Summer 2022

Using Scorecards to Improve the Performance of Behavior Technicians in an Autism Treatment Clinic

Carly M. Ruether
Missouri State University, Carly122@live.missouristate.edu

USING SCORECARDS TO IMPROVE THE PERFORMANCE OF BEHAVIOR TECHNICIANS IN AN AUTISM TREATMENT CLINIC

A Master's Thesis

Presented to

The Graduate College of

Missouri State University

In Partial Fulfillment

Of the Requirements for the Degree

Master of Science, Applied Behavior Analysis

By

Carly Marie Ruether

August 2022

USING SCORECARDS TO IMPROVE THE PERFORMANCE OF BEHAVIOR TECHNICIANS IN AN AUTISM TREATMENT CLINIC

Psychology

Missouri State University, August 2022

Master of Science

Carly Marie Ruether

ABSTRACT

Staff performance is important to the success of businesses and employers often use different methods to increase staff performance and motivate their staff. Performance scorecards, goal setting, and feedback are some methods of increasing staff performance. The goal of the current study was to partially replicate and extend Griffin et al.'s (2019) study utilizing performance scorecards to increase targeted behaviors of employees. In the current study, pre-intervention survey results were used to inform the choice of intervention based on deficiencies. Based on the results of the pre-intervention questionnaire, it was concluded that "task clarification and prompting" and "performance consequences, effort, and competition" were deficiencies. The scorecard intervention was linked to reinforcers from the beginning of the study, and in Phase 2 of the intervention, two contingencies were in place, an individual contingency as well as a group contingency. Graphic feedback was also utilized weekly. Interpretation of the results suggest that this intervention was effective in increasing scorecard scores, especially in Phase 2 when specific contingencies and goals were in place. A reversal and one-month follow-up were conducted, showing that intervention effects were maintained.

KEYWORDS: scorecards, performance, goal setting, intervention, feedback, reinforcers, pre-intervention questionnaire

USING SCORECARDS TO IMPROVE THE PERFORMANCE OF BEHAVIOR TECHNICIANS IN AN AUTISM TREATMENT CLINIC

By

Carly Marie Ruether

A Master's Thesis
Submitted to the Graduate College
Of Missouri State University
In Partial Fulfillment of the Requirements
For the Degree of Master of Science, Applied Behavior Analysis

August 2022

Approved:

Michael Clayton, Ph.D., Thesis Committee Chair

D. Wayne Mitchell, Ph.D., Committee Member

Amber Abernathy, Ph.D., Committee Member

Julie Masterson, Ph.D., Dean of the Graduate College

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Staff performance is vital to the effectiveness and success of any business (Plowman & Bailey, 2005). Staff must meet deadlines, provide good customer service, work as a team when necessary, be on-time to work, and complete their job duties appropriately to ensure that the business runs smoothly. Employers use a variety of different strategies to motivate their employees to perform well at work. These strategies can include raises (or other positive consequences), threats of losing their job (or other negative consequences), and verbal and/or written feedback. A prerequisite for successfully providing these consequences is a performance management system (Daniels, 2014).

LITERATURE REVIEW

Both goal setting and performance feedback have frequently been used to great effect in the performance management (PM) literature. Feedback generally consists of information that provides details on an individual's past performance, often in relation to a goal (Prue & Fairbank, 1981). Many studies have been conducted that have applied goal setting or feedback to increase performance, often in combination with each other, and have produced differing results. Pritchard et al. (1987) conducted a study using an intervention that included feedback, goal setting, and incentives to increase the productivity of organizational units at an Air Force base and found that feedback increased productivity significantly, and goal setting continued to increase productivity once introduced. Anderson et al. (1988) also used feedback and goal setting along with praise in their intervention to increase the hit rate of a university hockey team. Their results displayed that feedback produced a larger increase in performance than goal setting or praise. Wilk and Redmon (1990) and Ralis and O'Brien (1987) also found success in using goal setting and feedback to increase employee performance. Another study that employed the use of both feedback and goal setting conducted by Kim and Hamner in 1976 discovered that for some targeted behaviors, goal setting plus feedback produced larger increases in performance than only goal setting. Calpin et al. (1988) utilized self-monitoring as a form of feedback in which clinicians tracked their own session start and end times and then added goal setting to increase staff productivity in a mental health center. They observed significant increases during the self-monitoring (feedback) phase and only small increases during the feedback plus goal-setting phase. Unlike the other studies, Loewy and Bailey (2007) did not find significant increases when using an intervention including graphic feedback (graphs of group performance), goal setting,

and praise for customer service behaviors and saw only very modest improvements that did not maintain. The authors suspected this to be due to the absence of valuable consequences linked to the intervention.

As with many of the previous studies, other studies have applied feedback in conjunction with other treatment components to increase staff performance in various settings. Balcazar et al. (1985) in their review of performance feedback found that feedback, when used by itself, was associated with the least consistent effects and the highest number of no-effects and mixed effects. They also found that feedback when used in tandem with consequences or goal setting was much more effective than using feedback alone. Alvero et al. (1985-1998) also reviewed performance feedback and confirmed the previous findings of Balcazar et al. (1985). Austin et al. (2005) implemented an intervention package in a restaurant that included the use of checklists and both verbal and graphic feedback. The results indicated that the initial intervention of checklists and verbal feedback increased task completion but that the targeted behavior increased even more when graphic feedback was added. Richman et al. (1988) used self-monitoring with added feedback in a care facility for disabled individuals, and feedback was effective in increasing staff performance when the self-monitoring phase improvement effects did not maintain. Crowell et al. (1988) conducted a study using an intervention for bank tellers that included task clarification as well as feedback and praise. Feedback increased performance but not as quickly or significantly as did task clarification. In 1993, Hawkins et al. conducted a study that implemented an intervention consisting of verbal and written feedback to geriatric nursing assistants on completing assigned prompted voiding. Compliance of nursing assistants increased when written feedback was added to the verbal feedback already in place. Fox and Sulzer-Azaroff et al. (1990) investigated differences in who gave employees written feedback on

performance between supervisors and non-supervisors and found that feedback was effective no matter who delivered the feedback. Allison et al. (1993) utilized feedback and different incentive systems to employees in a treatment center and found that performance was increased during all incentive conditions as compared to feedback by itself, although it was not a significant difference. Brown et al. (1980) utilized training and feedback to increase customer-service behaviors in department stores and feedback significantly improved performance of all targeted behaviors. Brown et al. (1981) used an intervention consisting of verbal feedback and feedback plus approval to decrease off-task behaviors while at work. Verbal feedback did result in noticeable decreases in off-task behaviors.

