA (accurate) condition resulted in increases in accurate self-recording. However, this condition did not result in increases in appropriate task engagement. Instead, participants exhibited off-task behavior and accurately self-monitored their behavior, resulting in frequent reinforcer deliveries. During the self-monitoring materials + DRA (accurate & engagement) condition, participants showed clinically acceptable increases in appropriate task behavior. For Chris and Scott, this outcome was obtained only after the reinforcer was changed from edibles to tokens. For Bob and Scott, subsequent replications of the self-monitoring + DRA (accurate) condition resulted in decreases in appropriate task engagement even though self-monitoring accuracy maintained at high levels. By contrast, Chris' appropriate task engagement maintained during subsequent replications of the self-monitoring materials + DRA (accurate) condition after exposure to self-monitoring materials + DRA (accurate & engagement), suggesting that the presence of the self-monitoring materials or the tokens may have functioned as a discriminative stimulus for appropriate task engagement.

For Chris, although self-monitoring materials + DRA (accurate) and DRA (engagement) were effective interventions relative to presenting self-monitoring materials alone, self-monitoring materials + DRA (accurate) was associated with higher levels of appropriate task engagement than DRA (engagement). For Bob and Scott, self-monitoring + DRA (accurate & engagement) and DRA (engagement) conditions were equally effective relative to self-monitoring materials alone, suggesting that DRA was the critical component of the intervention.

The results of the current study add to the existing self-monitoring literature in several ways. Although some studies describe procedures for teaching self-monitoring (e.g., Fritz et al., 2012; Koegel & Koegel, 1990), few studies have empirically evaluated their training procedures for promoting self-monitoring accuracy, limiting conclusions that can be drawn regarding the efficacy of their training method. To address this limitation, we evaluated self-monitoring training prior to intervention by conducting video and in-vivo training phases using a multiple baseline across participants experimental design to demonstrate experimental control. We also included a detailed description of our training procedure to permit replication. Vocal and point prompts were systematically faded in 2-s increments following two consecutive sessions with 80% accuracy. An error correction procedure was implemented if the participant inaccurately self-monitored his behavior. Self-monitoring training continued until participants met a criterion of 80% accuracy and independence across two consecutive trials. We reported accuracy data across baseline and training phases to ensure that participants could accurately and independently self-monitor their behavior. These modifications may promote a technology for teaching individuals with ID to self-monitor. Specifically, by empirically evaluating the effects of training through the use of an experimental design, providing clear and concise procedural details, and including accuracy data, researchers can assess the relative efficacy and efficiency of various self-monitoring training procedures.

We extended the training procedure described by Fritz et al. (2012), Koegel and Koegel (1990), and Stahmer and Schreibman (1992) in several ways. In the Fritz et al., Koegel and Koegel, and Stahmer and Schreibman studies, training involved teaching participant's to record occurrences and nonoccurrences of stereotypy or appropriate behavior modeled by the therapist. However, in the current study, to promote accurate and independent self-monitoring of the participant's target behavior, we conducted occurrence, nonoccurrence, and interspersed trials of the participant emitting the target behavior, using video clips and in vivo trial formats. By contrast, Fritz et al. presented only interspersed occurrence and non-occurrence trials using an in-vivo format. For one participant who did not acquire the self-monitoring response during interspersed trials, Fritz et al. presented conducted additional training trials in which one occurrence trial was interspersed with nine nonoccurrence trials, and vice versa. In the current study, we first presented occurrence trials only, then nonoccurrence trials only, and then interspersed trial types. Participants were prompted to record occurrences or nonoccurrences on a worksheet at the sound of a tone, and reinforcement was delivered for accurate self-monitoring. If the participant did not self-monitor accurately (checked the incorrect box), we erased the error and physically guided the participant to check the correct box.

Although some authors have asserted that self-monitoring is easy to use (e.g., Agran et al., 2005), data from the current study and those from Fritz et al. (2012) suggest otherwise. In the current study, training required 4.25 hours for Chris, 2.58 hours for Bob, and 3.5 hours for Scott. During the video phase of training in the current study, participants met criterion performance quickly during the occurrence only and nonoccurrence only video clip trials (approximately 4 or 5 sessions, or 20 min). However, when the video clips were interspersed across trials, requiring participants to make a conditional discrimination on each trial, rate of acquisition declined and participants emitted frequent sequencing errors (i.e., participants frequently checked the boxes on the self-monitoring worksheet out of order, rather than sequentially). However, a simple visual aid in the form of a cardboard cutout window was effective in remediating this problem. The interspersed video training phase of the study required approximately 1.9 hours for Chris, 1 hour for Bob, and 2 hours for Scott. In the Fritz et al. study, two of three participants demonstrated accurate self-monitoring during trials when the therapist modeled the participant's stereotypy. For the other participant, additional training components were required, including an overcorrection procedure following errors, interspersing occurrence and nonoccurrence trials, and the use of video clips of the participant engaging in stereotypy or engaging in appropriate behavior. Although this participant acquired the self-monitoring response, it took over 5 hours. Therefore, teaching self-monitoring may require a significant time investment.

We also extended previous research by using a variable momentary (VM) self-monitoring procedure instead of a whole-interval recording method. In previous research, whole-interval recording schedules are most often used. For example, Koegel and Koegel (1990) taught participants to place a mark in a box following intervals with no stereotypy, whereas Stahmer and Schreibman (1992) taught participants to place a mark in a box if appropriate play occurred for the entire interval. In the in-vivo training procedure of the current study, we conducted a VM recording method by sounding a tone when the participant was exhibiting appropriate task engagement or non-work behavior. At the sound of the tone, we prompted the participant to self-monitor the behavior that was occurring *at that moment*. In a momentary DRA procedure, reinforcement is delivered for the occurrence of appropriate behavior at the end of the interval. In addition, because the interval duration is varied, it is less discriminable when the tone will sound. Lindberg, Iwata, Kahng, and DeLeon (1999) noted some advantages of VM reinforcement schedules over other reinforcement schedules. First, VM schedules may be more practical than whole interval schedules because the therapist does not need to continuously monitor the participant's behavior throughout the interval. Instead, the therapist merely needs to observe the participant's behavior at the end of the interval to determine whether reinforcement should be delivered. This advantage may also be true for participants who are learning to self-monitor their behavior. It may be easier and more practical for participants to self-monitor their behavior at the moment the interval ends, rather than continuously throughout the interval. A second advantage of VM schedules is that, because the criterion for reinforcement is less stringent than it is with whole interval schedules, a higher percentage of reinforcers may be earned.

The current study demonstrated one method for teaching participants to self-monitor their behavior. All three participants acquired the self-monitoring response during video and in-vivo self-monitoring training. However, future research is needed to determine the best training procedures for teaching self-monitoring. Given the paucity of research in this area, it is unclear which teaching strategies are most effective and efficient. For example, researchers could evaluate whether it is most effective to conduct training prior to or simultaneous with intervention. In addition, researchers could evaluate the relative utility of various self-monitoring strategies, such as direct instruction (e.g., lecture + demonstration), modeling, prompting, rehearsal, and reinforcement, alone and in combination. By conducting comparative analyses, one can determine which component or combination of components is most effective for teaching self-monitoring. In addition, researchers could report correct responding and sessions or trials to criterion to inform the reader about the relative efficacy and efficiency of the method used.

Similarly, more research is needed to determine what, if any, prerequisite skills may be necessary for self-monitoring acquisition. No studies to date have systematically evaluated the acquisition of self-monitoring in relation to participants' discrimination repertoire. Because self-monitoring interventions require the participant to discriminate between the occurrence and nonoccurrence of a target response, a necessary prerequisite for this intervention may be a demonstration of competency with conditional discriminations. Other variables to evaluate might include the participant's receptive and expressive communication repertoire, counting skills, and motor skills related to performing the self-monitoring response. With accumulated data, it may be possible to identify individuals for whom self-monitoring is and is not appropriate (e.g., based on diagnosis, IQ scores, discrimination skills, verbal abilities, etc.; Tiger, Fisher, & Bouxsein, 2009).

