12-2019

The Effect of Social Comparison Feedback on Task Completion in a Human-Service Setting

Curtis Thomas Phillabaum

The Effect of Social Comparison Feedback on Task Completion in a Human-Service Setting

by

Curtis Thomas Phillabaum

A thesis submitted to the School of Behavior Analysis of
Florida Institute of Technology
in partial fulfillment of the requirements
for the degree of

Master of Science
in
Applied Behavior Analysis and Organizational Behavior Management

Melbourne, Florida
December, 2019

We the undersigned committee hereby approve the attached thesis, "The Effect of Social Comparison Feedback on Task Completion in a Human-Service Setting" by Curtis Thomas Phillabaum.

Rachel Tilka, Ph.D.
Assistant Professor
School of Behavior Analysis
Major Advisor

Darby Proctor, Ph.D.
Assistant Professor
School of Psychology

David Wilder, Ph.D., BCBA-D
Professor
School of Behavior Analysis

Nicholas Weatherly, Ph.D., BCBA-D
Associate Professor and Head
School of Behavior Analysis

Abstract

Title: The Effect of Social Comparison Feedback on Task Completion in a Human-Service Setting

Author: Curtis Thomas Phillabaum

Advisor: Rachael Tilka, Ph. D.

There has been much debate over the most successful types of feedback, but little research has examined the use of social comparison feedback. The purpose of the present study was to assess the effects of social comparison feedback on the percentage of end-of-shift cleaning tasks completed by three behavior technicians identified as exhibiting lower performance. The social comparison feedback consisted of a bar graph depicting the percentage of tasks completed by the individual participant in relation to two behavior technicians identified as high performers. Following the client shift, the researcher shared the graph and delivered vocal social comparison feedback during an individual meeting with each participant. All 11 cleaning tasks were being performed below the mastery criterion, but performances exceeded the criterion on seven of 11 tasks with the addition of social comparison feedback. Completion of each task increased between 55% and 97%. Overall, the average percentage of tasks completed by all three

participants was 6% in baseline and 83% during intervention. The results suggest social comparison enhanced the performance of participants and that the intervention is a cost-effective strategy for increasing cleaning behaviors.

***Keywords*: social comparison feedback, comparative feedback, peer comparison, normative comparison, social-norms**

Table of Contents

List of Figures

List of Tables

Acknowledgement

I would like to thank my advisor, Dr. Rachael Tilka, for her guidance and support throughout my graduate school career. It was an honor to be her first thesis student and I am proud of the work we accomplished together.

I would also like to acknowledge my thesis committee: Dr. David Wilder and Dr. Darby Proctor. Thank you for challenging me. As a result of your insight and suggestions, I am a stronger writer and researcher.

A special thank you is given to Gina, Estefania, and Missy for their dedication and assistance with data collection, interobserver agreement, and feedback delivery. It was a privilege to work with you.

Lastly, I want to thank my family and friends for their unconditional love, support, and encouragement. I could not have made it this far without you!

Curtis T. Phillabaum

The Effect of Social Comparison Feedback on Task Completion in a Human-Service Setting

Establishing a feedback system can be a challenging feat, especially when the job is demanding and the environment offers competing responsibilities. Often managers will prioritize the supervision of client services at the expense of secondary tasks, such as providing feedback on infection-control practices (Babcock, Sulzer-Azaroff, Sanderson, & Scibak, 1992). Research conducted within human-service organizations has focused primarily on the treatment of individuals with intellectual disabilities and the integrity with which the behavioral intervention plans are delivered. In addition, few studies have examined staff preparation and cleaning tasks at these locations (Gravina, Villacorta, Albert, Clark, Curry, & Wilder, 2018). Yet, these behaviors have a direct impact on both consumers and staff and poor performance can negatively impact health (Carr, Wilder, Majdalany, Mathisen, & Strain, 2013; Herzing & Jacobsson, 2019; Rose & Ludwig, 2009), client outcomes (Carr et al., 2013), and ultimately business results (Gravina et al., 2018). Therefore, these organizations would benefit from cost-effective strategies to enhance staff cleaning behaviors, especially since cleaning tasks are often unpleasant and/ or physically demanding (Anderson, Crowell,

Hantula, & Siroky, 1988; Doll, Livesey, McHaffie, & Ludwig, 2007; Shier, Rae, & Austin, 2003). One strategy which has proven successful in a variety of settings is the use of performance feedback.

The Effectiveness of Feedback

Feedback has commonly been referred to as a popular, inexpensive, and straightforward intervention that informs a person on past performance, notifies where a performer stands in relation to a goal or target, and indicates how to improve future performance (Alvero, Bucklin, & Austin, 2001; Daniels & Bailey, 2014; Durgin, Mahoney, Cox, Weetjens, & Poling, 2014). Supplementary to cost effectiveness and ease of delivery, there are many other characteristics that contribute to the success of performance feedback. These characteristics include the source, frequency, participants, modality, privacy, and content of the feedback (Alvero et al., 2001).

The first feedback characteristic to consider is the source, or who presents the feedback to the performer. According to Lechermeier and Fassnacht (2018), there are only 28 empirical studies comparing feedback sources. Based on these studies, the researchers identified five original sources of feedback: self, computer, coworker, supervisor, and teacher generated feedback. In most organizational settings, the supervisor/ manager will deliver feedback to their direct reports. This practice is consistent with what has been recommended by research. For instance,

Alvero, Bucklin, and Austin (2001) found that feedback is most effective when delivered by supervisors/ managers and researchers. One possible explanation for this finding is the credibility and power of the feedback source. Specifically, feedback sources perceived to have more expertise (experience with the task) and knowledge (observations/ information) are more likely to influence an individual's performance (Andiola, 2014; Giffin, 1967; Steelman & Rutkowski, 2004). For example, researchers studying feedback can provide supervisors with insight into when and how feedback should be delivered and ensure that procedures are implemented with integrity. Also, a supervisor or researcher who observes employees, records data, and delivers feedback is aware of what is going on and will therefore be more likely to communicate accurate information concerning how the performance impacts the organization (Daniels & Bailey, 2014). In addition, a supervisor typically has the authority to deliver consequences (e.g., monetary rewards, professional advancement, etc.), that may be contingent on satisfactory or outstanding performance (Austin, Weatherly, & Gravina, 2005; Crowell, Anderson, Abel, & Sergio, 1988; Durgin et al., 2014; Gravina, Van Wagner, & Austin, 2008; LaFleur & Hyten, 1995; Reetz, Whiting, & Dixon, 2016; Therrien, Wilder, Rodriguez, & Wine, 2005). For all these reasons, recipients are more likely to comply with the feedback.

Frequency of feedback is also an important characteristic to consider. Although weekly feedback is the most common frequency of delivery, daily

feedback has been shown to have the greatest impact (Alvero et al., 2001; Cooper, 2009). For instance, So, Lee, and Oah (2013) identified four critical customer-service behaviors for employees to engage in and compared the effects of weekly and daily feedback on the percentage of completion. The percentages substantially increased for three of the four behaviors following weekly feedback. However, when the same feedback was delivered daily, the researchers observed additional improvements in all behaviors, but especially in the completion of the fourth behavior. So et al. (2013) concluded that the fourth behavior required a larger response effort and discovered that the smaller interval between performance and feedback contributed to increased performance. Likewise, Pampino, MacDonald, Mullin, and Wilder (2003) assessed the effects of weekly and daily feedback on the completion of secondary duties at a retail store. First, the researchers delivered verbal feedback during weekly staff meetings which consisted of the provision of a positive statement, the percentage of group task completion, and suggestions for improvement. Researchers also presented a graph to illustrate the percentage of tasks completed by the group for each day. The same feedback content was then delivered daily during individual meetings with staff. As a result, tasks related to the general maintenance and appearance of the store increased from a mean of 18% completion in baseline, to 66% completion following weekly feedback, to 86% completion with the addition of daily feedback. However, while both studies suggest that daily feedback is more effective than weekly feedback, they failed to

examine how participants would respond to individualized feedback rather than group feedback.

Along with the frequency of delivery, the number of individuals receiving the feedback has been shown to directly contribute to its effectiveness. For instance, Cooper (2009) found that group feedback was six times more effective at reducing workplace injuries than individual feedback and, therefore, suggested that group dynamics and peer pressure contributed to safe performance. Despite the consistent effects found among studies implementing group feedback (Alvero et al., 2001; Balcazar, Hopkins, & Suarez, 1985), many researchers advocate for individual feedback. Group feedback often includes a generalized statement about overall performance whereas individual feedback is typically task-specific (Andiola, 2014) and communicates how to accurately complete the task (Goodman, Wood, & Chen, 2011). In addition, Nadler (1979) delivered group feedback and found that individuals took credit for positive feedback and attributed the corrective feedback to others. Moreover, delivering feedback to a group can lead to social loafing, or the tendency of individuals to reduce their effort and contributions in a team context (Karau & Williams, 1993; Shepperd, 1995). When recognition and rewards are contingent on group outcomes rather than individual contributions, there is less motivation for individuals to perform well and a greater probability of rewarding poor performances (Daniel & Bailey, 2014). Few studies have utilized both individual and group feedback and, of the studies that have, group feedback

was typically delivered publicly in graphic form (Alvero et al., 2001). Therefore, one may argue that the feedback was effective not because it was delivered to a group, but because of the modality of feedback delivery.

Of the eight different modalities, verbal or written feedback combined with a graphic display of performance has been shown to be the most effective. Though written feedback is the most commonly used type, it should be noted that its effectiveness is enhanced with the addition of graphic feedback (Alvero et al., 2001). For instance, to increase cleaning behaviors at a ski shop, Doll, Livesey, McHaffie, and Ludwig (2007) implemented a packaged intervention that included task clarification, graphic feedback, and written feedback. Graphic feedback was posted weekly to illustrate the group percentage of task completion from the previous week. During the next phase, the researchers posted written feedback daily to identify the neglected tasks. As a result, the percentage of completion improved to near-perfect levels across all behaviors. Similarly, Austin, Weatherly, and Gravina (2005) found that the combination of verbal and graphic feedback had the greatest impact on the completion of closing tasks at a restaurant. During bi-weekly pre-shift meetings with staff, managers delivered verbal feedback that identified the items on a checklist that were completed well along with tasks that needed improvement. At the end of the week, the researchers posted a graph of group performance and, as a result, task completion improved by an average of 38%. The aforementioned examples are just two of many studies illustrating the

success of verbal or written feedback in combination with graphic feedback (Crowell et al., 1988; Gravina et al., 2008; Pampino et al., 2003; Rohn, Austin, & Lutrey, 2003; Slowiak, 2014, So et al., 2013; Wilk & Redmon, 1997). Yet, if both types of feedback were delivered more frequently and the graphs displayed individual performances, then perhaps researchers may have observed more rapid changes in behavior.