Performance scorecards are another way of tracking employee behavior and communicating the feedback clearly to employees. Scorecards allow the employee and manager to track agreed upon performance indicators that also provide the employer a basis for providing feedback. The performance indicators are reviewed periodically, and goals adjusted as needed. Scorecards track behaviors that are important to the employees' performance success and therefore affect the success of the business. Scorecards give specific information to the employee and employer on areas of improvement and how much improvement can be made. Eikenhout and Austin (2005) used graphic feedback, goal setting, praise, and a performance matrix to provide feedback to employees and assess the effects of the matrix and additional intervention components on staff performance measures. Prior to beginning, the authors administered the Performance Diagnostic Checklist (PDC) to employees. The PDC is used to identify variables that might impact poor performance in organizational settings (Austin, 2000). The assessment suggested that effective consequences were the main deficiency. The performance matrix measured 14 customer service behaviors across 115 staff members at a department store. The

participants were in three different categories based on departments of the store and consisted of cashiers, “Hardlines” sales, and “Softlines” sales. Graphic feedback was used in the form of graphs depicting performance in the first intervention phase (Phase B). During the second intervention phase (Phase C), an intervention package was implemented that included the graphic feedback as well as goal setting, verbal feedback/praise, and an incentive contingency. Results showed that behaviors decreased when feedback was removed. All departments showed an increase in desirable staff performances during the intervention phases. Greeting customers with a smile increased by 30.7%, eye contact increased by 64.9% (softlines) and 49.9% (hardlines), and friendly small talk increased by 29.3%. The performance matrix and graphic feedback were linked in a way that prevented analysis of the contribution of either treatment alone.

In an unpublished dissertation, Plowman & Bailey (2005) utilized a performance matrix with incentives as well as goals and graphic feedback to increase beneficial customer service behavior at a retail establishment. The relative weights assigned in the matrix were also manipulated. The author found that a weight increase of 15 or more points consistently resulted in a 10 percent increase in the corresponding behavior. Helpful behaviors increased when their respective matrix weights increased and decreased when weighting decreased. This study was the first to experimentally manipulate the relative value or weights of individual behaviors. It is difficult to make firm statements about any one independent variable due to the package intervention, but employee performance behaviors did improve during the intervention.

In 2012, Szabo et al. conducted a study that also employed performance scorecards. Using a package intervention consisting of the scorecards, service review, training, public feedback, and reinforcement, the authors attempted to improve staff performance. Graphs were

used during biweekly meetings and the service review performance scorecards were scored during meetings. Consumable reinforcers were provided at the end of designated meetings if a team scored at or above criterion for two consecutive months. Time to criterion lasted from five to 20 weeks, depending on the group of staff. Criterion was set at 1000 points or 100%. Significant, sustained staff behavior changes resulted from the implementation of service review and scorecards. Significantly, positive consumer outcomes were also correlated with staff performance improvements. After intervention, staff scores remained at criterion for the remainder the study and were maintained at subsequent 3-month, 6-month, and one-year probes. In 2019, Griffin et al. conducted a study that used performance scorecards along with a leveled lottery system to increase the behaviors of behavior technicians at two locations of the same organization. The performance scorecards consisted of five differentially weighted behaviors, including attendance (weighted 30%), average trial count (weighted 25%), cleaning duties (weighted 15%), goal completion (weighted 20%), and time/attendance sheet (weighted 10%). Each behavior was rated on a scale ranging 4-13 points in which a score of 7 was average and 10 was the goal. A lottery component was later implemented and linked to the scorecards. Scorecard scores increased approximately 10% in both locations after introduction. Scorecard feedback and the added lottery increased scores above that of scorecards by themselves. Social validity surveys showed that many participants reported that the scorecards took more than a reasonable amount of time to complete but that they were useful and improved the staff's ability to do their jobs well and that they would like to continue using the scorecards after the study was completed. There were several limitations in this study. First, staff were already performing at high levels, so a ceiling effect was evident. Another weakness was the initial employee resistance

to scorecard feedback. Finally, some variables that affected scoring were not under the employee's control.

The present study partially replicated and extended Griffin et al's (2019) study. This study sought to reduce potential staff resistance, which was a limitation stated in the Griffin study, by implementing the scorecard intervention linked to graphic feedback, setting a goal, and earning reinforcers from the beginning. This study also placed higher assigned weights to the behaviors that have the greatest opportunity for improvement and are more important to job success. Another extension is the use of a pre-intervention assessment (PDC) to guide treatment selection. The purpose of the present study was to improve Registered Behavior Technicians' (RBT) job performance across five different criteria including timeliness, preparedness for sessions, wearing nametags, monitoring clients' goals, and completing documentation using a scorecard intervention, graphic feedback, and reinforcement in a clinic that provides ABA therapy to clients with diagnoses of autism spectrum disorder.

METHODS

Participants

Participants of this study were 15 Registered Behavior Technicians (RBTs). The RBTs included 13 females and 2 males, ranging in age from 21-30 (mean age of 24.5). One RBT had their G.E.D., one had their associate degree, five were current college students, seven had their bachelor's degrees, and one had their master's degree.

Three of the RBTs ended their employment during the study and seven RBTs began employment during the study. The rest of the RBTs were employed at the clinic throughout the entire duration of the study. See Table 1 for detailed information about the participants.

Overall, the clinic that served as the setting for this study employed a Clinic Director, a Lead Behavior Analyst (BCBA), and the 15 behavior technicians described above.

Setting

This study was conducted at a local clinic in southwest Missouri that provides therapy to children (2- to 12-years-old) primarily with diagnoses of autism spectrum disorder (ASD). The RBTs in this study worked in several locations, including the clinic, a local private elementary school, and occasionally in the client's home.

Materials

The behavior technicians were given an electronic copy of their scorecard each week that depicted their scores on five different criteria. See Figure 1 for an example of the scorecard that RBT's received each week. RBTs also received an electronic thermometer graphic each week

depicting their weekly group average scorecard score. See Figure 2 for an example of the thermometer graphic. Reinforcers were utilized including bottled drinks, fountain drinks, candy bars, small snacks, five-dollar gift-cards, and a lunch meal.

Pre-Intervention Assessment

To assess possible relevant variables among the staff, the Performance Diagnostic Checklist-Human Services (PDC-HS) was provided to the RBT's by the clinic director. Each RBT completed the PDC-HS and returned it to the clinic director. Any questions that the RBTs had were directed to the researcher. See Appendix A. Results (Figure 3) showed that "task clarification and prompting" and "performance consequences, effort, and competition" categories were the variables contributing the most to poor performance levels. The categories with the highest percentages show the variables that are most deficient. The category of "equipment and processes" and "knowledge and skills" were less of a concern because they scored 0% and 6%, indicating that these were not areas of deficiency. Equipment and processes included materials and items used for therapy, as well as the treatment rooms and physical space. Knowledge and skills were likely an area of mastery due to RBTs going through training and getting feedback during supervision each month.