The second contribution of the current study was the inclusion of a treatment component analysis. Although previous studies have shown that self-monitoring interventions including self-monitoring materials combined with differential reinforcement for other behavior (e.g., Shabani, Wilder, & Flood, 2001), alternative behavior (e.g., Koegel & Frea, 1993) or accurate self-monitoring (e.g., Frea & Hughes, 1997) produced positive treatment outcomes, few studies have evaluated the independent effects of only presenting self-monitoring materials. To address this limitation, we conducted an additive component analysis to evaluate the effects of self-monitoring materials alone and in combination with differential reinforcement on appropriate task engagement, motor stereotypy (Chris and Scott), productivity, and self-monitoring accuracy. We conducted a component analysis by initiating treatment and successively adding single components until clinically significant outcomes were obtained. By conducting this design, the individual and additive effects of each treatment component were identified. First, we trained self-monitoring until a criterion level of self-monitoring accuracy and independence was achieved, and then subsequently introduced self-monitoring materials alone. Because self-monitoring alone was ineffective, DRA for accurate self-monitoring was added. Initially, DRA for accurate self-monitoring was ineffective for increasing appropriate task engagement for all participants (although self-monitoring accuracy increased). Thus, we evaluated self-monitoring + DRA for accurate self-monitoring *and* for appropriate behavior. Once this combination of components was found to be effective, we evaluated DRA for only appropriate behavior, without self-monitoring materials. For all three participants, this single component was effective. For Bob and Scott, self-monitoring + DRA for accurate self-monitoring *and* appropriate engagement and DRA for engagement (without self-monitoring) were equally effective, suggesting that self-monitoring was not a critical intervention component. However, for Chris, self-monitoring + DRA (accurate) proved to be slightly more effective than DRA for appropriate engagement, suggesting that self-monitoring enhanced the effects of DRA.

When evaluating self-monitoring interventions, authors often do not include data on self-monitoring accuracy (e.g., Christian & Poling, 1997; Ganz & Sigafoos, 2005; Koegel & Frea, 1993). When accuracy data are reported, they are typically reported as a mean value (e.g., Koegel & Koegel, 1990; Shabani, Wilder, & Flood, 2001; Stahmer & Schreibman, 1992). For instance, Koegel and Koegel (1990) reported mean self-monitoring accuracy for three of the four participants, and found that participant accurately self-monitored the absence of their stereotypy but not the occurrence of their stereotypy. Although the authors noted that accuracy improved over the course of the study, accuracy data were not graphically displayed. A surprising finding of Koegel and Koegel was that although the self-monitoring intervention was effective, participants did not accurately record the occurrence of their stereotypy. These findings suggest that some intervention component other than accurate self-monitoring may have been responsible for the effects obtained.

We extended previous research by measuring self-monitoring accuracy across all self-monitoring phases of the component analysis, and reported session-by-session accuracy data. When only self-monitoring materials were presented in the absence of differential reinforcement for accuracy, accurate self-monitoring did not occur. When DRA was in effect for accurate self-monitoring, self-monitoring accuracy increased to high levels for all participants. Chris, Bob, and Scott were off-task on each self-monitoring opportunity and accurately self-monitored their off-task behavior. Thus, including consequences for accurate self-monitoring increased this behavior but did not facilitate increases in the self-monitored target behavior. Appropriate task engagement increased and self-monitoring accuracy remained high when consequences were included for both behaviors during the self-monitoring + DRA (accurate and engagement) phase.

Future research is needed to further clarify the role of accuracy in self-monitoring interventions. First, researchers could continue to report session-by-session accuracy data during training and subsequent treatment. These data would permit the detection of trends or patterns of self-monitoring accuracy and corresponding changes in the self-monitored behavior. If accuracy data are low, other variables could be examined more closely (such as the arrangement of reinforcement contingencies, compliance with self-monitoring and the type of recording device used) and systematic changes could be made to improve accuracy. Second, future research should attempt to clarify the degree to which accuracy of self-monitoring enhances treatment effects. For instance, Koegel and Koegel (1990) found that a self-monitoring intervention was effective for reducing participants' stereotypy, even though self-monitoring accuracy was variable. From a practical standpoint, if the goal of treatment is to affect behavior, ensuring self-monitoring accuracy may not be critical for achieving desired outcomes. However, from a scientific point of view, it is important to design empirically sound analyses to determine the potential controlling variables included in behavioral interventions. Future research could evaluate self-monitoring when reinforcement for accurate recording is and is not in effect to identify the relative contribution of this independent variable.

A third contribution of the current study was the inclusion of a treatment acceptability assessment. Although several studies have evaluated the social validity of self-monitoring interventions with parents or teachers, few have directly assessed participant preference for self-monitoring interventions. However, the importance of determining individuals' preference for behavioral interventions increases when two or more intervention strategies prove to be equally effective (Luczynski & Hanley, 2010), as was the case in the current study. Studies that have assessed participant preference for self-monitoring have most commonly used indirect methods (e.g., Copeland et al., 2002; Hildebrand, Martin, Furer, & Hazen, 1990; Hughes et al., 2002; Kaplan, Hemmes, Motz, & Rodriguez, 1996; Rae, Martin, & Smyk, 1990). For example, Copeland et al. (2002) conducted open-ended interviews with participants before and after a self-monitoring intervention for worksheet completion in the classroom. Prior to intervention, participants were asked whether they were motivated to change their behavior, and after intervention, they were asked if they liked the treatment and if their behavior changes goals had been met. Although this strategy is efficient and may yield qualitatively rich information with the direct recipients of treatment, it is unlikely to be effective for use with individuals with ID who possess less sophisticated vocal repertoires.

In the current study, we used a concurrent-chains procedure to assess Bob's preference for treatments. In this arrangement, Bob was prompted to make an initial choice response (e.g., touch a colored card initial link), and each response resulted in a different treatment (terminal link). For Bob, DRA alone and self-monitoring combined with DRA were found to be equally effective for increasing appropriate behavior. Thus, each treatment was included as a terminal link in our concurrent-chains procedure. A control condition (self-monitoring materials *without* reinforcement) was included to distinguish between indiscriminate initial link responding (represented by equal responding on all initial links) and indifferent initial link responding (represented by approximately 50% of responding to each treatment condition). We reported data on the cumulative number of initial link responses to permit visual inspection of patterns of responding. Bob consistently selected the initial link associated with the self-monitoring + DRA treatment, suggesting that he preferred to self-monitor his appropriate task engagement.

Instead of relying on a restricted set of treatment procedures assumed to be in the best interest of the individual, use of a concurrent-chain preference assessment may help clinicians identify treatments that are both effective and preferred. Objective, empirically based methods of assessing treatment acceptability may offer advantages to indirect methods (i.e., interviews or questionnaires). For example, objective methods can be conducted with direct consumers, some of whom may have limited expressive skills that prevent them from answering questionnaires or interviews. Thus, treatment preference assessments may allow individuals with ID to participate in the habilitative planning process in more meaningful ways.

One noteworthy limitation of the current study is that we did not evaluate the generality of treatment effects. We did not evaluate the maintenance of treatment effects while thinning the reinforcement schedule, nor did we implement the effective intervention across people, settings, or tasks. In previous literature, maintenance has been evaluated in several ways. Some authors have evaluated the durability of treatment when the self-monitoring materials were withdrawn (Koegel & Koegel, 1990; Rudrud, Ziarnek, & Colman, 1984; Stahmer & Schreibman, 1992). For example, after conducting a self-monitoring condition that included self-monitoring materials, DRO, and DRA for accurate self-monitoring, Koegel and Koegel (1990) removed all self-monitoring components, including the self-monitoring materials and observed increases in problem behavior. When the self-monitoring intervention was reinstated, problem behavior decreased. The finding that problem behavior increased during a return to baseline is not surprising because no discriminative stimuli or consequences were provided. Alternatively, Stahmer and Schreibman (1992) conducted a withdrawal condition in which only the self-monitoring materials were removed and appropriate behavior maintained. Because reinforcement (praise) continued to be delivered for appropriate behavior, this result suggests that self-monitoring may not have been necessary for treatment effects. However, authors have argued that self-monitoring is a useful intervention because it may enhance the generality of treatment effects for individuals with ID (e.g., Brooks, Todd, Tofflemoyer, & Horner, 2003; Koegel et al., 2001; Loftin, Odom, & Lantz, 2008). For example, Loftin et al. (2008) stated that self-monitoring may enhance maintenance of treatment effects because the control of some aspect of the intervention is shifted from the treatment provider to the recipient. Koegel et al. (2001) argued that self-monitoring is a pivotal skill that may promote maintenance of treatment gains, produce improvements in behavior across various contexts, and increase independence in individuals with ID. Thus, it may be most useful to keep the self-monitoring component in effect after desirable treatment effects have been obtained. Other strategies for programming maintenance with self-monitoring are to add a self-monitoring component following a successful intervention (e.g., Feldman, 1986; Kiburz, Miller, & Morrow, 1984; Newman & Ten Eyck, 2005), implement a self-monitoring intervention and subsequently withdraw all treatment components except for the presentation of self-monitoring materials (e.g., Agran et al., 2005; Agran, Wehmeyer, Cavin, & Palmer, 2008), and present self-monitoring materials and subsequently thin the schedule of reinforcement (e.g., Gardner, Clees, & Cole, 1983; Koegel & Frea, 1993).