Experts within organizational behavior management recommend reserving public posting for group feedback since the performance data is openly displayed for others to see (Daniels & Bailey, 2014) and displaying individual data can have adverse effects on motivation and morale (Andrasik, Heimberg, & McNamara, 1981). However, group performance data often disguises whether or not a particular individual improves (Nordstrom, Lorenzi, & Hall, 1991). For that reason, many studies have implemented public posting of individual performances (Crowell et al., 1988; Emmert, 1978; Newby & Robinson, 1983; Rose & Ludwig, 2009; Van Houten, Hill, & Parsons, 1975; Welsch, Ludwig, Radiker, & Krapfl, 1973; Wikoff, Anderson, & Crowell, 1982). For instance, Newby and Robinson (1983) found that group feedback was not very effective given the independent nature of the tasks, the low impact on the performance of others, and the lack of information to enhance individual performance. On the other hand, the posted individual feedback significantly improved the punctuality, cashiering, and check-out performance of 15 employees in a retail setting. In addition, researchers can minimize adverse

effects of publicly posting individual performance by assigning code names. Providing some identification and posting individual contributions can invoke social comparison and reduce social loafing (Chen, Zhang, & Latimer, 2014). For example, bank tellers were given a code name so their mean transaction quality points could be publicly posted (Crowell et al., 1988). The daily graphic feedback was then combined with verbal feedback consisting of a brief meeting with each teller. When delivering the feedback, the manager acknowledged the individual's graph, compared the current score to the previous one, and explained how completion of specific tasks led to certain scores. Consequently, teller transaction scores exceeded the levels of acceptable performance determined by management.

Lastly, feedback can be identified by the type of information provided to the employee. Based on the literature (e.g., Alvero et al., 2001), there are two popular and effective contents of feedback: 1. Comparison of the individual's performance with his or her previous performance; and 2. Comparison of the individual's performance with the individual's standard performance. For instance, when Wilk and Redmon (1997) compared an individual's performance with the individual's standard performance, productivity and satisfaction increased among university admissions staff. Additionally, researchers who have implemented feedback that compared an individual's performance with a past performance found an increase in desired behaviors (Anderson et al., 1988; Crowell et al., 1988). The studies previously discussed illustrate that feedback is an effective intervention and the

effects of feedback can be enhanced through incorporating the proper components during delivery. Perhaps the marked effects experienced through the utilization of feedback alone is one reason it continues to be the most popular procedure (Alvero et al., 2001; Balcazar et al., 1985). However, within behavior analysis, one content area that has not been heavily explored is comparing an individual's performance to a peer's performance.

Social Comparison Outside of Organizational Behavior Management

The concept of comparing one's opinions, abilities, or characteristics to someone else's is known as social comparison (Festinger, 1954; Garcia, Tor, & Schiff, 2013; Myers & Twenge, 2019). While not much has been published on social comparison in the behavior analytic literature, there have been many publications outside the field that merit mention. With that in mind, it should be emphasized that when considering the studies that follow, the interpretation of the findings is based on those provided by the original authors. The intention of presenting these studies is that they have the potential to incite ideas for exploring this topic, specifically in relation to new questions that can be explored. Also, it is important to note that future research is needed that will interpret the findings from a behavioral viewpoint.

Some common measurements of social comparison include grades, test scores, teacher comments, physical appearance, athleticism, clothing, and hobbies. According to social comparison theory, the comparison allows individuals to form an accurate appraisal of their performance (Dijkstra, Kuyper, van der Werf, Buunk, & van der Zee, 2008; Festinger, 1954). There are several individual and situational characteristics that increase social comparison: the similarity to the target (Festinger, 1954), performance-related attributes (Goethals & Darley, 1977), relevance of the comparison dimension (Hoffman, Festinger, & Lawrence, 1954; Huguet, Dumas, Monteil, & Genestoux, 2001; Tesser & Smith, 1980), psychological closeness (Meisel & Blumberg, 1990; Tesser & Smith, 1980), and proximity to a standard (Garcia & Tor, 2007).

When given a choice, people typically make comparisons to others within their reference group because they share similar abilities. When a comparison performance is too far above or below, it is difficult for an individual to estimate their own performance (Dijkstra et al., 2008; Festinger, 1954). Therefore, choosing a target who is similar can minimize the discrepancies between performances (Garcia et al., 2013; Gilbert, Giesler, & Morris, 1995; Whittemore, 1925). For instance, when a confederate was given a significantly higher score on an intelligence test, the other individuals in the group ceased to compare their scores with the top performer and instead, compared only with each other (Hoffman, Festinger, & Lawrence, 1954). Researchers suggest the redefinition of the

comparison group minimized the disparity and the individuals could make a subjectively better assessment of their own performance. Thus, future studies aiming to improve lower performers should include comparison individuals who are performing slightly better.

As mentioned, sometimes a performer may select a comparison target based on sharing similar attributes despite the fact that these attributes may be inconsequential to the context (Garcia et al., 2013; Goethals & Darley, 1977). For instance, Miller (1982) found undergraduates selected a comparison target similar in attractiveness to compare test scores. These measures are highly subjective and, therefore, it is unclear how physical appearance directly relates to academic performance. As an alternative, researchers suggest that the attributes with the greatest influence on performance and social comparison include age (Suls, Gastorf, & Lawhon, 1978), gender (Blanton, Buunk, Gibbons, & Kuyper, 1999; Golden & Cherry, 1982), race (Aboud, 1976; Meisel & Blumberg, 1990), and socioeconomic status (Régner & Monteil, 2007). According to Dijkstra, Kuyper, van der Werf, Buunk, and van der Zee (2008), students are more likely to select comparison targets because they can identify with others sharing these characteristics. Within applied settings, populations are diverse and the comparison target is often selected by management. Therefore, researchers can control for these attributes by providing participants with anonymity and directly observing their behaviors.

The relevance of the comparison dimension can also impact the effect of social comparison. Being outperformed on a dimension can be threatening, especially when that dimension is highly important to the individual's self-definition. On the other hand, if an individual is outperformed on something of low relevance, then the social comparison has less impact on performance (Dijkstra et al., 2008; Festinger, 1954; Garcia et al., 2013). For example, Tesser and Smith (1980) had participants identify a word based on clues of varying difficulty presented by other participants. Individuals in the high relevance condition were told the task was an exercise on verbal skills and success had been linked to college exams (e.g., SAT, GRE, and MCAT) and hiring opportunities at large companies. Those in the low relevance condition were told that the task was a game in which researchers were trying to identify the skills needed to win. Participants in the high relevance condition perceived a threat and, therefore, gave harder clues to hinder the performance of others. On the other hand, individuals in the low relevance condition were not threatened and social comparison had little impact on performance. Also, Huguet, Dumas, Monteil, and Genestoux (2001) found an increase in dimension relevance was correlated with an increase in the comparison target. Students were asked to rate the importance of seven academic subjects to their self-identity and to select two students in each class for the purposes of comparing grades. They chose a different target depending on the subject and, when the comparison was perceived as a threat, students frequently engaged in

upward social comparison (i.e., comparing their performance to someone who performed slightly better). While studies show that relevance can influence one's likelihood to engage in social comparison, Dijkstra et al. (2008) warn that forcing individuals to compare themselves on irrelevant dimensions might not produce an effect. Furthermore, studies have relied on self-report and subjective measures (e.g., self-identity). Future research is needed that implements a more behavioral approach and provides a behavioral interpretation of the findings.

Other studies have found that "psychological closeness," or the degree to which individuals relate to one another, contributes to social comparison (Meisel & Blumberg, 1990; Tesser & Smith, 1980). People often compare themselves to friends, siblings, or others with whom they have positive feelings towards (Garcia et al., 2013). For instance, elementary students preferred to compare academic achievements with popular students whereas ninth graders tended to compare their performance to best friends (Meisel & Blumberg, 1990). Likewise, when seventh, eighth, and ninth graders could choose two students to compare exam grades, they chose close friends (Dumas, Huguet, Monteil, Rastoul, & Nezlek, 2005). However, when a friend outperformed the student on a task that was relevant to one's self-definition, participants were more likely to give harder clues to the friend than to a stranger (Tesser & Smith, 1980). Thus, social comparison may have had a negative impact on self-esteem and, therefore, participants were less likely to give help. Additional studies suggest that psychological closeness and personal history

intensify social comparison and competitiveness (Charness, Masclet, & Villeval, 2014; Garcia et al., 2013; Harbring & Irlenbusch, 2008). For instance, participants artificially inflated their performance when the ranking feedback was based on the performance of group members from the same college (Charness et al., 2014). While group identity seemed to decrease acts of sabotage between these participants, researchers observed an increase in rivalry. This is perhaps one reason why behavior analysts and practitioners have expressed concern about competition and the potential for individuals to engage in unethical behaviors (Daniels & Bailey, 2014; Johnson & Dickinson, 2010). In addition, terms such as psychological closeness, self-definition, and self-esteem would benefit from a more behavioral interpretation and, thus, future research is needed to develop this line of research within the area of behavior analysis.

The proximity to a standard can also intensify social comparison and competitiveness (Garcia et al., 2013; Vandegrift & Holaday, 2012). In a series of experiments, participants were given hypothetical scenarios in which they had to choose between getting paid equally or accepting a larger profit which favored the other individual. One study by Garcia and Tor (2007) assigned participants to a top rank (i.e., number one versus number two), an intermediate rank (i.e., number 101 versus number 102), or a no scale comparison condition (i.e., the top or intermediate standing was not affected by choice). Those in close proximity to the top standard experienced a threat to their rank and, therefore, chose to get paid

equally whereas those with an intermediate rank were far more likely to maximize their joint gains. Likewise, when the rank was not affected by the participant's choice, the participants behaved more cooperatively. In the second scenario, participants were CEOs of a Fortune 500 company and assigned to similar conditions. Those in the scale comparison group were given a top rank (i.e., number two versus number three), an intermediate rank (i.e., number 231 versus number 232), or bottom rank (i.e., number 500 versus number 501). Results showed 35% of top-ranked participants, 70% of intermediately ranked participants, and 13% of bottom-ranked participants chose to maximize profits. Garcia and Tor (2007) illustrate that proximity to any meaningful standard triggers upward comparison and competition. In a follow-up study, the researchers showed that participants who behaved competitively were driven by the threat of upward comparison on a scale and not by the comparison on the task alone. Consequently, they caution the use of forced ranking in organizations. When individuals focus on where they stand in relation to others on a scale, the contributions to the organization and individual performance may suffer (Garcia & Tor, 2007).

These findings also align with what has been argued by behavior analysts and practitioners (e.g., Daniels & Bailey, 2014; Johnson & Dickinson, 2010). Although rankings and employee-of-the-month incentive programs are commonly implemented to motivate and reward good performance, they are not without their flaws. These employee-comparison methods rank staff from high to low on a

performance dimension that is based on a standard that is often arbitrarily decided by management (Daniels, 2009). Furthermore, since organizations primarily focus on business results, employees may engage in undesirable, unethical, or illegal activities to achieve the top spot (Johnson & Dickinson, 2010). Accordingly, it has been suggested that these programs promote lying, cheating, stealing, and backstabbing (Daniels & Bailey, 2014) and sabotaging a peer's chances at winning (Harbring & Irlenbusch, 2008). In addition, the winner-take-all design can punish employees who performed slightly under the top performer and potentially extinguish desirable behaviors (Johnson & Dickinson, 2010). For these reasons, Daniels and Bailey (2014) advocate against ranking systems and argue that publicly posting data leads to unhealthy and destructive competition. To overcome this problem, future research could explore the effect of social comparison that is not based on ranking and provides everyone with the opportunity to earn a good score independently of other workers' performance.