Measures

The primary dependent variable in this study was the weekly scorecard score of each participant. The five behaviors tracked on the scorecards were 1) wearing nametags, 2) timeliness, 3) preparedness for sessions, 4) completed documentation, and 5) monitoring of client goals.

Wearing Nametag. RBTs were required to wear their nametags in a visible location on their shirt or pants upon entering the work site. To ensure that RBTs received their points for this behavior due to being at different locations, RBTs could send a picture of their nametag on their shirt or pants to the clinic director or lead BCBA. Points were still given if the RBT was observed wearing their nametag. This measure was scored on a scale of 4-10, with 10 being the goal. A score of 4 denoted that a RBT wore their nametag for 40% of their sessions. A score of 7 denoted that a RBT wore their nametag for 70% of their sessions. A score of 10 denoted that a RBT wore their nametag for 100% of their sessions. This measure was assigned a weight of 10%.

Preparedness for Sessions. At the beginning of each clinic session, RBTs were required to have their work-issued iPad completely charged as well as access to their keyboard charger and iPad charger, if needed. RBTs at both sites (clinic and school) were required to have all necessary equipment in their possession before their session began. This equipment included session materials used for therapy goals and client's communication devices, if needed. For this target behavior, RBTs were required to send the clinic director or lead Board Certified Behavior Analyst (BCBA) a screenshot of their iPad screen each morning showing that their iPad was fully charged (100%) by 9:30 a.m. As there were only five opportunities to complete this behavior each week (M-F), scoring was slightly different than other measures. This measure was measured on a scale of 4-10, with 10 being the goal. A score of 4 denoted that a RBT has sent a screenshot of their iPad with 100% charge for 20% of sessions. A score of 5 denoted that a RBT has sent a screenshot of their iPad with full charge for 40% of sessions; 6 denoted full charge for 60% of sessions. A score of 7 denoted that an RBT sent a screenshot of their iPad with 100% charge for 60% of sessions and all screenshots were sent on time. A score of 8 denoted that an

RBT sent a screenshot of their iPad with full charge for 80% of sessions. A score of 9 denoted that an RBT sent a screenshot with full charge for 80% of sessions on time. A score of 10 denoted that an RBT sent a screenshot of their iPad with 100% charge for 100% of session and all screenshots were sent on time. This measure was assigned a weight of 25%.

Timeliness. RBTs were required to clock-in and clock-out of their sessions each day using the "Microsoft Teams Shifts" application on their phone. When an RBT arrived at their session site, they clocked in. RBTs generally worked two sessions per day with a 30-minute break for lunch in the middle of their day. RBTs were required to be at their session site a minimum of five minutes prior to their session. To earn points for being "on time," an RBT would clock in at (or before) the start time of their session. Timeliness was scored on a scale of 4-10, with 10 being the goal. A score of 4 denoted that a RBT clocked in on-time for 40% of sessions. A score of 7 denoted that a RBT clocked in on-time for 70% of sessions. A perfect score of 10 denoted that a RBT clocked in on time for 100% of sessions. This target behavior was assigned a weight of 20%.

Completed Documentation. RBTs were required to complete documentation for each session they conducted. RBTs used a data collection system on the Rethink application (rethinkbehavioralhealth.com) for entering ABC data (data detailing the antecedent of a behavior, a description of behavior, and the consequence or what occurred immediately after the behavior) when necessary, and session notes. Digital data sheets were to be completed the day of each session and signed online by the RBT. Completed documentation was scored on a scale of 4-10, with 10 being the goal. A score of 4 denoted that an RBT completed 40% of documentation on time. A score of 7 denoted that an RBT completed 70% of documentation on

time. A score of 10 denoted that an RBT completed 100% of documentation on time. This target behavior (completed documentation) was assigned a weight of 25%.

Self-Monitoring of Goals. RBTs were required to specify goal progress in their daily session notes. In the Rethink application session note, there was a section titled "Summary of progress towards treatment goals." RBTs were required to specify the progress of at least three client goals in each session note. The Rethink application had a different kind of session note option for when an RBT conducted a "pairing" session. This kind of session note did not have a section for the summary of progress towards treatment goals. To still earn points for this target behavior, RBTs were required to discuss three goals worked on (if they worked on goals which is often not completed during a pairing session), prompting levels required to complete tasks, or describing at least three specific activities that the RBT engaged in with his or her client during the pairing session. Monitoring of goals was scored on a scale of 4-10. A score of 4 denoted that a RBT summarized progress for at least three client goals for 40% of their session notes. A score of 7 denoted that a RBT summarized progress for at least three client goals for 70% of their session notes. A score of 10 denoted that a RBT summarized progress for at least three client goals for 100% of their session notes. This measure will have an assigned weight of 20%.

Procedure

A concurrent multiple baseline across tasks design was used to assess the effects of the independent variables. The first group of tasks tracked 1) timeliness, 2) wearing nametags, and 3) preparedness for sessions. The second group of tasks tracked 1) documentation and 2) monitoring client goals.

During baseline, participants completed their shifts as normal. During this time, the researcher discreetly completed scorecards for each of the RBTs each week without sharing the scorecards with the RBTs. An overall total score was calculated by summing the five monitored behaviors after they are weighted. A questionnaire assessing reinforcers was distributed to all RBTs and kept on file. See Appendix B. Baseline lasted until stability was evident. Baseline for Group 1 behaviors lasted for 14 weeks, and baseline for Group 2 continued for 5 additional weeks, totaling 19 weeks. Prior to data collection beginning, the institutional review board (IRB) reviewed and approved the proposed study. IRB-FY2021-517 was approved on May 10, 2021, and a subsequent modification was approved on February 9, 2022. See Appendix C.

Intervention Phase 1

After baseline was completed, the clinic director sent a message to the RBTs explaining the scorecards via "Microsoft Teams." RBTs were asked to review the information and confirm with the clinic director that they had done so. At the end of each week, the researcher completed the scorecards. Due to RBTs working at different sites, the scorecards were sent out electronically. Scorecards were distributed to RBTs the following week.

The individual scorecard scores of each RBT were then averaged together as a weekly average group score. This score was displayed on a thermometer graphic including the weekly average score, previous scores, and the percentage of the goal reached for that week. This thermometer graphic was sent in a group message to all RBTs on the Microsoft Teams application after RBTs had received their individual scorecards. To reduce pressure and avoid negative feelings about the scorecards, formal goals were not set each week. RBTs received small reinforcers delivered by the BCBA or clinic director. These reinforcers included bottled

drinks, fountain drinks, candy bars, and small snacks. RBTs received reinforcers at the end of each week for the previous week's scorecard scores. Phase 1 lasted for 10 weeks.