Treatment generality has also been evaluated in previous research by assessing the effects of the self-monitoring intervention across people or settings (e.g., Kern, Marder, Boyajian, Elliot, & McElhattan, 1997; Koegel & Koegel, 1990; Mancina, Tankersley, Kamps, Kravits, & Parrett, 2000). For instance, Koegel and Koegel (1990) assessed stimulus generalization of treatment effects (low stereotypy) across settings (i.e., from a clinic to a home or school). Results showed that stimulus generalization did not occur; stereotypy maintained at high levels in the untreated settings. Therefore, the authors sequentially introduced the self-monitoring intervention (i.e., self-monitoring materials, instructions, and reinforcement for self-monitoring) in the previously untreated setting, resulting in decreases in participants' stereotypy. Although application of the self-monitoring treatment was required in each setting, it should be noted that additional self-monitoring training was not required prior to treatment implementation in the untreated settings. In the treatment setting, prompts to self-monitor were initially delivered and subsequently faded by delivering vocal and physical prompts to accurately self-monitor only when the participant did not exhibit correct self-monitoring. In addition, the schedule of reinforcement was gradually thinned by increasing the number of boxes (from 1 to several) per self-monitoring worksheet and by increasing the duration of the self-monitoring interval from 15-30 s to 15-20 min such that the participant requested a reinforcer from the therapist after intervals of at least 1 hour. In the novel setting, the therapist simply presented the self-monitoring materials to the participant, instructed him to use the materials, and implemented the terminal reinforcement schedule.

In addition, only a few studies have assessed maintenance of treatment effects under conditions of reduced supervision (e.g., Belfiore et al., 1989; Stahmer & Schreibman, 1992), which is surprising in light of the fact that a proposed advantage of self-monitoring is that it may promote maintenance of treatment effects under conditions of reduced supervision (Koegel et al., 2001). A comprehensive evaluation of the generality of self-monitoring is an important area in need of further research as it is unclear to what extent self-monitoring may facilitate maintenance of treatment effects. To this end, researchers could evaluate whether self-monitoring materials may enhance maintenance effects by comparing differential reinforcement schedule thinning procedures with and without self-monitoring materials present. Comparative evaluations of treatment generality with and without a self-monitoring component may yield particularly useful information about the utility of self-monitoring as a maintenance strategy.

References

Ackerman, A. M., & Shapiro, E. S. (1984). Self-monitoring and work productivity with mentally retarded adults. *Journal of Applied Behavior Analysis, 17*(3), 403-407.

Agran, M., Sinclair, T., Alper, S., Cavin, M., Wehmeyer, M., & Hughes, C. (2005). Using self-monitoring to increase following-direction skills of students with moderate to severe disabilities in general education. *Education and Training in Developmental Disabilities, 40*(1), 3-13.

Agran, M., Wehmeyer, M. L., Cavin, M., & Palmer, S. (2008). Promoting student active classroom participation skills through instruction to promote self-regulated learning and self-determination. *Career Development for Exceptional Individuals, 31*(2), 106-114.

American Association on Intellectual and Developmental Disabilities (2012). Retrieved January 14, 2012, from http://www.aaidd.org/content_100.cfm?navID=21

Belfiore, P. J., Mace, F. C., & Browder, D. M. (1989). Effects of experimenter surveillance on reactive self-monitoring. *Research in Developmental Disabilities, 10*(2), 171-182.

Blick, D. W., & Test, D. W. (1987). Effects of self-recording on high-school students' on-task behavior. *Learning Disability Quarterly, 10*(3), 203-213.

Brooks, A., Todd, A. W., Tofflemoyer, S., & Horner, R. H. (2003). Use of functional assessment and a self-management system to increase academic engagement and work completion. *Journal of Positive Behavior Interventions, 5*(3), 144-152.

Christian, L., & Poling, A. (1997). Using self-management procedures to improve the productivity of adults with developmental disabilities in a competitive employment setting. *Journal of Applied Behavior Analysis, 30*(1), 169-172.

Copeland, S. R., Hughes, C., Agran, M., Wehmeyer, M. L., & Fowler, S. E. (2002). An intervention package to support high school students with mental retardation in general education classrooms. *American Journal on Mental Retardation, 107*(1), 32-45.

Dube, W. V., & McIlvane, W. J. (1997). Reduction of stimulus selectivity with nonverbal differential observing responses. *Journal of Applied Behavior Analysis, 32,* 25-34.

Feldman, M. A. (1986). Parent education project: II. Increasing stimulating interactions of developmentally handicapped mothers. *Journal of Applied Behavior Analysis, 19*(1), 23-37.

Fisher, W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, I. (1992). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis, 25*(2), 491-498.

Frea, W. D., & Hughes, C. (1997). Functional analysis and treatment of social-communicative behavior of adolescents with developmental disabilities. *Journal of Applied Behavior Analysis, 30*(4), 701-704.

Fritz, J. N., Iwata, B. A., Rolider, N. U., Camp, E. M., Neidert, P. L. (2012). Analysis of self-recording in self-management interventions for stereotypy.

Ganz, J. B., & Sigafoos, J. (2005). Self-monitoring: Are young adults with MR and autism able to utilize cognitive strategies independently? *Education and Training in Developmental Disabilities, 40*(1), 24-33.

Gardner, W. I., Clees, T. J., & Cole, C. L. (1983). Self-management of disruptive verbal ruminations by a mentally retarded adult. *Applied Research in Mental Retardation, 4*(1), 41-58.

Gilberts, G. H., Agran, M., Hughes, C., & Wehmeyer, M. (2001). The effects of peer delivered self-monitoring strategies on the participation of students with severe disabilities in general education classrooms. *Journal of the Association for Persons with Severe Handicaps, 26*(1), 25-36.

Hanley, G. P., Piazza, C. C., Fisher, W. W., Contrucci, S. A., & Maglieri, K. A. (1997). Evaluation of client preference for function-based treatment packages. *Journal of Applied Behavior Analysis, 30*(3), 459-473.

Harrower, J. K., & Dunlap, G. (2001). Including children with autism in general education classrooms: A review of effective strategies. *Behavior Modification, 25*(5), 762-784.

Hildebrand, R. G., Martin, G. L., Furer, P., & Hazen, A. (1990). A recruitment-of-praise package to increase productivity levels of developmentally handicapped workers. *Behavior Modification, 14*(1), 97-113.

Howlin, P., Goode, S., Hutton, J., & Rutter, M. (2004). Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry, 45*(2), 212-229.

Hughes, C., Copeland, S. R., Agran, M., Wehmeyer, M. L., Rodi, M. S., & Presley, J. A. (2002). Using self-monitoring to improve performance in general education high school classes. *Education & Training in Mental Retardation & Developmental Disabilities, 37*(3), 262-272.

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis, 27*, 197-209. (Reprinted from Analysis and Intervention in Developmental Disabilities, 2, 3-20, 1982).

Kapadia, S., & Fantuzzo, J. W. (1988). Training children with developmental disabilities and severe behavior problems to use self-management procedures to sustain attention to preacademic/academic tasks. *Education & Training in Mental Retardation, 23*(1), 59-69.

Kaplan, H., Hemmes, N. S., Motz, P., & Rodriguez, H. (1996). Self-reinforcement and persons with developmental disabilities. *The Psychological Record, 46*(1), 161-178.

Kern, L., Marder, T. J., Boyajian, A. E., Elliot, C. M., & McElhattan, D. (1997). Augmenting the independence of self-management procedures by teaching self-initiation across settings and activities. *School Psychology Quarterly, 12*(1), 23-32.

Kiburz, C. S., Miller, S. R., & Morrow, L. W. (1984). Structured learning using self-monitoring to promote maintenance and generalization of social skills across settings for a behaviorally disordered adolescent. *Behavioral Disorders, 10*(1), 47-55.