Three Functions of Social Comparison

While there are a variety of variables that can contribute to social comparison, there are three main functions: self-evaluation, self-enhancement, and self-improvement (Dijkstra et al., 2008). The first aligns with social comparison theory (Festinger, 1954) which states that people compare themselves to similar individuals to form an accurate appraisal of their own performance. As previously

discussed, there are several studies to support that people engage in social comparison for self-evaluation purposes (Aboud, 1976; Blanton et al., 1999; Dumas et al., 2005; Golden & Cherry, 1982; Hoffman et al., 1954; Huguet et al., 2001; Meisel & Blumberg, 1990; Régner & Monteil, 2007; Suls et al., 1978). The second function, self-enhancement, suggests that social comparison can protect or enhance self-esteem (Dijkstra et al., 2008). Specifically, individuals under stress often engage in downward comparison, or the comparison to others who are worse off, in order to cope with difficult situations (Collins, 1996; Garcia et al., 2013; Wade & Tavris, 2017). Lastly, people may engage in social comparison in an act of self-improvement. People often compare upward to a target who is slightly better in order to improve their skills and abilities (Garcia et al., 2013; Huguet et al., 2001). Upon learning that someone scores better on a task, the tendency is to aspire to that socially favorable level (Festinger, 1954; Moon, Lee, Lee, & Oah, 2017; Myers & Twenge, 2019), to identify with and imitate the comparison target (Huguet et al., 2001), or to set higher standards and personal goals (Huguet et al., 2001; Dijkstra et al., 2008).

Social Comparison in Schools

To date, an extensive number of studies have examined social comparison within educational settings, particularly with elementary and secondary school students. These classrooms consist of a stable group that attends classes together

for years and shares similar characteristics and abilities. Also, students become accustomed to evaluative feedback from teachers, rewards for excellent performances, pressures from parents, and social comparison with peers (Dijkstra et al., 2008). According to Feldman and Ruble (1977), young children use social comparison as a means of competency testing, or learning about norms and standards of behavior. Specifically, young children compare themselves to get answers and to determine if they are doing a task correctly. As children age, they become more confident in their abilities and use social comparison to evaluate their own performance (Ruble, Boggiano, Feldman, & Loebl, 1980), fine tune their skills (Dijkstra et al., 2008), and outperform peers (Charness et al., 2014; Seidner, Stipek, and Feshbach, 1988).

In addition to identifying variables, moderators, and motives for social comparison, researchers have identified contexts in which students benefit from comparing their performance to others. For instance, Monteil and Michinov (2000) discovered that low performers chose upward comparison in private and downward comparison in public following a poor performance. Others have found that grouping students by ability can encourage social comparison. High performers often serve as a role model, a source of information, and motivation (Dijkstra et al., 2008) and, therefore, low performers are likely to compare upward (Reuman, 1989). Alternatively, when high performers are grouped with lower performers, their performance may deteriorate; therefore, social comparison can have positive

or negative effects depending on the standard (Dijkstra et al., 2008). These studies suggest that the context and level of the performer may contribute to the effectiveness of social comparison.

Social-norm Campaigns

Over the past decades, social-norm campaigns have gained attention and popularity. Rather than using moral and fear-inducing messages to reduce problematic behaviors, these campaigns communicate peer norms to serve as a descriptive standard of prosocial behaviors (Schultz, Nolan, Cialdini, Goldstein, & Griskevicius, 2007). The standard creates a "norm to perform" and, as a result, underperformers are more likely to change their behaviors. In addition to activating personal and social motives, social-norm campaigns create awareness of the problem and link actions to specific outcomes (Herzing & Jacobsson, 2019). This strategy has been used to address socially significant behaviors such as energy consumption (Ayres, Raseman, & Shih, 2012; Egan, 1999; Haakana, Sillanpää, & Talsi, 1997; Ueno, Inada, Saeki, & Tsuji, 2005; Wilhite, Hoivik, & Olsen, 1999), healthy food choices (Mollen, Rimal, Ruiter, & Kok, 2013; Robinson, Fleming, & Higgs, 2014; Thorndike, Riis, & Levy, 2016), and good hygiene (Herzing & Jacobsson, 2019; Schmiege, Klein, & Bryan, 2010).

The use of social comparison to conserve energy. Social comparison to reduce energy consumption has had mixed results. Based on a review by Fischer

(2008), twelve studies have implemented feedback that compared consumption to other households in an effort to stimulate competition and ambition. While consumers in Finland and Japan favored the normative comparison, individuals in Sweden and the UK preferred to compare their current consumption to a previous one. Regardless, Haakana, Sillanpää, and Talsi (1997) and Ueno, Inada, Saeki, and Tsuji (2005) could not demonstrate an effect because the social comparison feedback encouraged households with low consumption rates to increase usage, thus cancelling out any reductions observed in high consumers.

In an effort to combat this boomerang effect, Schultz, Nolan, Cialdini, Goldstein, and Griskevicius (2007) delivered social comparison feedback and injunctive messages to households on their energy consumption. Those in the social-norm only condition received a handwritten letter comparing their energy usage from the previous week with the average energy usage per household within the neighborhood. Additionally, strategies were provided to reduce energy consumption. Households in the social-norm plus injunctive message condition received the same normative information with the addition of a happy face for consumption below the average or a sad face for consumption above the average. As a result, high consumption households that received the social-norm feedback, with or without the sad face, significantly reduced their energy usage. Households with low consumption that received just the social-norm feedback increased in consumption, while those who received the feedback with a happy face maintained

low levels. The researchers showed that the addition of approval can prevent households that abstain from high consumption from trying to meet the standard. Overall, however, there is a lack of studies on this topic and Fischer (2008) cautions others on applying the results.

The use of social comparison to promote healthy choices. In an effort to increase healthy food choices in a hospital cafeteria, Thorndike, Riis, and Levy (2016) compared purchases made by employees who received no feedback, peer-comparison feedback, and peer-comparison feedback plus a monetary incentive. Participants assigned to feedback received a monthly letter displaying color-coded bar graphs (i.e., green – healthy, yellow – less healthy, and red – unhealthy) for their personal purchases, the average of all employees, and the average of the healthiest employees. Also, some participants received a 10-dollar reward for meeting goals (40%, 60%, or 80% of green-labeled purchases) based on their purchases during the previous month. The researchers observed a 2.2% and 1.8% rate of change in healthy purchases for employees who received feedback plus the incentive and feedback alone, respectively. These results indicate that the incentive did not have a significant impact over the information on social norms. When Thorndike et al. (2016) analyzed subgroups, they discovered that employees who made the least healthy choices during baseline did not significantly change their purchases with the addition of feedback or feedback plus incentives. Based on population demographics, these individuals were prone to weight gain and obesity.

While the authors suggest increasing the incentive, the immediacy of payment, and providing social norms that target specific groups, a better approach might be to examine the relevance of the comparison dimension.

The use of social comparison to promote good hygiene. Researchers have examined the effect of social comparison feedback on various hygienic practices. For instance, Schmiege, Klein, and Bryan (2010) examined the effect of peer comparative information on the attitudes and beliefs toward teeth flossing and its future behavior. Undergraduate participants completed a pre- and posttest survey on perceived risk, worry about outcomes, and behavioral intentions. Behavior was self-reported during the pre-test and at a three-month follow-up. All participants received either lateral comparison (i.e., they flossed the same as peers) or upward comparison (i.e., they flossed five times less per week). In addition, half received objective information in the form of expert recommendations (i.e., by flossing fourteen times per week, you are five times less likely to get gum disease). Researchers found that social comparison had a significant effect on later behavior. Specifically, upward social comparison was correlated with a greater perceived risk, tendency to worry, and higher intentions of flossing. Despite participants rating the expert recommendation to be high in value and believability, there were no main effects. The researchers concluded that the recommendation was exceedingly high in relation to participant self-reports and, therefore, the objective information was not as meaningful.

Consequently, Schmiege et al. (2010) modified the study to include upward comparison (i.e., six times per week) and downward comparison (i.e., once per week), a condition that did not include social comparison, and they reduced the expert recommendation to seven times per week. Results showed that the peer comparison information (i.e., none, once, or six times per week) did not influence the perceived risk or worry, but upward comparison was linked to favorable attitudes toward flossing and greater intentions. This study also suggests that social comparison could perhaps be more effective at promoting health behaviors and wellness than objective information. However, the interpretations are limited given that results are based on self-report data which were not verified through direct, observational measures. Despite this limitation, the study is still integral in that it promotes the formulation of new hypotheses relevant to the topic that can be assessed using more objective methods.

Additionally, Herzing and Jacobsson (2019) used more direct measures (i.e., permanent products) to assess the cleanliness of preschools in three Swedish municipalities. Annually, inspectors swabbed the bathroom handles for organic material and high levels of adenosine triphosphate (ATP) which are indicative of poor cleaning behaviors. All preschools in one municipality received a written feedback letter consisting of (a) their individual ATP level, (b) the average and medium ATP levels in the municipality, (c) the percentage of preschools that had a lower level, (d) information on acceptable levels, and (e) the importance of

preventing the spread of infections. This feedback had a strong and statistically significant effect on poorly performing schools; those with the highest ATP levels, showed the most improvement and reduced ATP levels by 42%. On the other hand, the cleanest schools got significantly worse. Feedback which indicated that these schools were in the top 33% led to undesirable behaviors in these schools. The researchers concluded that social comparison can be an effective tool, but that further research is necessary to address the boomerang effect. Furthermore, while this research allows one to conclude that social comparison feedback may be an effective intervention for promoting cleaning behaviors (in lower performers), there are still limitations. These conclusions were based on permanent product measures. Although permanent product measures may be more valid than self-report measures, one could argue that collecting direct observational measures at the time the behavior occurs would further strengthen the argument.

Social Comparison in Organizational Behavior Management

As mentioned, there are few studies in the behavior analytic literature that incorporate social comparison feedback (Alvero et al., 2001; Balcazar et al., 1985). These studies have occurred in classrooms (Van Houten et al., 1975), industrial and manufacturing plants (Emmert, 1978; Wikoff et al., 1982), a pharmacy (Newby & Robinson, 1983), at a community swimming pool (Rose & Ludwig, 2009), and in human-service organizations (Panyan, Boozer, & Morris, 1970; Welsch et al., 1973;

Gil & Carter, 2016). Researchers have implemented social comparison feedback to increase punctuality, cashiering, and check-out performance (Newby & Robinson, 1983) and to increase the completion of closing tasks (Rose & Ludwig, 2009). Within the human-service setting, this type of feedback has been used to increase the percentage of sessions delivered by staff (Panyan et al., 1970), completion of daily projects (Welsch et al., 1973), and submission of data sheets (Gil & Carter, 2016).