Intervention Phase 2

To increase the weekly scorecard scores, two specific contingencies were implemented. A message was sent to all RBTs by the clinic director detailing the added contingencies. If an RBT reached a score of 800 or more (at least 80% of overall goal) on their individual scorecard, they would receive a reinforcer ($5 gift card) at the end of the following week when their scorecard is sent out. A group contingency was also in place. If RBTs collectively reached 80% of the goal (a score of 800 or above) for three consecutive weeks for the weekly group score, they earned lunch provided by the company. RBTs continued to receive electronic scorecards at the end of the next week and individual reinforcers if they met the established criteria. Individual reinforcers were provided to those who met criteria to ensure that RBTs were accessing reinforcement regardless of the group contingency. The thermometer graphic was sent out in the same way as in Phase 1. Phase 2 lasted for 3 weeks.

Following Intervention Phase 2, a reversal phase was implemented for 3 weeks. The intervention was not in place during this time. One month later, a one-week follow-up was conducted, and the intervention was still not in place.

Social Validity

At the end of the study, a questionnaire was distributed to RBTs to assess the acceptability of the procedures. The survey was modeled on the social validity questionnaire in Plowman & Bailey (2005). The questionnaire consisted of 4 questions rated on a Likert-type scale ranging

from Strongly Disagree to Strongly Agree. Questions assessed the 1) degree to which the incentive program motivated them, 2) if they felt they performed better on the different behaviors, 3) if they would like to continue using the intervention, and 4) if it is generally easier to get their work done each day. See Appendix D.

Interobserver Agreement

Interobserver agreement was collected by a research assistant for four of the five behaviors that were permanent product data and showed an average of 98% agreement. An agreement of 100% was reached for the fifth behavior (wearing nametag) and was collected for 30% of the data collected, partially permanent product data. IOA scores were measured by dividing agreements by agreements plus disagreements and multiplying by 100.

RESULTS

Figure 4 shows the percentage of completed target behaviors in each group of behaviors over the course of the study. Group 1 behaviors (top) included nametags, preparedness, and timeliness. During baseline, employee scorecards averaged about 46% of completed tasks. During the first intervention, the average percentage of completed tasks was 69%. The second intervention increased the average to 75.3% of completed tasks. Following the second intervention, a reversal phase lasted for three weeks. During this time the average percentage of completed tasks was 71.6%. A one-week follow-up was done one month after the end of the reversal phase and the average percentage of completed tasks was 70%. Group 2 behaviors (bottom) included 1) documentation and 2) monitoring. During baseline, employee scorecards averaged about 58.7% of completed tasks. During the first intervention, the average percentage of completed tasks was 74.6%. The second intervention increased the average to 89% of completed tasks. Following the second intervention, a reversal phase lasted for three weeks. During this time the average percentage of completed tasks was 78.7%. A one-week follow-up was done one month after the end of the reversal phase and the average percentage of completed tasks was 84%.

A comparison of the Group 1 behaviors during baseline (m=46.4, sd=1.91) and then during the first treatment (m=68.9, sd=3.64) resulted in a significant improvement t(12.56) = 17.87, p=.0001. Additionally, A comparison of the Group 2 behaviors during baseline (m=58.68, sd=3.90) and then during the first treatment (m=74.60, sd=5.32) resulted in a significant improvement $t(5.19) = 6.26$, $p = .0013$.

Results of the social validity questionnaire are shown in Figure 5. The average scores for the social validity survey were within the 3–4-point range out of 5 points with the statement “I believe the scorecard program helped me overall perform better at my job” receiving the lowest score and the statement “It is generally easier now to get my work done each day” receiving the highest score. The ranges for question 1 was 1-5 with a standard deviation of 2.1. The range for question 2 was 2 - 4 with a standard deviation of 2.9. The ranges for question 3 was 1 - 5 with a standard deviation of 2.3. The ranges for question 4 was 2 - 5 with a standard deviation of 2.7.

DISCUSSION

Following the introduction of scorecards, the percentage of required work behaviors completed rose 23% for Group 1 and 15.9% for Group 2. The addition of the added contingencies in Phase Two increased the percentage of required work behaviors 6.3% for Group 1 and 14.4% for Group 2. The independent variables were suspended for 3 weeks following Phase Two and the average percentage of required work behaviors remained high. This suggests that the behavior change had generality. A one-week follow-up one month later showed the behavior change remained strong. The employees were still completing most of the required work behaviors.

The scorecards were effective at increasing the percentage of required work behaviors and the staff were generally supportive of scorecard use. Anecdotally, the lead BCBA stated that she would like to continue using scorecards in some format moving forward due to the increase in the target behaviors.

The Griffin et al. (2019) study results exhibited increases in scorecard scores during their scorecard phase and scores increased even more when the lottery system was implemented. Similarly, to the current study, scorecard scores increased when reinforcement (or a higher level of reinforcement) was applied. The Griffin et al. (2019) study and the current study both exhibited that their intervention effects maintained with a follow-up data point showing high results. However, the Griffin study did not conduct a reversal phase and during the current study's follow-up point, the scorecards were no longer in place. The Griffin et al. (2019) social validity results were positive, with the majority agreeing that scorecards helped the employees perform well at their job and were useful, and the current study social validity results were

similar. Unlike the Griffin et al. (2019) study, the phase changes in the current study were based on the data reaching stability rather than dictated by the organization. Also, the current study did not see staff resistance when the scorecards were first introduced while the Griffin et al. (2019) study did report initial staff resistance.

LIMITATIONS

It is important to note that week 5 of the second phase of the intervention fell on a holiday week, with only 3.5 days of data (3 full days and 1 half day). Additionally, the first week of withdrawal was a shorter week. The follow-up data week was originally scheduled for the first week of February, but due to inclement weather there were only two days of data. The follow-up data week was instead collected the following week and the partial week data was included in this data point. It is also important to note that a documentation training that was not specifically associated with this project was conducted on 11/1 and could have affected the scores for the documentation and monitoring client goals behaviors.