Koegel, L. K., Koegel, R. L., Hurley, C., & Frea, W. D. (1992). Improving social skills and disruptive behavior in children with autism through self-management. *Journal of Applied Behavior Analysis, 25*(2), 341-353.

Koegel, R. L., & Frea, W. D. (1993). Treatment of social behavior in autism through the modification of pivotal social skills. *Journal of Applied Behavior Analysis, 26*(3), 369-377.

Koegel, R. L., & Koegel, L. K. (1990). Extended reductions in stereotypic behavior of students with autism through a self-management treatment package. *Journal of Applied Behavior Analysis, 23*(1), 119-127.

Koegel, R. L., Koegel, L. K., & McNerney, E. K. (2001). Pivotal areas in intervention for autism. *Journal of Clinical Child Psychology, 30*(1), 19-32.

Lindberg, J. S., Iwata, B. A., Kahng, S., & DeLeon, I. G. (1999). DRO contingencies: An analysis of variable-momentary schedules. *Journal of Applied Behavior Analysis, 32(2),* 123-136.

Loftin, R. L., Odom, S. L., & Lantz, J. F. (2008). Social interaction and repetitive motor behaviors. *Journal of Autism and Developmental Disorders, 38(6)*, 1124-1135.

Luczynski, K. C., & Hanley, G. P. (2010). Examining the generality of children's preference for contingent reinforcement via extension to different responses, reinforcers, and schedules. *Journal of Applied Behavior Analysis, 43*(3), 397-409.

Mancina, C., Tankersley, M., Kamps, D., Kravits, T., & Parrett, J. (2000). Reduction of inappropriate vocalizations for a child with autism using a self-management treatment program. *Journal of Autism and Developmental Disorders, 30*(6), 599-606.

Newman, B., Buffington, D. M., & Hemmes, N. S. (1996). Self-reinforcement used to increase the appropriate conversation of autistic teenagers. *Education & Training in Mental Retardation & Developmental Disabilities, 31*(4), 304-309.

Newman, B., & Ten Eyck, P. (2005). Self-management of initiations by students diagnosed with autism. *Analysis of Verbal Behavior, 21*, 117-122.

Newman, B., Tuntigian, L., Ryan, C. S., & Reinecke, D. R. (1997). Self-management of a DRO procedure by three students with autism. *Behavioral Interventions, 12*(3), 149-156.

Rae, S., Martin, G. L., & Smyk, B. (1990). A self-management package versus a group exercise contingency for increasing on-task behaviour of developmentally handicapped workers. *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences Du Comportement, 22*(1), 45-58.

Rudrud, E. H., Ziarnik, J. P., & Colman, G. (1984). Reduction of tongue protrusion of a 24-year-old woman with Down syndrome through self-monitoring. *American Journal of Mental Deficiency, 88*(6), 647-652.

Sainato, D. M., Goldstein, H., & Strain, P. S. (1992). Effects of self-evaluation on preschool children's use of social interaction strategies with their classmates with autism. *Journal of Applied Behavior Analysis, 25*(1), 127-141.

Shabani, D. B., Wilder, D. A., & Flood, W. A. (2001). Reducing stereotypic behavior through discrimination training, differential reinforcement of other behavior, and self monitoring. *Behavioral Interventions, 16*(4), 279-286.

Shear, S. M., & Shapiro, E. S. (1993). Effects of using self-recording and self-observation in reducing disruptive behavior. *Journal of School Psychology, 31*(4), 519-534.

Stahmer, A. C., & Schreibman, L. (1992). Teaching children with autism appropriate play in unsupervised environments using a self-management treatment package. *Journal of Applied Behavior Analysis, 25*(2), 447-459.

Storey, K., & Gaylord-Ross, R. (1987). Increasing positive social interactions by handicapped individuals during a recreational activity using a multicomponent treatment package. *Research in Developmental Disabilities, 8*(4), 627-649.

Tiger, J. H., Fisher, W. W., & Bouxsein, K. J. (2009). Therapist- and self-monitored DRO contingencies as a treatment for the self-injurious skin picking of a young man with Asperger syndrome. *Journal of Applied Behavior Analysis, 42*(2), 315-319.

Vollmer, T. R., Marcus, B. A., Ringdahl, J. E., & Roane, H. S. (1995). Progressing from brief assessments to extended experimental analyses in the evaluation of aberrant behavior.

Journal of Applied Behavior Analysis, 28(4), 561-576.

Table 1. *Operational Definitions of Appropriate Task Engagement and Productivity*

Task	Response Definition	Productivity
Beading	Pick up bead from pile, pick up string, thread string through bead opening, push bead to end of string	Number of beads strung
Filing cards	Pick up index card from pile, flip through filing box, file index card behind correct color tab	Number of index cards filed
Building Legos	Take Lego from bin or pile, stack Lego on baseboard, or stack Lego on Lego	Number of Legos stacked
Completing math facts	Orient toward computer screen, scan math fact and answer, click answer	Number of math facts completed
Assembling puzzle	Take puzzle piece from bin or pile, scan or arrange puzzle pieces on table, put or try to put puzzle pieces together	Number of puzzle pieces fit together
Sorting money	Pick up bill from pile, match bill to correct envelope, place bill with correct envelope	Number of bills sorted
Sorting silverware	Pick up utensil, match utensil to appropriate row, place utensil in tray	Number of utensils sorted
Stamping	Pick up stamp, press stamp onto ink pad, press stamp onto paper	Number of designs stamped
Stenciling	Pick up stencil, place stencil on paper, pick up pencil, stencil letter on paper	Number of letters stenciled
Stuffing envelopes	Pick up letter, fold letter, pick up empty envelope, put folded letter in envelope, put in bin	Number of enveloped stuffed
Weight lifting	Pick up one weight in each hand, bend elbows and raise weights above chest level, low weights below chest level	Number of repetitions

Table 2. *Operational Definitions of Motor Stereotypy*

Participant	Topography	Response Definition
Chris	Wrist twisting	Repetitively moving one or both wrists in a side-to-side or up-and-down motion 2 or more times in rapid succession
	Hand shaking	Holding one or both hands above the midline and repetitively moving the hands back-and-forth or up-and-down 2 or more times in rapid succession
	Body rocking	Moving the torso from side-to-side 2 or more times in rapid succession
	Head shaking	Moving the head from side-to-side, left-to-right, 2 or more times in rapid succession
Scott	Hand flapping	Repetitively moving one or both hands two or more times in rapid succession
	Body rocking	Moving the back-and-forth or side-to-side 2 or more times in rapid succession
	Ear holding	Placing one or both hands over the ear(s), one or several fingers in the ear(s), or lifting the shoulder to come into contact with the ear while tilting the head to that side
	Nose picking	Inserting one or more fingers into the nostril(s)
	Finger mouthing	Placing one or more fingers past the plane of the lips into the mouth.

Table 3. *Interobserver Agreement Values*

Participant	Phase	Dependent Variable	Mean (Range)
Chris	Skills Assessment	Task Completion	100%
	Task Assessment	Engagement	97% (90-100%)
		Motor STPY	99% (93.3-100%)
		RPM	99% (93.3-100%)
	Functional analysis	Motor STPY	95% (83-100%)
	SM Training	Accuracy	99% (75-100%)
	Component analysis	Engagement	95% (90-100%)
		Motor STPY	92% (80-96.7%)
		RPM	99% (96.7-100%)
		Accuracy	94% (75-100%)
Bob	Skills Assessment	Task Completion	100%
	Task Assessment	Engagement	99% (96.7-100%)
		RPM	92% (80-96.7%)
	SM Training	Accuracy	98% (93.3-100%)
	Component analysis	Engagement	96% (86.7-100%)
		RPM	92% (80-96.7%)
		Accuracy	98% (87.5-100%)
	Treatment PA	Initial link selection	100%
Scott	Skills Assessment	Task Completion	99% (90-100%)
	Task Assessment	Engagement	91% (80-96.7 (%)
		Motor STPY	96% (80-100%)
		RPM	
	Functional analysis	Motor STPY	98% (93.3-100%)
	SM Training	Accuracy	90% (62.5-100%)
	Component analysis	Engagement	96% (80-100%)
		Motor STPY	92% (73.3-100%)
		RPM	99% (96-100%)
		Accuracy	98% (87.5-100%)