Researchers have delivered feedback comparing groups (Gil & Carter, 2016), individuals (Moon et al., 2017; Rose & Ludwig, 2009; Van Houten et al., 1975; Wikoff et al., 1982), or a combination of both social comparison strategies (Emmert, 1978; Newby & Robinson, 1983; Panyan et al., 1970; Welsch et al., 1973). Feedback was publicly posted in the form of charts (Newby & Robinson, 1983; Van Houten et al., 1975; Welsch et al., 1973), graphs (Gil & Carter, 2016; Rose & Ludwig, 2009; Wikoff et al., 1982), or a rank (Moon et al., 2017; Panyan et al., 1970). In addition to public posting, one study (Panyan et al., 1970) implemented verbal feedback which involved the unit psychologist reviewing permanent products with each attendant. Lastly, feedback was delivered daily with the exception of Panyan, Boozer, and Morris (1970) which occurred weekly and Moon, Lee, Lee, and Oah (2017) which occurred immediately by computer. Similar to other disciplines, the effectiveness of social comparison feedback in behavior analysis has been mixed (Alvero et al., 2001; Balcazar et al., 1985; Nordstrom et

al., 1991). While these studies do incorporate social comparison within an intervention package, its effectiveness has rarely been examined in isolation.

Successful applications of social comparison feedback. The effects of social comparison are enhanced by incorporating the most effective characteristics of feedback into the delivery. For instance, Gil and Carter (2016) posted bar graphs in thirteen residential homes so staff could easily compare their current group percentage to the percentages of other homes in their group. Homes were assigned to groups based on baseline percentages of submitting data sheets; those in Group 1 had the highest levels of compliance. Following graphic feedback, compliance increased from 43% to 48% in Group 1, 26% to 59% in Group 2, and 28% to 62% in Group 3. The reason the social comparison feedback was likely effective was due to the fact that these researchers used bar graphs rather than a ranking system. Consequently, every home had the possibility of getting a high score independently of the other homes. The bar graphs were easy to interpret, facilitated comparison between multiple homes, and provided a graphic display of social-norms. While this study provides a nice illustration of the effect of social comparison, the emphasis was not placed as heavily on that aspect of the intervention. With that in mind, this important contribution to the literature should be noted. Additionally, future research might attempt to replicate these results through similar procedures but place a heavier emphasis on the direct effects of social comparison, especially given that it is a relatively cost-effective intervention.

Within the laboratory setting, Moon et al. (2017) found that performance level can impact the effectiveness of social comparison feedback. Researchers divided undergraduate students into high and low performers, gave them a simulated banking task, and the students received either objective or social comparison feedback. The objective feedback consisted of the number of tasks completed correctly and the social comparison feedback provided a performance rank. High performers performed better in the social comparison feedback condition while the low performers performed better in the objective feedback condition. The results suggest that the ranking system was motivating for high performers, but the winner-take-all design might have extinguished task completion in lower performers. In addition, the objective feedback provided relevant information so low performers could improve upon their performance. These findings suggest that social comparison feedback can have different functions depending on the level of the performer. Further research is necessary to determine if social comparison feedback has similar results in applied settings.

In summation, an extensive number of studies have been completed on social comparison outside of behavior analysis. While these studies are beneficial in that they may help promote the development of hypotheses for future research, the current research tends to accredit the phenomenon to a mental or inner dimension (e.g., reference group, self-definition, psychological closeness, self-esteem, and group-identity). Self-report data is also typically used to assess

performance which may not always be valid. The current body of literature would benefit from a behavioral interpretation and/or approach to measurement. In addition, the effects of social comparison have been mixed and seem to depend on the level of the performer. Whereas low performers tend to compare upwards in order to improve their performance, researchers have observed a boomerang effect in which individuals performing better than average frequently reduce performance to meet the social-norm.

Furthermore, while some social comparison research does exist within the behavior analytic literature, researchers have frequently examined social comparison as one component within a package. Thus, the effects of social comparison in isolation are not as clear. Perhaps one reason for the lack of research is due to the fact that behavior analysts and practitioners have advised against peer comparison, citing that publicly posting individual data promotes unhealthy competition and has adverse effects on lower performers. However, as mentioned, this may be more relevant if a ranking system is applied given that only one performer has the ability to be at the top and the performance of each individual directly impacts the standing of others within the group. Thus, a bar graph may be optimal as it will allow all performers to achieve high levels independently of the performance of others. Therefore, the purpose of the present study was to assess the effects of social comparison feedback on the percentage of end-of-shift cleaning tasks completed by three behavior technicians identified as exhibiting lower

performance. The tasks were directly observed by researchers as they were being completed and the social comparison feedback consisted of bar graphs depicting the percentage of tasks completed by the individual participant in relation to two high performers. We predicted that social comparison feedback would increase the task completion of low performers.

Method

Settings and Participants

This study took place in two large classrooms in the early intervention (EI) wing of an autism treatment center located on a university campus in the southeastern United States. In addition to their primary responsibilities, therapists working in these rooms were responsible for general cleaning tasks to maintain the facility and to prepare the client for pickup from caretakers. With the help of management, researchers identified 11 tasks all therapists were expected to engage in at the end of each client shift (see Appendix A).

Participants included three full-time therapists who provide early intensive behavioral intervention services to children diagnosed with autism. Employees were given a written consent form to review and an opportunity to ask questions (see Appendix B). The researcher emphasized that participation was voluntary and if they agreed to participate, their identities would be kept anonymous, data would be kept confidential, and participation or non-participation would not affect their standing with the company. Participants were prompted to select their own code name to be used for the duration of the study.

Low performers. Three individuals who performed the cleaning duties at the lowest levels were selected as participants in this study and their data were included for research purposes.

High performers. Two high performing individuals were included in this study to motivate low performers. While it may have been possible to use pseudo performers, the director of the clinic required the use of real performers. Their data were collected for the purpose of social comparison only and they received an intervention to ensure their performance reached and remained at high levels. However, their data were not included for research purposes.

Data Collection and Interobserver Agreement

Researchers directly observed participants while they performed various end-of-shift cleaning tasks. The observation period started 10 minutes prior to the end of the client session and ended 15 minutes after completion of the shift. Researchers observed from an observation room window and recorded items as they were completed (see Appendix C). Throughout the study, participants were unaware of when they were being observed.

During 33% of sessions, a Research Assistant along with the Primary Investigator observed and independently recorded if tasks were completed accurately. The data were compared against the data obtained by the Primary Investigator. To calculate interobserver agreement (IOA), the number of point-to-point agreements was divided by the total number of items on the checklist and then converted to a percentage. If IOA fell below 100% during an observation session, researchers identified where the mistake was made and discussed how to

correct the error. The overall IOA was 98.4% (range, 91% to 100%). Prior to data collection, the Research Assistants received extensive training which consisted of task analyses, modeling, role play, and feedback.

Dependent Variables

The dependent variable assessed in this study was the percentage of tasks completed by each individual during one session (i.e., one end-of-shift cleaning). All 11 tasks could be completed in approximately 5-10 minutes. For tasks to be considered complete, they had to meet the requirements detailed on the memos posted on the main bulletin board in the EI wing and doors to each classroom (see Appendix D).

Measurement

Percentage of tasks completed by each participant. To calculate the percentage of tasks completed by each participant, the researcher divided the number of tasks completed by the total number of tasks on the checklist and multiplied this amount by 100%.

Independent Variable

The independent variable in the study was graphic and verbal social comparison feedback delivered to participants following each client session. During an individual meeting, the Primary Investigator shared a bar graph

depicting the percentage of tasks completed by the participant and the percentage of tasks completed by each high performer for each individual's *previous session*. The verbal feedback consisted of: (a) the current percentage of tasks completed correctly by the participant; (b) a comparison to the percentage of tasks performed correctly by the top high performer; and (c) the value of completing the tasks (see Appendix E). While one might argue that delivering a value statement could influence performance, the researcher felt it necessary since the target behaviors were secondary job responsibilities occurring at low rates. Also, if participants did not see the relevance of the tasks, then the social comparison feedback was unlikely to increase performance.

As mentioned, the social comparison feedback had to be based on the performance of high performers. Two performers were selected who exhibited the highest performance in baseline so that their performance could be shared with low performers during each feedback meeting. Although the two individuals outperformed the low performers, they were also performing at relatively low levels. Therefore, feedback techniques that are effective (described below) were applied to increase the performance and to ensure performance maintained at high levels. Their data were continuously collected and updated according to their own progression in performance. Over the course of the study, the high performers completed an average of 92% (range, 73% to 100%) and 95% (range, 73% to 100%).

Research Design

A multiple baseline design across participants was used to assess the effects of social comparison feedback on performance. As mentioned, only the data from the low performers were included in the analysis. The intervention was implemented in a staggered fashion based on visual inspection of trends in the data. This design was selected because it was believed that once the low performers learned the standard that others were performing, it would be difficult to reverse the effects of the intervention and participants would continue to maintain high performance after the independent variable was removed.

Procedures

Pre-baseline task clarification and assessment. One month prior to data collection, the researcher distributed a memo that listed the tasks to be performed as well as definitions of each task. The memo was posted in various places around the facility where it was visible to all staff. The floor supervisor periodically reminded staff during team meetings to complete the cleaning tasks. The researcher assessed task completion from an observation room window. Three full-time employees exhibiting the lowest levels of performance were recruited as participants. In addition, two full-time employees exhibiting the highest performance were selected as comparison targets. The specific selection criteria were not specified to the

participants. In other words, they were not told they were selected due to low or high-performance levels.

Baseline. The researcher observed the individuals identified as exhibiting low or high-performance levels at the end of each client shift. No feedback was provided to participants during this phase. As mentioned, while the performance of high performers was not directly of interest, it was critical to select an intervention that has proven to be consistently effective at maintaining performance. Therefore, the researcher delivered verbal evaluative feedback and praise to each high performer (see Appendix F). Following each client shift, the researcher met with the high performer privately in the observation room to review the performance. The feedback consisted of: (a) a statement about the current performance in relation to the previous one; (b) a praise statement contingent on performance; (c) positive feedback on all the tasks performed well; (d) corrective feedback on an incomplete task and how to improve for next time; (e) the percentage of tasks completed following the present shift; and (f) the value of completing the tasks. The researcher did not compare the performance of high performers with that of the low performers since it was uncertain the impact this sort of comparison might have on high performers. On average, the evaluative feedback meetings for high performers lasted 70 seconds (range, 44 seconds to 120 seconds).

Social comparison feedback. The researcher observed a participant and then delivered graphic and verbal social comparison feedback in private. Each

participant began the intervention with a social comparison that was based on the lowest scores achieved by the high performers. By the time participants entered this phase of the study, high performers had been receiving feedback and achieving near perfect scores. Therefore, to minimize the discrepancy, it was important to display the task completion of high performers following their first feedback meeting (see Appendix G). From that point on, the graphs reflected percentages achieved during the previous session, as described above. The average duration of a social comparison feedback meeting was 43 seconds (range, 32 seconds to 71 seconds).