One limitation of the study was that reinforcers were small due to limited funds. Reinforcer questionnaires were used to provide preferred choices of small reinforcers including various and drinks and snacks. Reinforcers may have had a larger effect had they been larger more reinforcing/valuable items. Second, the social validity questionnaire results were in the 3–4-point range out of 5 possible points. RBTs scored close to a score of 3 meaning "neutral" for if they would like to continue using scorecards, if they were motivated by the intervention, and if their job was easier when using scorecards. There was significant variability between individuals, and it was observed anecdotally that RBTs had varying opinions about the usefulness of the scorecards, which influenced their scorecard scores as well as their questionnaire results. Third, completing scorecards was time-consuming for the researcher and would require time set aside for clinic directors or BCBAs to complete. Future research should seek to use larger reinforcers when possible and simplify scoring or behaviors to reduce time required to complete scorecards. Additionally, scoring was not always equal across RBTs due to

some RBTs working part-time and some RBTs working full-time, meaning that they did not all have the same number of sessions or opportunities to earn points. In the future, an adjusted scoring system based on number of sessions could help with this.

The results of this study showed that scorecard interventions linked to a reinforcement system as well as graphic feedback can improve employee behaviors and that using pre-intervention survey methods may be beneficial in informing employers on what kind of intervention to use. It is important for businesses to provide feedback in some form to help employees understand their roles and expectations within a company. It may benefit companies to allocate time and resources to providing adequate feedback and reinforcement to employees to motivate them to improve their performance in their job.

REFERENCES

Allison, D. B., Silverstein, J. M., & Galante, V. (1993). Relative effectiveness and cost-effectiveness of cooperative, competitive, and independent monetary incentive systems. *Journal of Organizational Behavior Management*, *13*(1), 85–112. https://doi.org/10.1300/j075v13n01_06

Alvero, A. M., Bucklin, B. R., & Austin, J. (2001). An objective review of the effectiveness and essential characteristics of performance feedback in organizational settings (1985-1998). *Journal of Organizational Behavior Management*, *21*(1), 3–29. https://doi.org/10.1300/j075v21n01_02

Anderson, D. C., Crowell, C. R., Doman, M., & Howard, G. S. (1988). Performance posting, goal setting, and activity-contingent praise as applied to a university hockey team. *Journal of Applied Psychology*, *73*(1), 87–95. https://doi.org/10.1037/0021-9010.73.1.87.

Austin, J. (2000). Performance analysis and performance diagnostics. In J. Austin & J.E. Carr (Eds.), *Handbook of applied behavior analysis* (pp. 321-349). Reno, NV: Context Press.

Austin, J., Weatherly, N. L., & Gravina, N. E. (2005). Using task clarification, graphic feedback, and verbal feedback to increase closing-task completion in a privately owned restaurant. *Journal of Applied Behavior Analysis*, *38*(1), 117–120. https://doi.org/10.1901/jaba.2005.159-03

Balcazar, F., Hopkins, B. L., & Suarez, Y. (1985). A critical, objective review of Performance Feedback. *Journal of Organizational Behavior Management*, *7*(3-4), 65–89. https://doi.org/10.1300/j075v07n03_05

Brown, M. G., Malott, R. W., Dillon, M. J., & Keeps, E. J. (1980). Improving customer service in a large department store through the use of training and feedback. *Journal of Organizational Behavior Management*, *2*(4), 251–265. https://doi.org/10.1300/j075v02n04_02

Brown, K. M., Willis, B. S., & Reid, D. H. (1981). Differential effects of supervisor verbal feedback and Feedback Plus approval on Institutional Staff Performance. *Journal of Organizational Behavior Management*, *3*(1), 57–68. https://doi.org/10.1300/j075v03n01_05

Calpin, J. P., Edelstein, B., & Redmon, W. K. (1988). Performance feedback and goal setting to improve mental health center staff productivity. *Journal of Organizational Behavior Management*, *9*(2), 35–58. https://doi.org/10.1300/j075v09n02_04

Crowell, C. R., Anderson, D. C., Abel, D. M., & Sergio, J. P. (1988). Task clarification, performance feedback and social praise: Procedures for improving the customer service of bank tellers. Journal of Applied Behavior Analysis, 21, 65-71.

Daniels, A. C. (2014). *Performance management: Changing behavior that drives organizational effectiveness*. Performance Management Publications.

Eikenhout, N., & Austin, J. (2005). Using goals, feedback, reinforcement, and a performance matrix to improve customer service in a large department store. *Journal of Organizational Behavior Management*, *24*(3), 27–62. https://doi.org/10.1300/j075v24n03_02

Fox, C. J., & Sulzer-Azaroff, B. (1990). The effectiveness of two different sources of feedback on staff teaching of fire evacuation skills. *Journal of Organizational Behavior Management*, *10*(2), 19–36. https://doi.org/10.1300/j075v10n02_03

Griffin, M., Gravina, N. E., Matey, N., Pritchard, J., & Wine, B. (2019). Using scorecards and a lottery to improve the performance of behavior technicians in two autism treatment clinics. *Journal of Organizational Behavior Management*, *39*(3-4), 280–292. https://doi.org/10.1080/01608061.2019.1632241

Hawkins, A. M., Burgio, L. D., Langford, A., & Engel, B. T. (1993). The effects of verbal and written supervisory feedback on staff compliance with assigned prompted Voiding in a nursing home. *Journal of Organizational Behavior Management*, *13*(1), 137–150. https://doi.org/10.1300/j075v13n01_09

Kim, J. S., & Hamner, W. C. (1976). Effect of performance feedback and goal setting on productivity and satisfaction in an organizational setting. *Journal of Applied Psychology*, *61*(1), 48–57. https://doi.org/10.1037/0021-9010.61.1.48

Loewy, S., & Bailey, J. (2007). The effects of graphic feedback, goal setting, and manager praise on Customer Service Behaviors. *Journal of Organizational Behavior Management*, *27*(3), 15–26. https://doi.org/10.1300/j075v27n03_02

Plowman, J. M., & Bailey, J. S. (2005). Measuring and changing employee work behavior: an application of the performance matrix (dissertation).