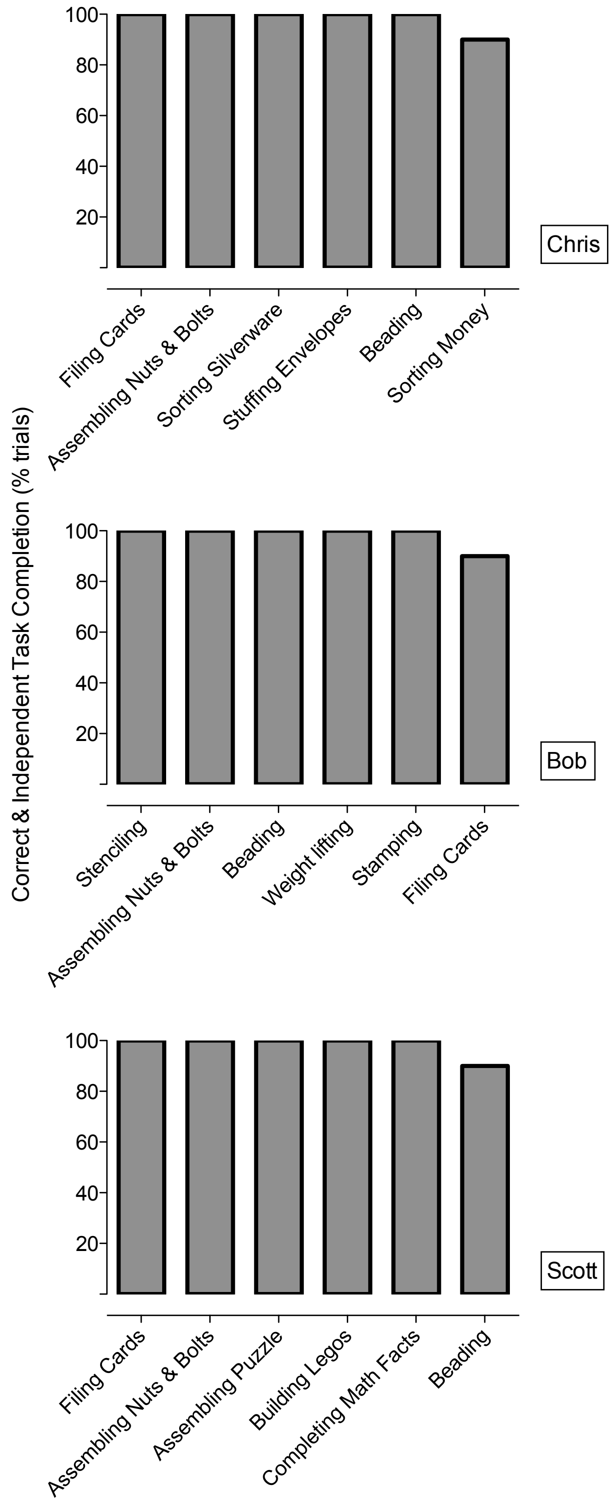


Figure 1. Percent of correct and independent responses across items during the skills assessment for Chris, Bob, and Scott.

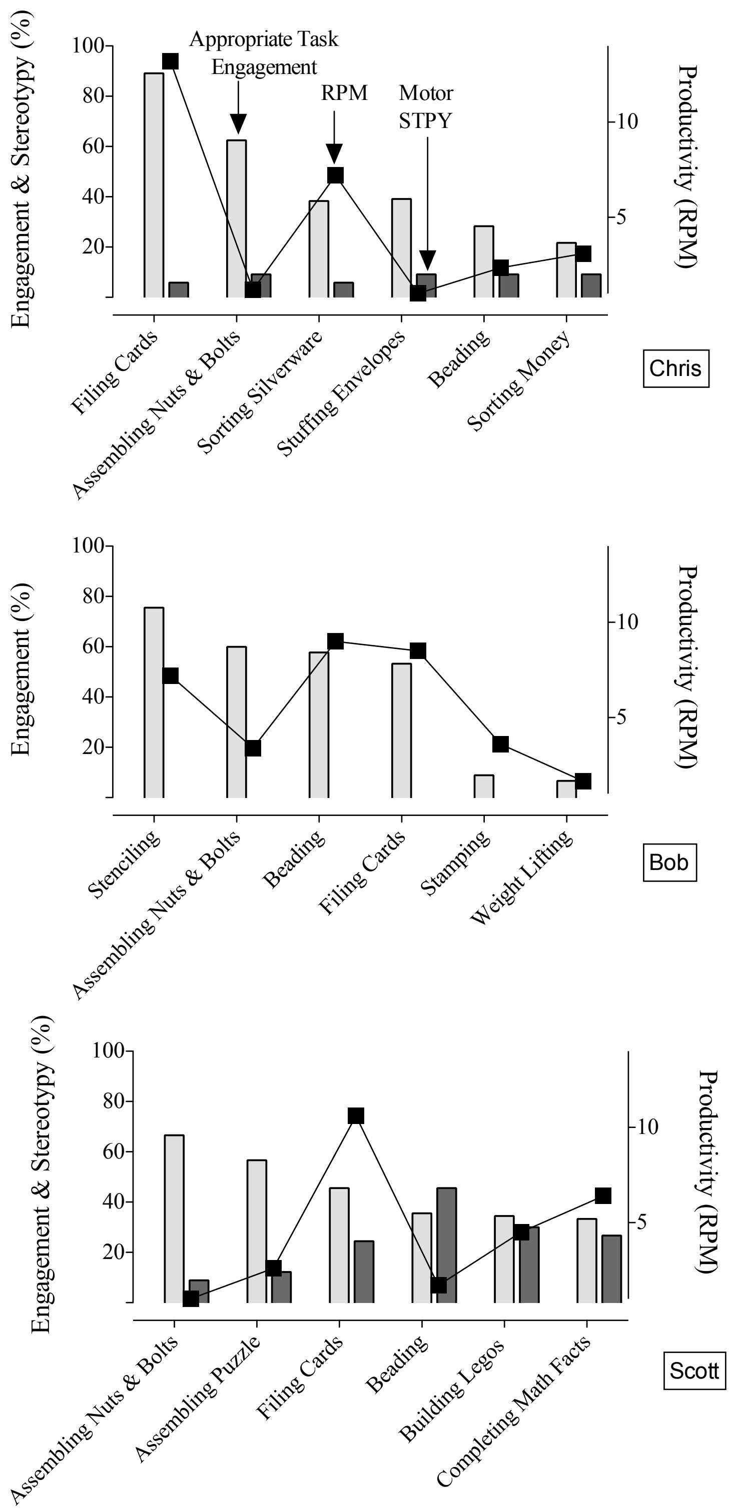


Figure 2. Percent occurrence of appropriate task engagement (grey bars) and motor stereotypy (black bars), and responses per min of productivity (closed squares) during the task assessment for Chris, Bob, and Scott.