Social Validity

Participants completed a 5-question social validity questionnaire (see Appendix H) following the conclusion of the study. Using a 5-point Likert scale, participants circled the number that best indicated their response. Questions were related to the participants' (a) task completion prior to receiving feedback, (b) task completion after receiving feedback, (c) perceptions of the social comparison graphs, (d) recommendation to continue the program with other staff, and (e) enjoyment of receiving the feedback. The researcher provided extra space below each question so participants could elaborate on their responses. Results were analyzed and shared with the floor supervisor.

Results

During baseline, the average percentage of end-of-shift cleaning tasks completed by all three participants was 6%. After the implementation of social comparison feedback, the average increased to 83%. Table 1 shows the average percentage of cleaning tasks performed by each participant during the baseline and intervention conditions.

Table 1

Overall Percentage of Cleaning Tasks Performed by Participants

Participant	Last Data Point in Baseline	First Data Point in Intervention	Average in Baseline	Average in Intervention
Mulan	0%	27%	4%	79%
Jafar	9%	82%	7%	93%
Rasputin	0%	73%	7%	76%

Figure 1 depicts the percentage of tasks completed per session by each participant during the baseline and intervention conditions. The intervention was first delivered to Mulan. During baseline, this participant completed an average of 4% of the cleaning tasks. The data were relatively stable between 0% and 9% completion. When the social comparison feedback was implemented, there was an immediate increase in level followed by a steady, increasing trend in task completion. By the fifth session, Mulan completed more than 80% of the tasks. The

performance maintained at or above this level for the remainder of the study, but there was some variability. On three occasions, Mulan completed 100% of the tasks. Social comparison feedback increased Mulan's performance to an average of 79%.

Next, the intervention was delivered to Jafar. During baseline, this participant completed an average of 7% of the end-of-shift cleaning tasks. The data were less stable and the percentage of task completion was between 0% and 18%. When the feedback was implemented, there was a dramatic increase in level with the participant completing more than 80% of the tasks. There was a slight dip in performance, but by the third session the participant was completing 100% of the tasks. The data remained stable at near-perfect levels for the remaining six sessions. Overall, the social comparison feedback increased Jafar's performance to an average of 93%.

Rasputin was the last participant to receive the social comparison feedback. In baseline, this participant completed an average of 7% of the end-of-shift cleaning tasks. Rasputin showed the most variability and completed between 0% and 27% of the tasks during this condition. There was a dramatic increase in level with the addition of feedback; however, the percentage decreased to 55% the following session. Task completion recovered to 82% and continued on an increasing trend for the remainder of the study. The participant's last data point was

at 91%. Social comparison feedback increased Rasputin's performance to an average of 76%.

The Effect of Social Comparison Feedback on Specific Tasks

Table 2 shows the overall percentage of completion per task during baseline and intervention conditions. During baseline, all 11 end-of-shift cleaning tasks were being performed below the expectations of management (80% mastery criterion). Following social comparison feedback, overall performances exceeded the mastery criterion on seven of 11 tasks. The percentage of completion increased the most for Task 2 (97%), Task 7 (92%), and Task 8 (96%). Although Tasks 4, 5, 9, and 10 fell short of the mastery criterion, completion of these tasks increased by more than 55%.

Table 2

Overall Percentage of Completion per Task During Baseline and Intervention

Condition	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7	Task 8	Task 9	Task 10	Task 11
Baseline	11%	0%	4%	0%	0%	28%	0%	0%	22%	0%	0%
Intervention	92%	97%	85%	71%	63%	94%	92%	96%	77%	58%	86%

Figure 2 illustrates the completed tasks during baseline and intervention for each participant. During baseline, only four tasks were performed: Task 1 – clearing off the desk; Task 3 – clearing off the client supply tower; Task 6 – putting

toys away; and Task 9 – stacking/ pushing-in chairs. Jafar was the only participant to engage in all four tasks. Tasks 6 and 9 were the two tasks performed by all participants in baseline. With the addition of social comparison feedback, the percentage of completion increased for all tasks across participants. All three participants met the mastery criterion on Tasks 2, 6, 8, and 11.

Table 3 shows the completion of each task per participant. Tasks highlighted in green were completed during 100% of intervention sessions. Those items highlighted in yellow fell between 80% and 99% completion (mastery criterion). The tasks highlighted in red indicate areas for improvement (below mastery). Mulan performed six of 11 tasks above standard during the intervention; two of those tasks were completed to 100%. This participant performed the lowest on Task 5 (sterilizing toys within the workspace) which was completed 42% in intervention. Jafar performed 10 of 11 tasks above standard; six of those tasks were completed to 100%. During intervention, this participant completed Task 7 (sanitizing the therapist's chair) 75% of the time which failed to meet the mastery criterion. Rasputin performed eight of 11 tasks above standard; five of those tasks were completed to 100%. There were two tasks in need of significant improvement; sanitizing the client supply tower (Task 4) and washing the client's hands (Task 10) were each completed during 20% of sessions.

Table 3

Average Percentage of Task Completion per Participant

Participant	Condition	Task Number										
		1	2	3	4	5	6	7	8	9	10	11
Mulan	Baseline	0%	0%	0%	0%	0%	20%	0%	0%	20%	0%	0%
	Intervention	75%	92%	75%	92%	42%	83%	100%	100%	50%	67%	92%
Jafar	Baseline	11%	0%	11%	0%	0%	33%	0%	0%	22%	0%	0%
	Intervention	100%	100%	100%	100%	88%	100%	75%	88%	100%	88%	88%
Rasputin	Baseline	23%	0%	0%	0%	0%	31%	0%	0%	23%	0%	0%
	Intervention	100%	100%	80%	20%	60%	100%	100%	100%	80%	20%	80%

Evaluation of the Feedback Meetings

The researcher also collected data on the feedback meetings and looked for trends in performance. Mulan met with the researcher on 12 occasions. During the feedback meetings, this individual was compared to high performer Loki seven times and to high performer Goofy five times. On average, Loki completed 94% of tasks (range, 73% to 100%) and Goofy completed 97% (range, 73% to 100%). Over the course of the study, Mulan's performance exceeded a high performer once, improved following 9 of 12 feedback meetings, and decreased three times after being compared to Loki.

Jafar met with the researcher on eight occasions. This individual was compared to Loki twice and to Goofy six times. Loki completed an average of 89% (range, 82% to 100%) and Goofy completed an average of 94% (range, 73% to 100%). Over the course of the study, Jafar's performance exceeded a high performer on four occasions, improved following three of eight feedback meetings,

maintained at 100% following three meetings, and decreased twice after being compared to Goofy.

Rasputin met with the researcher on five occasions. This individual was compared to Loki three times and Goofy twice. Loki completed an average of 93% (range, 82% to 100%) and Goofy completed an average of 95% (range, 73% to 100%). Overall, Rasputin's performance improved following three of five feedback meetings, maintained at 82% after one meeting, and decreased once after being compared to Goofy.

Social Validity

Prior to the social comparison feedback, low performers reported that they rarely completed nine or more of the end-of-shift cleaning tasks. In intervention, they reported completing nine or more tasks daily. The responses for the remaining questions revealed mixed results. One performer felt neutral about the social comparison graphs, the second agreed that the graphs were beneficial, and the third strongly agreed that the graphs were beneficial. When asked if they enjoyed receiving the feedback, participants reported the same responses. In other words, the same individual who was neutral about the graphs was neutral about receiving feedback. Two of the three low performers strongly agreed that the program should be available to all the employees, while the third participant agreed that the program should be available to all the employees.

Discussion

The aim of this study was to assess the effects of social comparison feedback on the completion of end-of-shift cleaning tasks by lower performing behavior technicians. The results of this study suggest that social comparison feedback involving upward comparisons is an effective strategy to increase cleaning behaviors. Consequently, organizations might implement this cost-effective intervention to address other performance issues.

This study incorporated the most effective components of feedback. For instance, a knowledgeable and experienced researcher collected data on cleaning behaviors, delivered feedback consistently, and communicated the significance of the tasks following the performance. When designing the procedures for feedback delivery, the researcher adhered to the recommendations found in the literature. By individualizing the feedback, the researcher avoided the use of generalized statements about a group performance (Andiola, 2014) and minimized the potential for social loafing (Shepperd, 1995). Also, participants were informed how their individual contributions impacted staff, clients, and the facility. Previous studies successfully implemented verbal plus graphic feedback to increase cleaning behaviors (Austin et al., 2005; Doll et al., 2007); however, graphs were posted weekly and included group data that may have disguised individual improvements (Rohn et al., 2003). As an alternative, the present study incorporated a personalized

bar graph during every feedback meeting to facilitate a quick and easy comparison with coworkers. Contrary to Daniels and Bailey (2014), sharing individual performances did not appear to have an adverse effect on the task completion of low performers. Instead, there was an immediate and rapid change in behavior, as observed by Newby and Robinson (1983). One possible explanation for the increase in task completion could be the use of code names to maintain anonymity; this technique provided participants with anonymity while also holding them accountable. While the results appear to contradict Daniels and Baily (2014), it should be mentioned that it is not a direct contradiction and in fact supports the later point made by the authors. Specifically, while the individual data were shared, the code names allowed individual performance to be unknown and therefore increased privacy. This is perhaps why the technique was effective. Lastly, delivering social comparison content may prove to be as successful as other types of feedback (Alvero et al., 2001). Additional research is needed in the field of organizational behavior management to confirm the function, contexts, and components that make social comparison feedback a viable intervention.

When implementing social comparison feedback, it is important to consider its success in the cognitive literature. For decades, researchers have identified personal and situational variables that may enhance its effectiveness as a consequence strategy. However, researchers have used many different methods and measures of social comparison and, as a result, findings have been difficult to

replicate (Wheeler & Miyake, 1992; Wood, 1996). Specifically, Wood (1996) cautions that the measures are sensitive to extraneous variables and fail to portray social comparison. Researchers in behavior analysis can continue this line of research by designing studies with objective measures. By doing so, they can directly observe how social comparison feedback affects performance. Furthermore, researchers can control for psychological and performance-related attributes by keeping participant identities anonymous. Thus, participants cannot select a target comparison based on age, sex, race, socioeconomic status, or other variable which may confound results. Future studies can focus on the function of social comparison, the contexts in which comparing peer performance improves behavior, and the components of social comparison feedback which are most effective.