Prue, D. M., & Fairbank, J. A. (1981). Performance feedback in organizational behavior management. *Journal of Organizational Behavior Management*, *3*(1), 1–16. https://doi.org/10.1300/j075v03n01_01

Pritchard, R. D., Jones, S. D., Roth, P. L., Stuebing, K. K., & Ekeberg, S. E. (1987). Feedback, goal setting, and incentives effects on organizational productivity. *PsycEXTRA Dataset*. https://doi.org/10.1037/e520532006-001

Ralis, M. T., & O'Brien, R. M. (1987). Prompts, goal setting and feedback to increase suggestive selling. *Journal of Organizational Behavior Management*, *8*(1), 5–18. https://doi.org/10.1300/j075v08n01_02

Richman, G. S., Riordan, M. R., Reiss, M. L., Pyles, D. A., & Bailey, J. S. (1988). The effects of self-monitoring and supervisor feedback on staff performance in a residential setting. *Journal of Applied Behavior Analysis*, *21*(4), 401–409. https://doi.org/10.1901/jaba.1988.21-401

Szabo, T. G., Williams, W. L., Rafacz, S. D., Newsome, W., & Lydon, C. A. (2012). Evaluation of the service review model with performance scorecards. *Journal of Organizational Behavior Management*, *32*(4), 274–296. https://doi.org/10.1080/01608061.2012.729408

Wilk, L. A., & Redmon, W. K. (1990). A daily-adjusted goal setting and feedback procedure for improving productivity in a university admissions department. *Journal of Organizational Behavior Management*, *11*(1), 55–75. https://doi.org/10.1300/j075v11n01_05

Table 1. Participant Information

RBT Initial	Sex	Age	Level of Education	Length of Employment at Clinic
A.A.	Female	24	Bachelor's degree	June 2021-September 2021
A.P.	Female	24	Current college student	March 2021-October 2021
C.B.	Female	25	Associate degree	October 2021-Present (Present for full study)
C.L.	Female	22	Bachelor's degree	March 2021-Present (Present for full study)
E.L.	Female	25	Master's degree	January 2020-Present (Present for full study)
E.R.	Female	23	Current college student	October 2021-Present
J.W.	Male	24	Current college student	November 2020-Present (Present for full study)
M.N.	Female	21	Bachelor's degree	April 2021-Present (Present for full study)
M.S.	Female	24	Current college student	October 2021-Present
O.S.	Female	28	Bachelor's degree	January 2021-October 2021
S.H.	Female	21	Current college student	November 2021-Present
S.M.	Female	23	Bachelor's degree	August 2021-Present
S.W.	Female	25	Bachelor's degree	July 2021-Present
T.W.	Female	30	GED	January 2021-Present (Present for full study)
Z.H.	Male	29	Bachelor's degree	August 2021-Present

Note. This table displays information about the participants in the current study including their sex, age, education level, and length of employment.

Name:								Date:		
Behavior	4	5	6	7	8	9	10	Weight	Raw score	Points
Nametag	40	50	60	70	80	90	100	10%		
Timeliness	80	100	120	140	160	180	200	20%		
Preparedness	100	125	150	175	200	225	250	25%		
Documentation	100	125	150	175	200	225	250	25%		
Monitoring goals	80	100	120	140	160	180	200	20%		
Perfect score=1000									Score	

Figure 1. Above shows an example of the scorecards used during the intervention phases of the current study. The top row of numbers (4-10) are the "raw scores" with the weighted points per each behavior in rows 3-7.

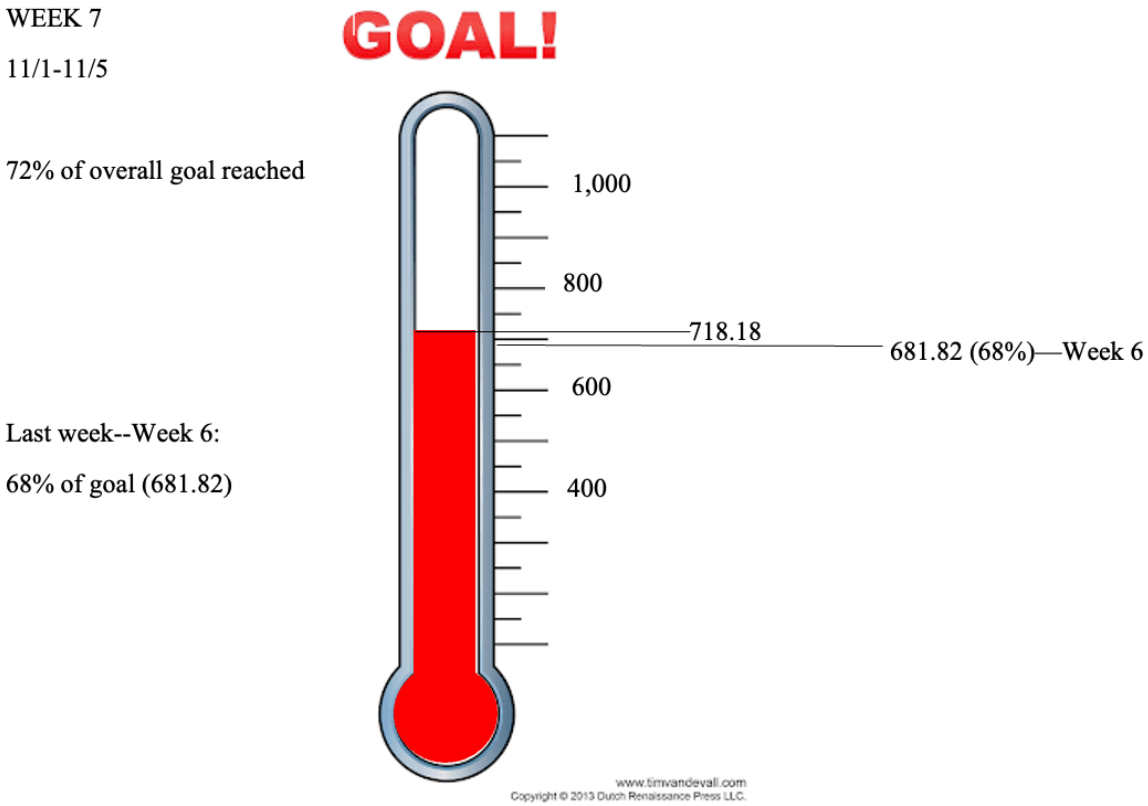

Figure 2. The above graphic is an example of the thermometer graphic used in the intervention phases of the current study. The graphic each week depicted the group's average scorecard score, the previous week's average scorecard score, and the percentage of the overall goal that was reached.