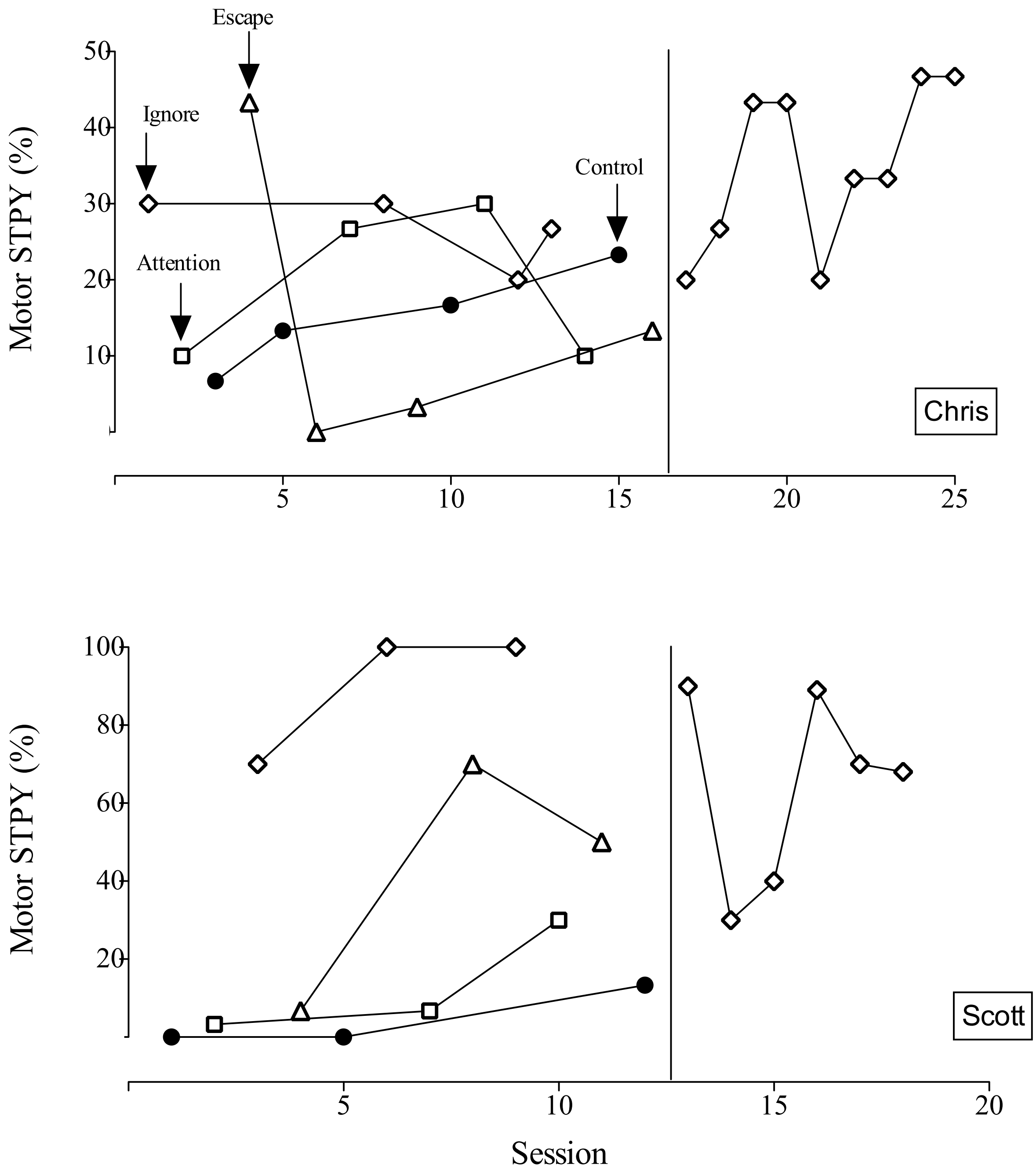


Figure 3. Percent occurrence of stereotypy across the attention (open squares), escape (open triangles), alone (open diamonds), and control (closed circles) conditions of the functional analysis for Chris and Scott.

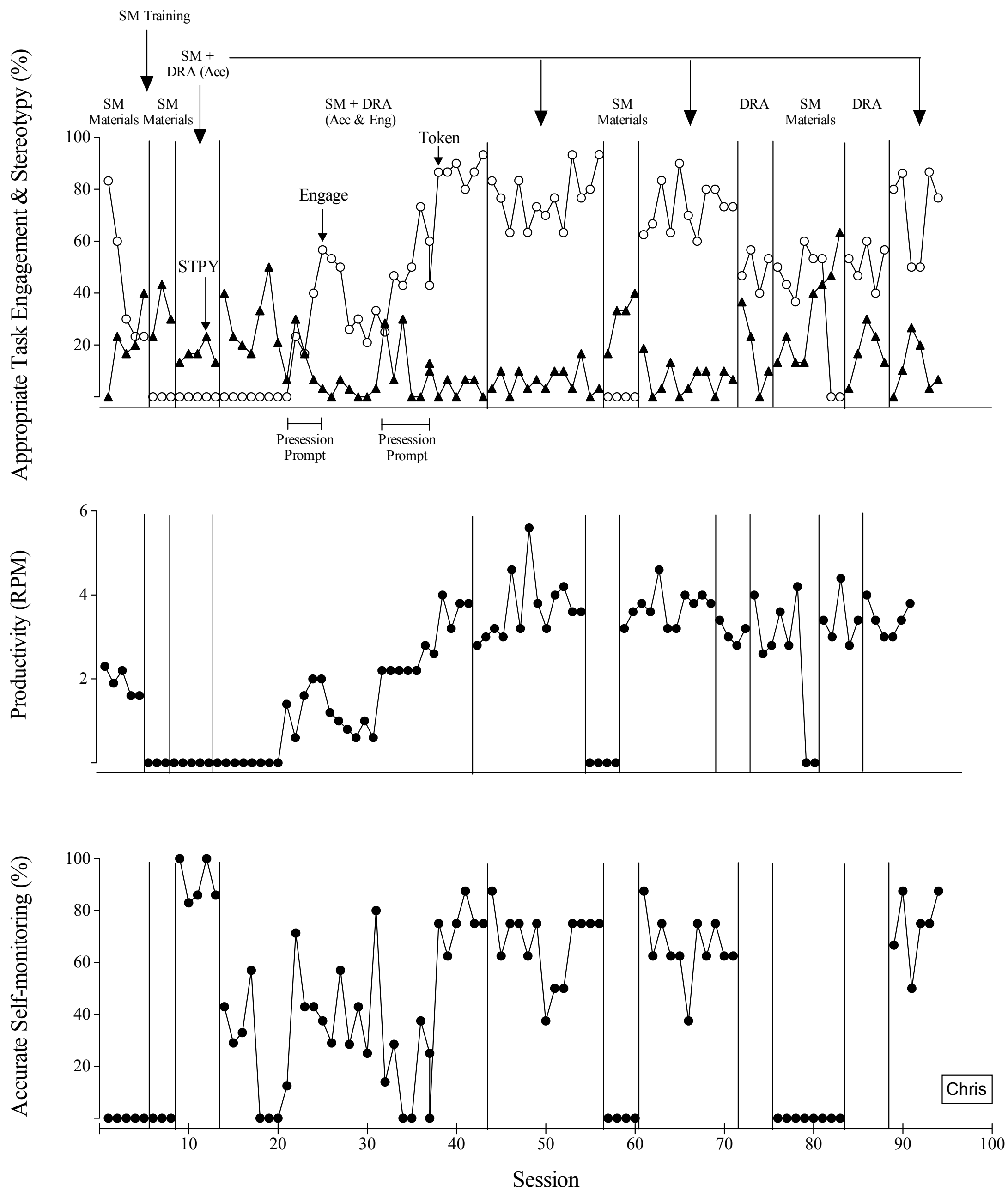


Figure 4. The top panel depicts percent occurrence of appropriate task engagement (open circles) and motor stereotypy (closed triangles), the middle panel depicts responses per minute of productivity, and the bottom panel depicts percent accurate self-monitoring across conditions of the treatment component analysis for Chris.