Studies suggest that the relevance of the comparison dimension can impact the likelihood to engage in social comparison and Dijkstra et al. (2008) warn that if the dimension is not important to the individual, social comparison may have no effect. Extra precaution was taken in the present study since the dependent variables included a variety of cleaning behaviors. Tasks were expected to be completed with low-probability since this class of behaviors can be unpleasant, physically demanding, and secondary to other job responsibilities (Doll et al., 2007; Shier et al., 2003). Therefore, the social comparison feedback included a brief value statement to "motivate" task completion. The researcher informed participants that

the completion of tasks (a) prepared the workspace for the next shift, (b) assisted other staff with their assigned room duties, (c) prevented clients and staff from getting sick, or (d) resulted in a sense of accomplishment. A different statement was selected during each feedback meeting to communicate the benefit of engaging in the tasks. Social-norm campaigns commonly deliver a value statement because it links prosocial behaviors to specific outcomes and increases the individual's sense of control (Fischer, 2008; Herzing & Jacobson, 2019; Schmiege et al., 2010). Also, it is important to note that this practice not only provides feedback on the current performance, but prompts subsequent behaviors. Future research could examine the effect of social comparison feedback with and without a value statement to determine if increases in behavior can be correlated with specific statements.

Along with relevance of the comparison dimension, literature on social comparison theorizes that individuals select a target based on similar abilities. Since high performers were performing near 100%, it was important to decrease the disparity in performance when sharing the data with lower performers for the first time. To do so, the researcher reviewed the graph of high performers following their first evaluative feedback meeting. At that time, the high performers completed 73% and 82% of tasks whereas low performers completed 0% or 9%. Had the researcher shared perfect scores, lower performers may have deemed the social-norm information as unattainable (Daniels & Bailey, 2014; Schmiege et al., 2010), therefore rendering the social comparison feedback ineffective. For the remainder

of the study, the researcher shared the percentages acquired during the most recent session, noted the highest performer, and made a point to alternate the highest performer when there was a tie. In Hoffman et al. (1954), participants were compared to the same confederate every trial and refrained from comparing performances with the confederate. Researchers suggested that it was due to large discrepancy in scores; however, it is possible that being outperformed by the same individual every time was aversive. Since this study was conducted in an applied setting, the researcher was limited to the performance of high performers. To minimize biases toward a single high performer, the researcher varied the target comparison as much as possible. However, the bar graph likely reminded participants that Goofy was the one to beat. Goofy completed all of the tasks during 83%, 63%, and 90% of feedback meetings with Mulan, Jafar, and Rasputin, respectively. If both comparison targets were performing well, then the participant was compared to a higher standard.

For instance, Mulan was compared to an average of 94% and an average of 97%. Anecdotally, on one occasion, the participant outperformed Goofy and appeared celebratory when referring to the performance. Despite increases in task completion, this participant never seemed to win (outperform both high performers). As a result, the task completion may have been negatively reinforced over time by the social comparison feedback. In other words, the participant completed the tasks to lessen the averseness of the feedback meeting or intensity of

the social comparison. When Mulan received feedback on a perfect performance, anecdotally, the participant's verbal behavior indicated a sense of relief. Future research should measure how the frequency of comparison and disparity between the target performance can contribute to social comparison. By controlling for these variables, researchers might determine under what circumstances social comparison feedback functions as negative reinforcement.

Rasputin's actions also suggest that the social comparison feedback was aversive and, consequently, many of the comments made by this participant were likely an act of self-enhancement. Rasputin appeared to look down when the researcher made a direct comparison. When the bar graph was shared, this participant seemed dismissive toward both high performers regardless of which one the researcher referenced. Individuals typically compare downward to protect their self-esteem (Collins, 1996; Wade & Tavris, 2017), but the participants in this study were selected as low performers. Therefore, they may have looked to strategies to help them cope when downward comparison was not possible. Further empirical research is necessary to determine the context in which low performers compare downward.

Cognitive research also indicates that proximity to a meaningful standard triggers upward comparison and competition (Garcia & Tor, 2007; Garcia et al., 2013; Vandegrift & Holaday, 2012). Whereas these studies incorporated a ranking system, the present study implemented bar graphs as a subtler form of social

comparison. The graphs illustrated the performance of the participant in relation to high performers, but without the winner-take-all design. Since each participant had the potential to reach 100%, they may have used the performance of higher performers to establish a standard, improve their abilities, and/ or set personal goals (Dijkstra et al., 2008; Festinger, 1954; Huguet et al., 2001). According to Daniels and Bailey (2014), goal setting can be an effective intervention when performers do not know which behaviors to engage in or how to perform them well.

In this study, Goofy was the first high performer to obtain a perfect score and to maintain high levels of task completion. Consequently, Goofy's performance most likely served as the standard. For instance, Mulan's performance always increased when specifically compared to Goofy whereas the performance was variable in relation to Loki. In addition, the data showed a steady, increasing trend to 100% which suggests Mulan made smaller, attainable goals. This practice is recommended when the original performance is a long way from the final goal (Daniels & Bailey, 2014). Furthermore, goal attainment (100%) failed to maintain over consecutive sessions. When there is not enough reinforcement to keep the performance going, Daniels and Bailey (2014) warn that extinction will occur and, in the case of low performers, negative reinforcement may be necessary to keep the performance from declining. Since social comparison feedback has the potential to function as negative reinforcement, additional research might investigate how long this intervention can maintain performance.

The motive for social comparison is not as obvious for Jafar. Based on the data, social comparison seemed to have had the greatest impact at the start of the intervention. Jafar may have compared the percentages to make an accurate appraisal of task completion and the self-evaluation could explain the rapid change in performance. Rather than setting smaller goals like Mulan, Jafar completed 100% of the tasks in half the time (three feedback meetings). In addition, the initial achievement of 100% was extra reinforcing; Jafar completed all of the tasks and outperformed both high performers. Since the target comparisons were performing at slightly lower levels (89% and 94%), Jafar tied with one high performer and outperformed the other during three subsequent sessions. Therefore, task completion was positively reinforced by the social comparison feedback. Consequently, the performance maintained at near-perfect levels for the remainder of the study.

A Strategy to Enhance Specific Behaviors

In addition to increasing overall performances, the social comparison feedback could be used to target specific behaviors. Overall, there were three tasks (Tasks 2, 7, and 8) which increased between 92% and 97% following social comparison feedback. The data show that social comparison feedback created awareness about these tasks, the behaviors were part of the participants' repertoire, and additional training was not necessary. As expressed by Geller (1999), the social

comparison feedback provided extrinsic motivation for employees to complete the tasks. In addition, these tasks shared a similar topography. A participant who used a Lysol wipe to sanitize one object was more likely to engage in the other sanitation tasks. That being said, one would have expected similar increases in the sanitation of the client supply tower (Task 4). However, Rasputin frequently neglected this task and brought the average down to 71%.

Although Rasputin completed Task 4 during one session, anecdotally, the Primary Investigator and Research Assistant collecting IOA at the time agreed the behavior was inadvertent. There were two additional sessions in which the researcher could not count the task as complete per the definition and noted that an attempt had been made. Rasputin sanitized the top of the supply tower with a Lysol wipe, but neglected to clean the front. Since the social comparison feedback emphasized the participant's percentage of task completion in relation to peers, Rasputin likely would have continued to make the same error given that task clarification would not have been provided. One alternative would be to implement specific/ social comparison feedback. Williams and Geller (2000) delivered weekly bar graphs to compare different work groups on percent safe scores for each targeted behavior. The researchers found that this intervention was slightly more effective than global/ social comparison feedback.

Another task that participants frequently seemed to neglect was disinfecting toys within the workspace (Task 5). Overall, participants completed the task 0%

during baseline and an average of 63% in intervention. Two of the participants engaged in this task below the company standard following the social comparison feedback; Mulan completed the task an average of 42% and Rasputin completed the task an average of 60%. Upon closer examination, the data suggest that the shift (AM or PM) might have influenced the completion of this task. Trends in the data indicate participants were more likely to skip this task in the morning than in the afternoon. Since clients interacted with many of the same toys during the afternoon shift, participants might have downplayed the benefit of disinfecting toys midday and, therefore, completed the task once at the end of the day. Additional observations are necessary to verify this prediction.

Task 9 showed the least improvement with the addition of social comparison feedback. These results are surprising since participants stacked/ pushed-in chairs an average of 22% during baseline. Following the intervention, the average increased by 55%. Mulan was the only participant to complete the task below the company standard. It is possible participants did not see the relevance or understand the importance of completing this task (Dijkstra et al., 2008; Fischer, 2008). Although the researcher delivered one of four generalized value statements during the feedback meeting, the completion of this task directly aligned with one outcome: helping other staff with their room duties. Had the task been related to the cleanliness of the facility or preventing illnesses, chances are the task would have been completed more frequently. For instance, Mulan sanitized both chairs

during 100% of the intervention sessions, but only stacked/ pushed them in 50% of the time. Future studies might create multiple value statements for each task (specific and general), deliver the statement when the corresponding task is completed incorrectly, and measure how the value statement enhances social comparison feedback. This strategy could provide clarification to participants without the labor-intensive task of creating graphs for each behavior, as in Williams and Geller (2000).

An alternative explanation for the lower percentage of Task 9 might be that participants diffused the responsibility to other staff (Nadler, 1979). In other words, they left the task for someone else to complete. A full-time employee was assigned to the general upkeep and cleanliness of the classroom. Stacking the chairs at the end of the day was on the room checklist, but this task was often neglected. Therefore, the researcher selected this task because it could be completed by all behavior technicians at the end of each client shift. It is possible participants intentionally skipped the task and assumed someone else would stack the chairs. Future studies could compare task completion after each shift, compare task completion for different days of the work week, and then implement social comparison feedback to minimize discrepancies in performance.

Washing hands with soap and water can minimize the spread of germs and prevent the flu. Despite the importance of the task, participants washed the client's hands (Task 10) on average 0% during baseline and 58% in intervention. Mulan

and Rasputin completed the task below company standard for an average of 67% and 20%, respectively. In comparison, these participants washed their own hands an average of 92% and 80%. Since the tasks are similar, one would have expected them to be completed at the same time. However, the data suggest that completion of Task 10 may have required a larger response effort. Many of the clients needed assistance to perform the task correctly (e.g., prompting, modeling, or graduated guidance). Additionally, the child might have engaged in problematic behaviors to escape the task, thereby potentially punishing the participant for making the child wash his or her hands. By completing this task, the participant was also required to collect data on the performance of the client. When the response effort is high, modifying one of the feedback characteristics could enhance performance (So et al., 2013). Therefore, researchers could implement evaluative feedback to increase this behavior in the future.

Return on Investment

No additional costs were procured during the study. The facility provides staff with cleaning supplies and their usage did not exceed the facility's monthly expenditure. In addition, the researchers volunteered their time to collect data and deliver feedback.

Since the data sheet was digitalized, minimal time and effort was necessary for data collection. As the researcher conducted observations, the data were entered

and simultaneously graphed. Participants were given a 25-minute window at the end of the morning and afternoon shifts, but often completed the tasks in less than 10 minutes. Given the set-up of the classrooms, the researchers were able to observe two participants per shift (range, 1 to 4). The majority of tasks involved a permanent product making it possible for a single researcher to collect data on two participants without impacting IOA. Staff members could be trained on the procedures and collect data for the supervisor.