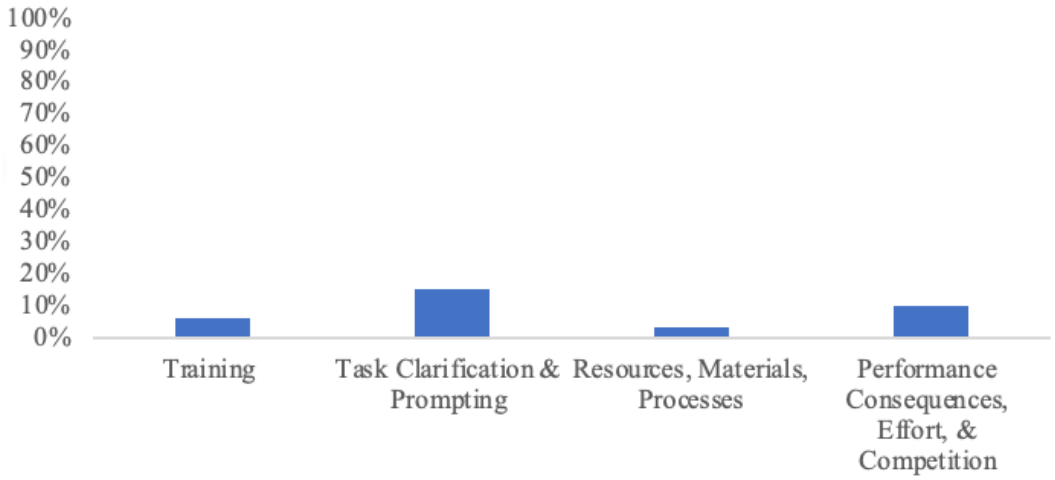

Figure 3. The above graph depicts the results of the Performance Diagnostic Checklist which exhibits that "task clarification and prompting" and "performance consequences, effort, and competition" were the areas of most deficiency.

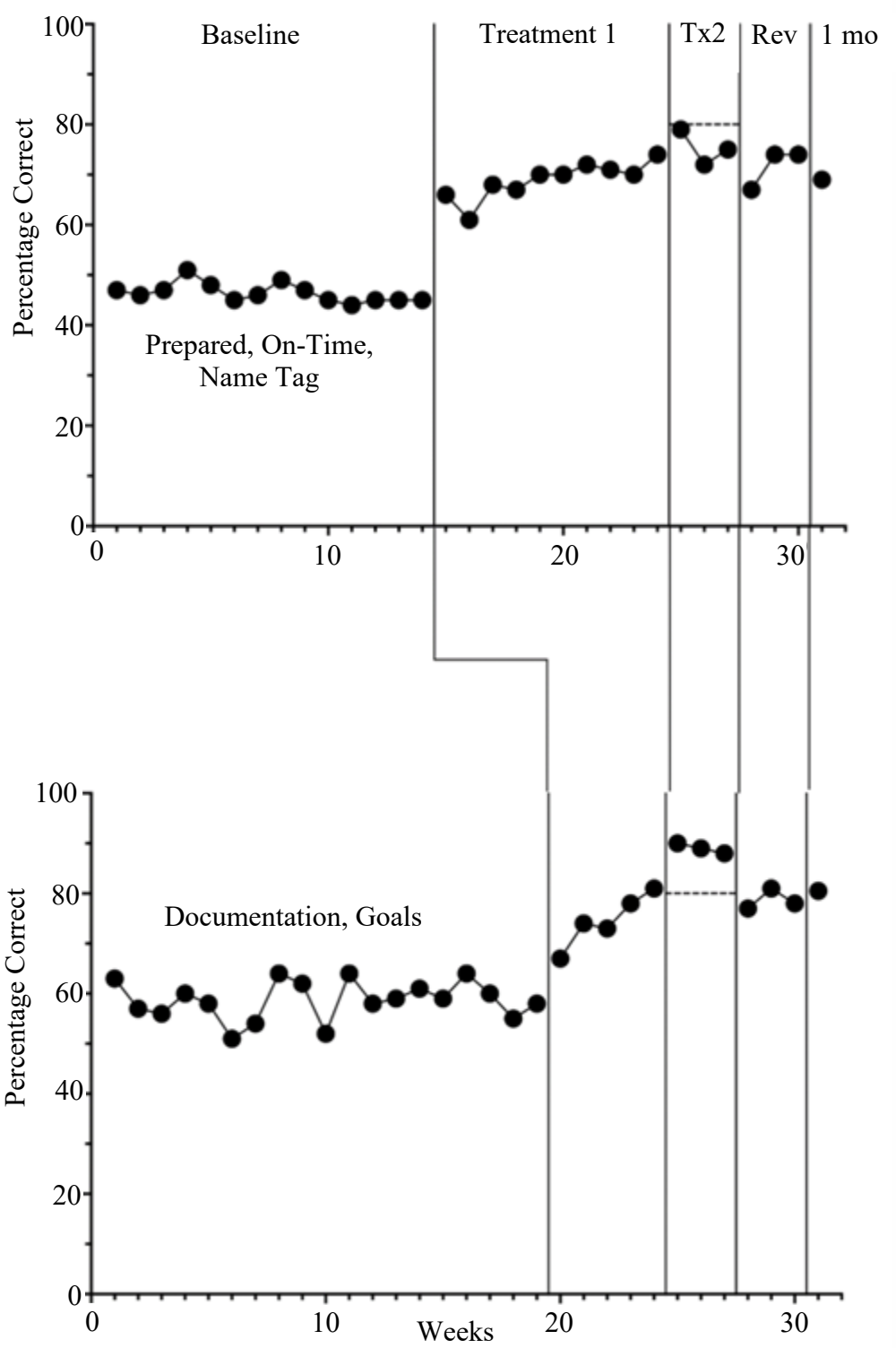

Figure 4. The above graph depicts scorecard scores across baseline, treatment 1, treatment 2, reversal, and a 1-month follow-up point with multiple baselines. Due to inclement weather during the follow-up week, there were only two days of sessions for the week. Because of this, a second week of follow-up the following week was collected in which a full week of data was collected.