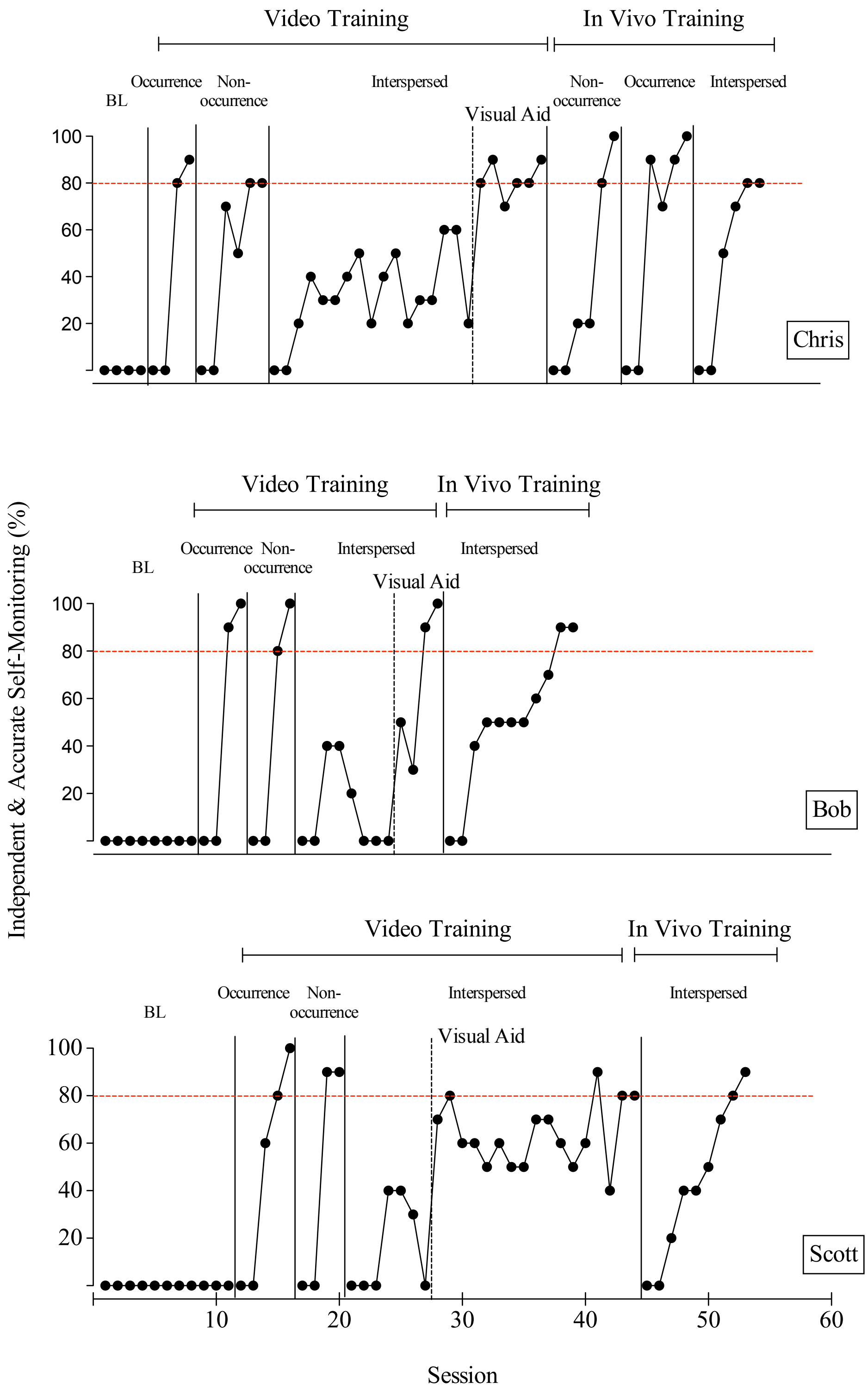


Figure 5. Percent correct and independent self-monitoring across phases of the video and in-vivo self-monitoring training for Chris, Bob, and Scott

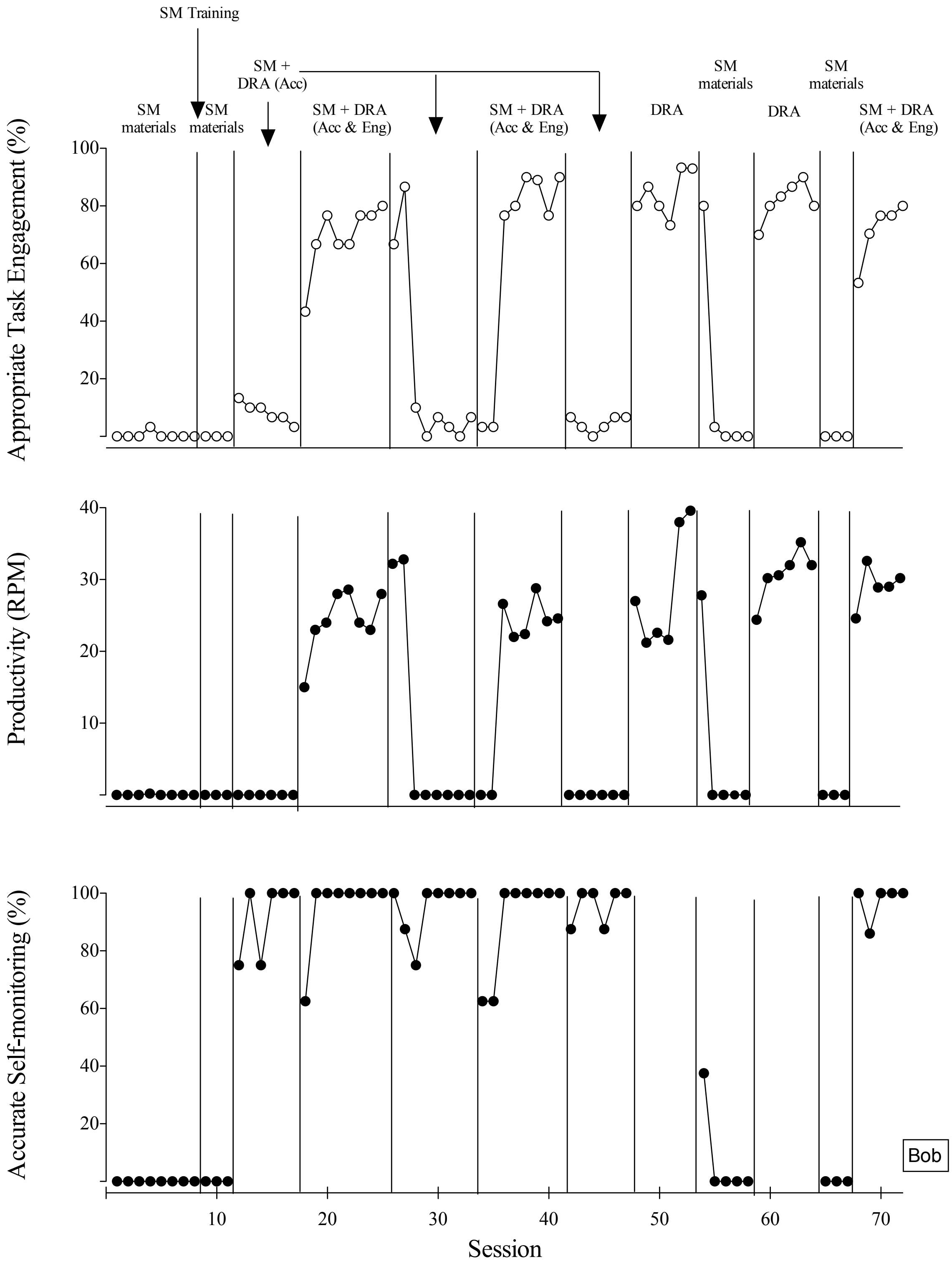


Figure 6. The top panel depicts percent occurrence of appropriate task engagement, the middle panel depicts responses per minute of productivity, and the bottom panel depicts percent accurate self-monitoring across conditions of the treatment component analysis for Bob.