Immediately following observations, a researcher met with the participant in private to review the performance. Social comparison feedback meetings lasted approximately 45 seconds (range, 32 to 71 seconds) whereas evaluative feedback meetings lasted approximately 70 seconds (range, 44 to 120 seconds). Meetings were efficient and did not deter participants from other tasks and activities. In addition, feedback delivery would not require much time from the supervisor.

Anecdotally, one performer commented on how the research was a good thing and that the tasks should be done. Another admitted to not paying attention to the posted memo or the cleaning tasks prior to the study, but plans to continue. Throughout the study, researchers observed more and more staff engaging in the cleaning tasks.

The aim of the study was to implement a cost-effective strategy to increase cleaning behaviors. The data show that social comparison is an effective

intervention; however, further research is necessary to determine if there was a reduction in client and staff call-outs due to sickness.

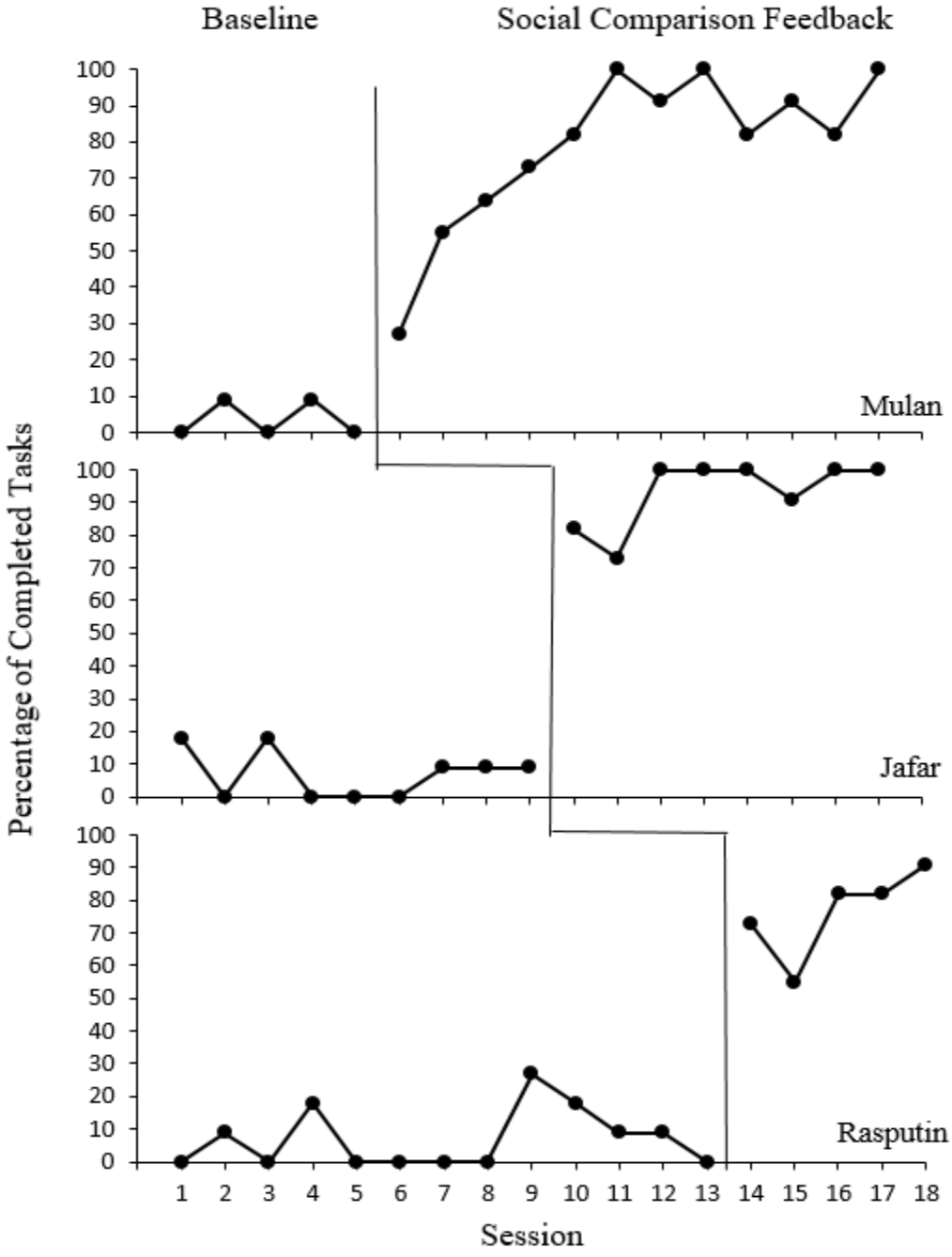


Figure 1. Percentage of completed end-of-shift cleaning tasks across participants during baseline and social comparison feedback conditions.

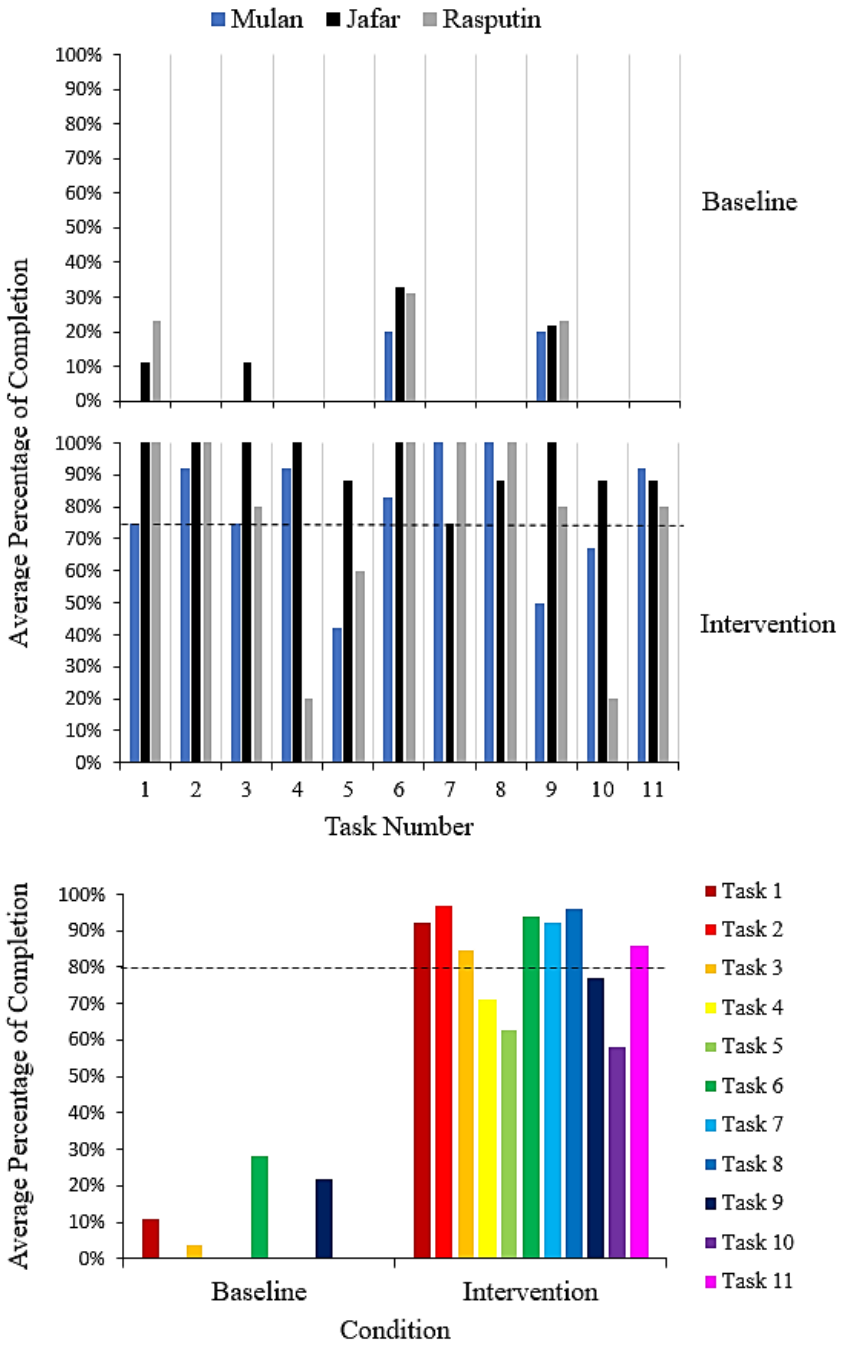


Figure 2. Percentage of end-of-shift cleaning tasks completed by participants during baseline and social comparison feedback conditions.

References

Aboud, F. E. (1976). Self-evaluation: Information seeking strategies for interethnic social comparisons. *Journal of Cross-Cultural Psychology, 7*(3), 289-300. doi:10.1177/002202217673003

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. doi:10.1300/J075v21n01_02

Anderson, D. C., Crowell, C. R., Hantula, D. A., & Siroky, L. M. (1988). Task clarification and individual performance posting for improving cleaning in a student-managed university bar. *Journal of Organizational Behavior Management, 9*(2), 73-90. doi:10.1300/J075v09n02_06

Andiola, L. M. (2014). Performance feedback in the audit environment: A review and synthesis of research on the behavioral effects. *Journal of Accounting Literature, 33*(1-2), 1-36. doi:10.1016/j.acclit.2014.07.001

Andrasik, F., Heimberg, J., & McNamara, J. (1981). Behavior modification of work and work-related problems. *Progress in Behavior Modification, 11*, 117-161. doi:10.1016/B978-0-12-535611-4.50009-9

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. doi:10.1901/jaba.2005.159-03

Ayres, I., Raseman, S., & Shih, A. (2012). Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage. *Journal of Law, Economics, and Organization, 29*(5), 992–1022. doi:10.1093/jleo/ews020

Babcock, R. A., Sulzer-Azaroff, B., Sanderson, M., & Scibak, J. (1992). Increasing nurses' use of feedback to promote infection-control practices in a head-injury treatment center. *Journal of Applied Behavior Analysis*, *25*(3), 621-627. doi:10.1901/jaba.1992.25-621

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. doi:10.1300/J075v07n03_05

Blanton, H., Buunk, B. P., Gibbons, F. X., & Kuyper, H. (1999). When better-than-others compare upward: Choice of comparison and comparative evaluation as independent predictors of academic performance. *Journal of Personality and Social Psychology, 76*(3), 420-430. doi:10.1037/0022-3514.76.3.420

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

Charness, G., Masclet, D., & Villeval, M. C. (2013). The dark side of competition for status. *Management Science*, *60*(1), 38-55. doi: 10.1287/mnsc.2013.1747

Chen, F., Zhang, L., & Latimer, J. (2014). How much has my co-worker contributed: The impact of anonymity and feedback on social loafing in asynchronous virtual collaboration. *International Journal of Information Management, 34*, 652-659. doi:10.1016/j.ijinfomgt.2014.05.001

Collins, R. L. (1996). For better or worse: The impact of upward social comparison on self-evaluations. *Psychological Bulletin, 119*, 51-69. doi:10.1037/0033-2909.119.1.51

Cooper, M. D. (2009). Behavioral safety interventions: A review of process design factors. *Professional Safety, 54*(2), 36-45. Retrieved from https://www.onepetro.org/journal-paper/ASSE-09-02-36

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*(1), 65-71. doi:10.1901/jaba.1988.21-65

Daniels, A. C. (2009). *Oops! 13 management practices that waste time and money (and what to do instead).* Atlanta, GA: Aubrey Daniels International.