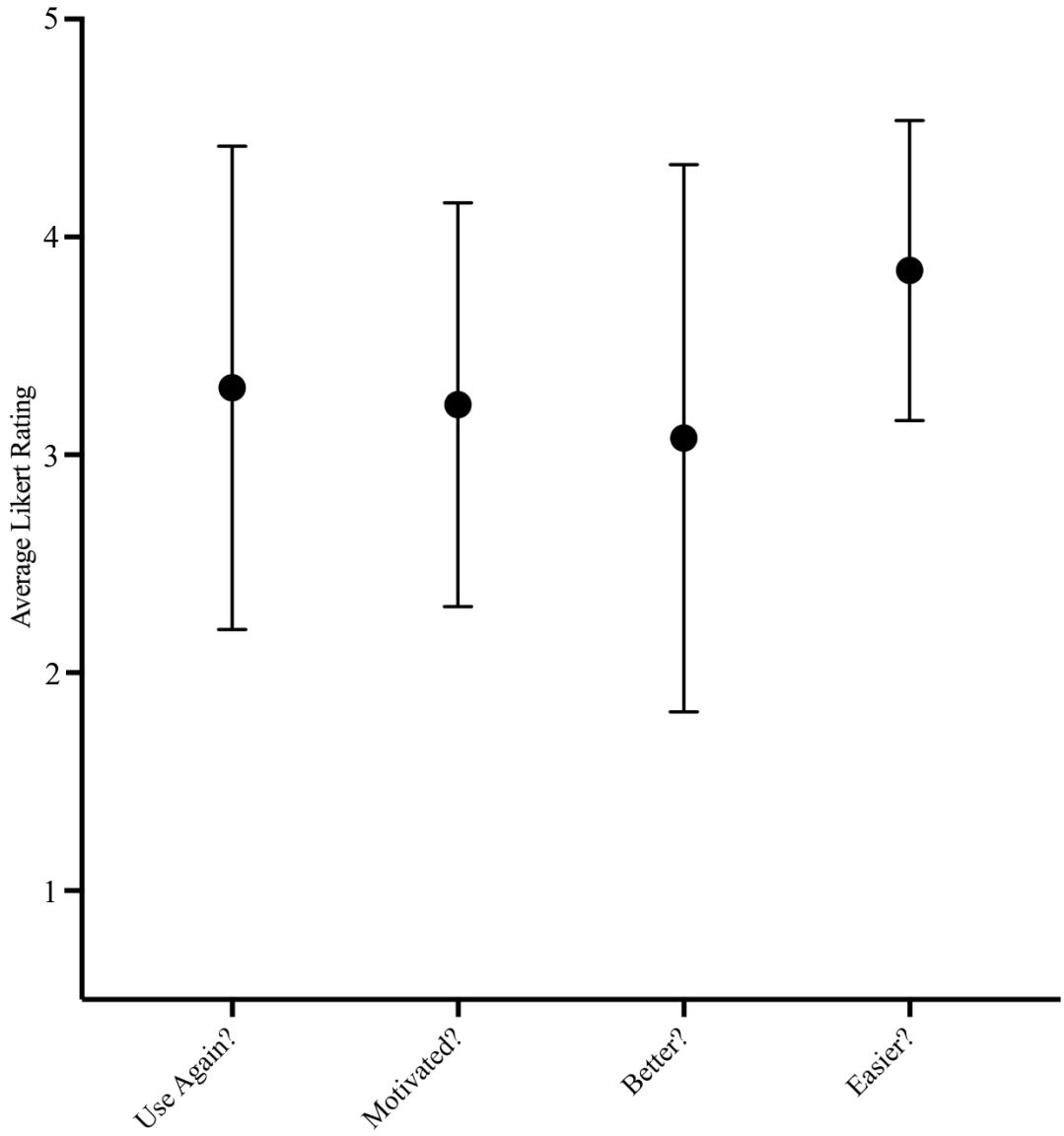

Figure 5. The above graph portrays the results of the social validity survey distributed to RBTs asking them four questions about the scorecard intervention. The whiskers represent the standard deviation around the means.

APPENDICES

Appendix A: Performance Diagnostic Checklist-Human Services

Answer each of the following questions, providing data in support of your answer if possible.

Training

Yes/No
O O Have you received formal training?
O O Can you accurately describe your job and when it is performed?

O O Is there evidence that you have accurately completed your duties previously?

O O If a task needs to be completed quickly, can you perform it at the appropriate speed?

Task Clarification & Prompting

Yes/No
O O Have you been informed of your supervisors' expectations for your job?

O O Can you state the purpose of your position?

O O Is a job aid (e.g., checklist, data sheet) for completing your duties visibly located in the task area?

O O Are you ever verbally, textually, or electronically reminded to complete your duties?
O O Are your duties being performed in an environment well-suited for task completion (e.g., not noisy or crowded?

Resources, Materials, & Processes

Yes /No
O O Are there enough trained staff available in the clinic?

O O If materials (e.g., teaching stimuli, preferred items, data sheets) are required for task completion, are they readily available (e.g., easy to find nearby?

O O Are the materials necessary to complete your duties well-designed for their intended purpose?
O O Are the materials necessary to complete your duties organized for their intended purpose?
O O Is the task suffering from other tasks not being completed first?

Performance Consequences, Effort, & Competition

Yes No

O O Is the employee ever directly monitored by a supervisor?

O O How often is the employee monitored?

-Hourly

-Daily

-Weekly

-Monthly

-Other

O O Does the employee ever receive feedback about their performance?

O O Does the employee ever see the effects of accurate task completion?

O O Does the duties of your job required a great deal of effort, or is it easy to perform?

O O Do other tasks appear to take precedence over the target task?

Appendix B: Reinforcer Survey

1. Please list your top 3 favorite places to receive a gift card to (could include restaurants, Starbucks, gas stations, etc.):
    1._____
    2._____
    3._____

2. Do you prefer being publicly recognized or privately recognized?
    a. Publicly—announced to team that you've done a good job/achieved something
    b. Privately—told one-on-one by a superior that you've done a good job/achieved something

3. Do you prefer words of affirmation or gifts?
    a. Affirmation
    b. Gifts

4. What is your favorite drink?
    a. Starbucks:
    b. Sonic:
    c. Gas station or store:

5. What is your favorite snack?

    _____

6. Would extra time with your supervisor be beneficial for you?

    _____
    _____

7. What is your favorite kind of sensory toy?

    _____

8. What therapy items would be helpful or enjoyable for you to have?

    _____

Appendix C: IRB Approval

To:
Michael Clayton
Psychology

RE: Notice of IRB Approval
Submission Type: Initial
Study #: IRB-FY2021-517
Study Title: Using Performance Scorecards to Improve the Performance of Registered Behavior Technicians
Decision: Approved

Approval Date: May 10, 2021

This submission has been approved by the Missouri State University Institutional Review Board (IRB). You are required to obtain IRB approval for any changes to any aspect of this study before they can be implemented. Should any adverse event or unanticipated problem involving risks to subjects or others occur it must be reported immediately to the IRB.

This study was reviewed in accordance with federal regulations governing human subjects research, including those found at 45 CFR 46 (Common Rule), 45 CFR 164 (HIPAA), 21 CFR 50 & 56 (FDA), and 40 CFR 26 (EPA), where applicable.

Researchers Associated with this Project:
PI: Michael Clayton
Co-PI:
Primary Contact: Carly Ruether
Other Investigators:

Appendix D: Social Validity Questionnaire

Please circle your answer using the following key:
1=strongly disagree 2=disagree 3=neutral 4=agree 5=strongly agree

1. I would like to continue the scorecard incentive program in some form.
   1 2 3 4 5

2. I feel the incentives offered in the incentive program motivated me.
   1 2 3 4 5

3. I believe the scorecard program helped me overall perform better at my job.
   1 2 3 4 5

4. It is generally easier now to get my work done each day.
   1 2 3 4 5

Additional comments:

1. _____
2. _____
3. _____