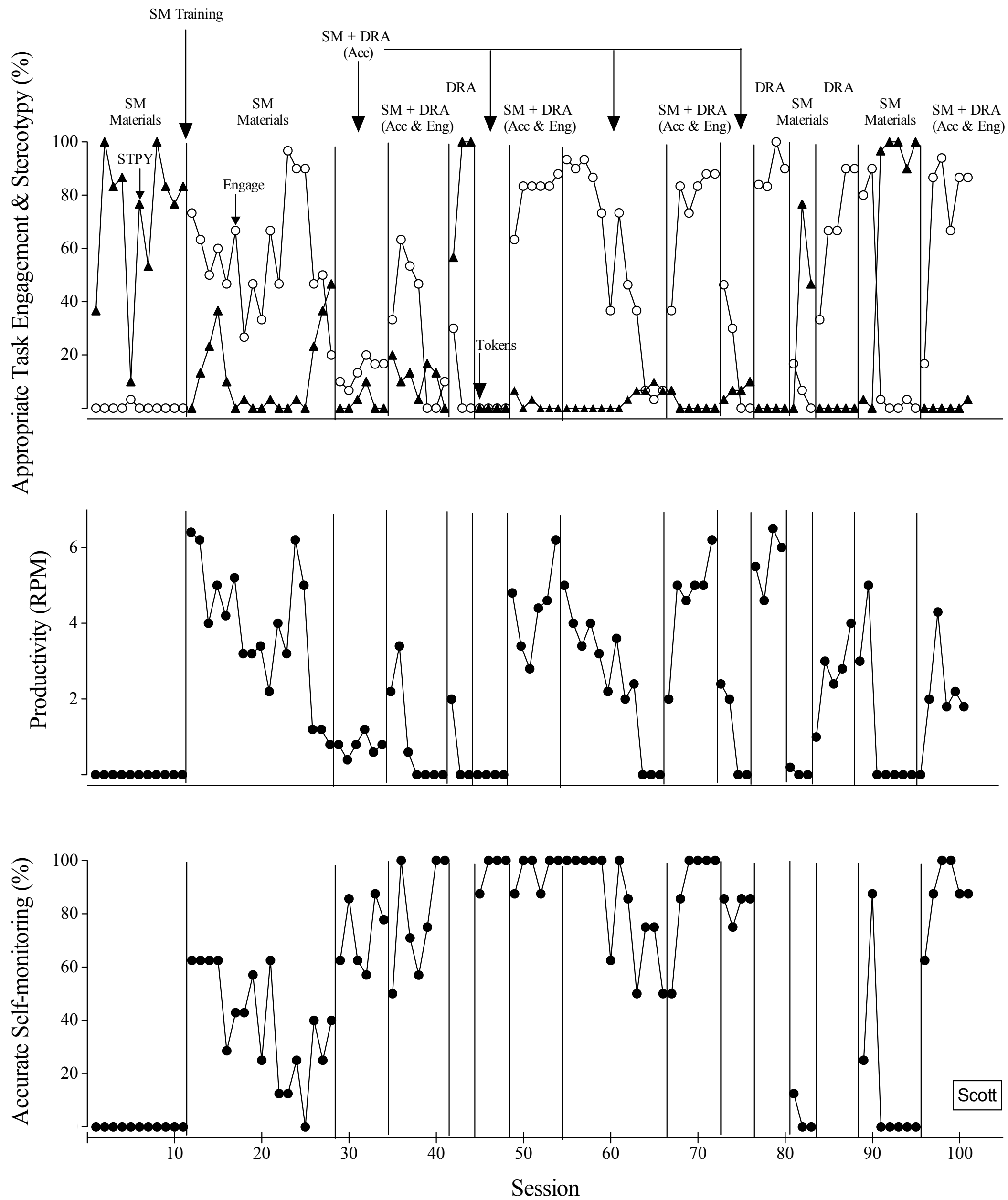


Figure 7. The top panel depicts percent occurrence of appropriate task engagement (open circles) and motor stereotypy (closed triangles), the middle panel depicts responses per minute of productivity, and the bottom panel depicts percent accurate self-monitoring across conditions of the treatment component analysis for Scott.

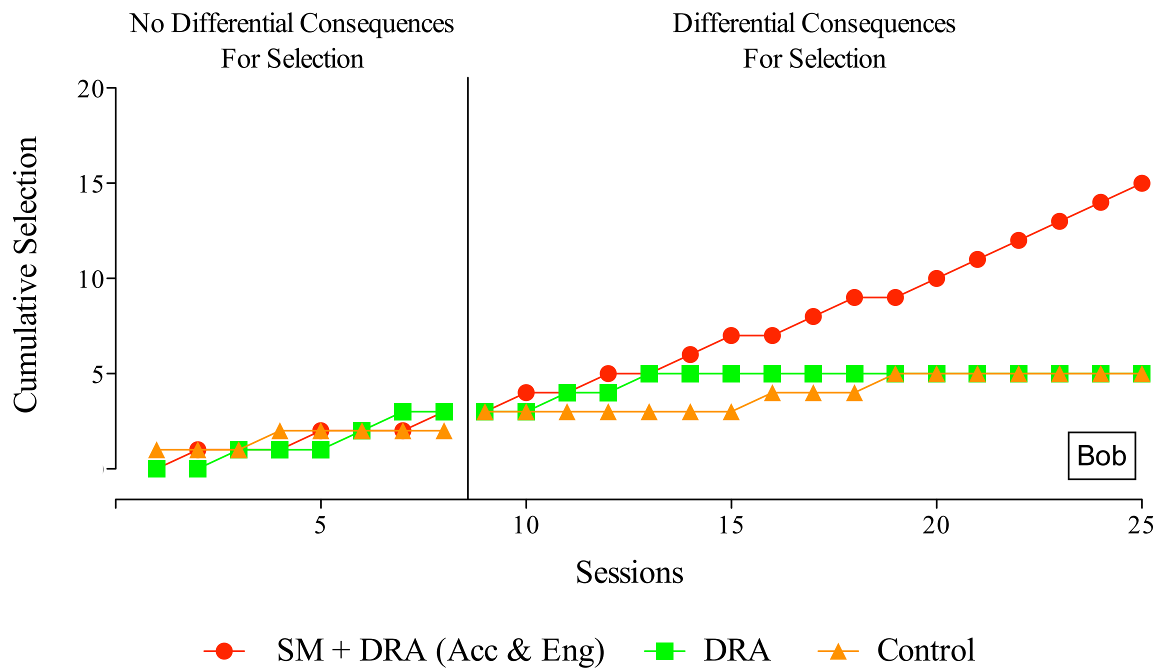
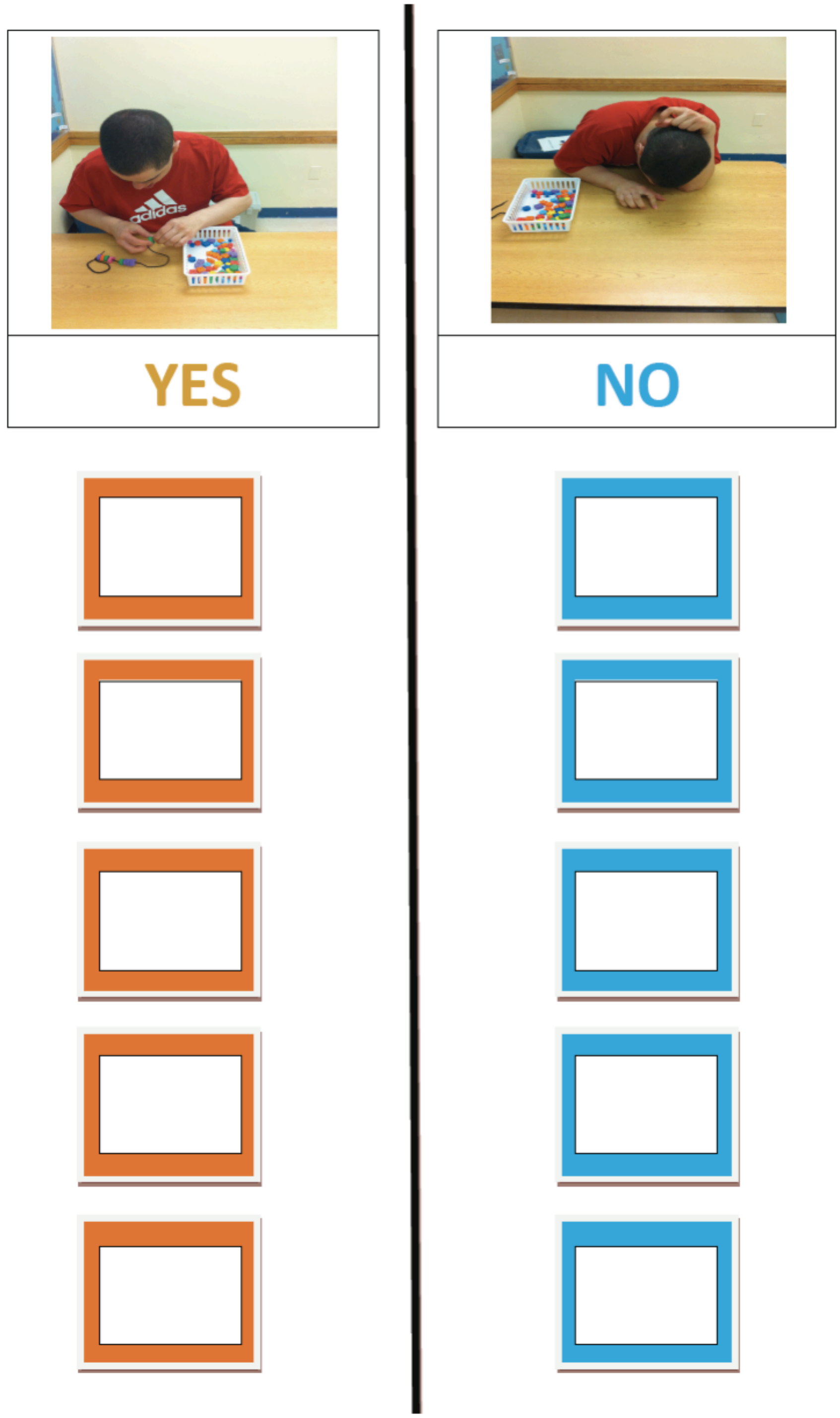


Figure 8. Cumulative number of selections for the no differential consequence for selections sessions and differential consequences for selections in the form of treatments for appropriate task engagement during the treatment preference assessment for Bob.



Appendix A. Sample self-monitoring data sheet used by the participant. The left column depicts the occurrence of appropriate task engagement, and the right column depicts the occurrence of off-task behavior. Columns are color-coded to enhance discrimination across box types