Daniels, A. C., & Bailey, J. S. (2014). *Performance management: Changing behavior that drives organizational effectiveness* (5th ed.). Tucker, GA: Performance Management Publications.

Dijkstra, P., Kuyper, H., van der Werf, G., Buunk, A. P., van der Zee, Y. G. (2008). Social comparison in the classroom: A review. *Review of Educational Research, 78*(4), 828-879. doi:10.3102/0034654308321210

Doll, J., Livesey, J., McHaffie, E., & Ludwig, T. D. (2007). Keeping an uphill edge: Managing cleaning behaviors at a ski shop. *Journal of Organizational Behavior Management, 27*(3), 41-60. doi:10.1300/J075v27n03_04

Dumas, F., Huguet, P., Monteil, J. M., Rastoul, C., & Nezlek, J. B. (2005). Social comparison in the classroom: Is there a tendency to compare upward in elementary school. *Current Research in Social Psychology, 10*(12), 166-187. Retrieved from https://search-proquest-com.portal.lib.fit.edu/docview/1821808673?accountid=27313

Durgin, A., Mahoney, A., Cox, C., Weetjens, B. J., & Poling, A. (2014). Using task clarification and feedback training to improve staff performance in an east African nongovernmental organization. *Journal of Organizational Behavior Management, 34*(2), 122-143. doi:10.1080/01608061.2014.914007

Egan, C. (1999). Graphical displays and comparative energy information: What do people understand and prefer? *ECEEE Summer Study 1999*.

Emmert, G. D. (1978). Measuring the impact of group performance feedback versus individual performance feedback in an industrial setting. *Journal of Organizational Behavior Management, 1*(2), 134-141. doi:10.1300/J075v01n02_04

Feldman, N. S., & Ruble, D. N. (1977). Awareness of social comparison interest and motivations: A developmental study. *Journal of Educational Psychology*, *69*(5), 579-585. doi:10.1037/0022-0663.69.5.579

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*(2), 117-140. doi:10.1177/001872675400700202

Fischer, C. (2008). Feedback on household energy consumption: A tool for saving energy. *Energy Efficiency, 1*(1), 79-104. doi:10.1007/s12053-008-9009-7

Garcia, S. M., & Tor, A. (2007). Rankings, standards, and competition: Task vs. scale comparisons. *Organizational Behavior and Human Decision Processes, 102*(1), 95-108. doi:10.1016/j.obhdp.2006.10.004

Garcia, S. M., Tor, A., & Schiff, T. M. (2013). The psychology of competition: A social comparison perspective. *Perspectives on Psychological science, 8*(6), 634-650. doi:10.1177/1745691613504114

Geller, E. S. (1999). Behavior-based safety: Confusion, controversy, and clarification. *Occupational Health & Safety, 68*(1), 40-49. Retrieved from https://search-proquest-com.portal.lib.fit.edu/docview/220996089?accountid=27313

Giffin, K. (1967). The contribution of studies of source credibility to a theory of interpersonal trust in the communication process. *Psychological Bulletin, 68*(2), 104-120. doi:10.1037/h0024833

Gil, P. J., & Carter, S. L. (2016). Graphic feedback, performance feedback, and goal setting increased staff compliance with a data collection task at a large residential facility. *Journal of Organizational Behavior Management, 36*(1), 56-70. doi:10.1080/01608061.2016.1152207

Gilbert, D. T., Giesler, R. B., & Morris, K. A. (1995). When comparisons arise. *Journal of Personality and Social Psychology, 69*(2), 227-236. doi:10.1037/0022-3514.69.2.227

Goethals, G. R., & Darley, J. M. (1977). Social comparison theory: An attributional approach. *Social comparison processes: Theoretical and empirical perspectives*, 259-278.

Golden, G. A., & Cherry, F. (1982). Test performance and social comparison choices of high school men and women. *Sex Roles: A Journal of Research, 8*(7), 761-772. doi:10.1007/BF00287571

Goodman, J. S., Wood, R. E., & Chen, Z. (2011). Feedback specificity, information processing, and transfer of training. *Organizational Behavior and Human Decision Processes, 115*, 253-267. doi:10.1016/j.obhdp.2011.01.001

Gravina, N., Van Wagner, M., & Austin, J. (2008). Increasing physical therapy equipment preparation using task clarification, feedback and environmental manipulations. *Journal of Organizational Behavior Management, 28*(2), 110-122. doi:10.1080/01608060802100931

Gravina, N., Villacorta, J., Albert, K., Clark, R., Curry, S., & Wilder, D. (2018). A literature review of organizational behavior management interventions in human service settings from 1990 to 2016. *Journal of Organizational Behavior Management*, *38*(2-3), 191-224. doi:10.1080/01608061.2018.1454872

Haakana, M., Sillanpää, L., & Talsi, M. (1997). The effect of feedback and focused advice on household energy consumption. In *Proceedings, European Council for an Energy-Efficient Economy*.

Herzing, M., & Jacobsson, A. (2019). Measuring the effects of feedback from inspections on cleanliness in Swedish pre-schools – a field experiment. *Journal of Behavioral and Experimental Economics, 78*, 1-9. doi:10.1016/j.socec.2018.11.002

Harbring, C., & Irlenbusch, B. (2008). How many winners are good to have: On tournaments with sabotage. *Journal of Economic Behavior & Organization*, *65*(3-4), 682-702. doi:10.1016/j.jebo.2006.03.004

Huguet, P., Dumas, F., Monteil, J. M., & Genestoux, N. (2001). Social comparison choices in the classroom: Further evidence for students' upward comparison tendency and its beneficial impact on performance. *European Journal of Social Psychology, 31*(5), 557-578. doi:10.1002/ejsp.81

Hoffman, P., Festinger, L., & Lawrence, D. (1954). Tendencies toward group comparability in competitive bargaining. *Human Relations, 7*, 141-159. doi:10.1177/001872675400700203

Johnson, D. A., & Dickinson, A. M. (2010). Employee-of-the-Month Programs: Do they Really Work. *Journal of Organizational Behavior Management, 30(4),* 308-324. doi:10.1080/01608061.2010.520144

Karau, S. J., & Williams, K. D. (1993). Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology, 65*(4), 681-706. doi:10.1037/0022-3514.65.4.681

LaFleur, T., & Hyten, C. (1995). Improving the quality of hotel banquet staff performance. *Journal of Organizational Behavior Management, 15*(1-2), 69-93. doi:10.1300/J075v15n01_05

Lechermeier, J., & Fassnacht, M. (2018). How do performance feedback characteristics influence recipients' reactions: A state-of-the-art review on feedback source, timing, and valence effects. *Management Review Quarterly, 68*(2), 145–93. doi.10.1007/s11301-018-0136-8

Meisel, C. J., & Blumberg, C. J. (1990). The social comparison choices of elementary and secondary school students: The influence of gender, race, and friendship. *Contemporary Educational Psychology, 15*(2), 170-182. doi:10.1016/0361-476X(90)90015-S

Miller, C. T. (1982). The role of performance-related similarity in social comparison of abilities: A test of the related attributes hypothesis. *Journal of Experimental Social Psychology*, *18*(6), 513-523. doi:10.1016/0022-1031(82)90070-1

Mollen, S., Rimal, R. N., Ruiter, R. A., & Kok, G. (2013). Healthy and unhealthy social norms and food selection. Findings from a field-experiment. *Appetite*, *65*, 83-89. doi:10.1016/j.appet.2013.01.020

Monteil, J. M., & Michinov, N. (2000). Effects of context and performance feedback on social comparison strategies among low-achievement students: Experimental studies. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition, 19*(5-6), 513-531.

Moon, K., Lee, K., Lee, K., & Oah, S. (2017). The effects of social comparison and objective feedback on work performance across different performance levels. *Journal of Organizational Behavior Management, 37*(1), 63-74. doi:10.1080/01608061.2016.1236059

Myers, D., & Twenge, J. (2019). *Social Psychology* (13th ed.). New York, NY: McGraw-Hill Education.

Nadler, D. A. (1979). The effects of feedback on task group behavior: A review of the experimental research. *Organizational Behavior and Human Performance, 23*(3), 309-338. doi:10.1016/0030-5073(79)90001-1

Newby, T. J., & Robinson, P. W. (1983). Effects of grouped and individual feedback and reinforcement on retail employee performance. *Journal of Organizational Behavior Management, 5*, 51-68. doi:10.1300/J075v05n02_03

Nordstrom, R., Lorenzi, P., Hall, R. V. (1991). A review of public posting of performance feedback in work settings. *Journal of Organizational Behavior Management, 11*, 101-124. doi:10.1300/JO75v11n02_07

Pampino, R. N., MacDonald, J. E., Mullin, J. E., & Wilder, D. A. (2003). Weekly feedback vs daily feedback: An application in retail. *Journal of Organizational Behavior Management, 23*(2-3), 21-43. doi:10.1300/J075v23n02_03

Panyan, M., Boozer, H., & Morris, N. (1970). Feedback to attendants as a reinforcer for applying operant techniques. *Journal of Applied Behavior Analysis*, *3*(1), 1-4. doi:10.1901/jaba.1970.3-1

Reetz, N. K., Whiting, S. W., & Dixon, M. R. (2016). The impact of a task clarification and feedback intervention on restaurant service quality. *Journal of Organizational Behavior Management, 36*(4), 322-331. doi://10.1080/01608061.2016.1201035

Régner, I., & Monteil, J. (2007). Low- and high-socioeconomic status students preference for ingroup comparisons and their underpinning ability expectations. *Revue Internationale De Psychologie Sociale, 20*(1), 87-104. Retrieved from https://search-proquest-com.portal.lib.fit.edu/docview/621742432?accountid=27313

Reuman, D. A. (1989). How social comparison mediates the relation between ability-grouping practices and students' achievement expectancies in mathematics. *Journal of Educational Psychology, 81*(2), 178-189. doi:10.1037/0022-0663.81.2.178

Robinson, E., Fleming, A., & Higgs, S. (2014). Prompting healthier eating: Testing the use of health and social norm based messages. *Health Psychology*, *33*(9), 1057-1064. doi: 10.1037/a0034213

Rohn, D., Austin, J., & Lutrey, S. M. (2003). Using feedback and performance accountability to decrease cash register shortages. *Journal of Organizational Behavior Management, 22*(1), 33-46. doi:10.1300/J075v22n01_03

Rose, H. M. S., & Ludwig, T. D. (2009). Swimming pool hygiene: Self-monitoring, task clarification, and performance feedback increase lifeguard cleaning behaviors. *Journal of Organizational Behavior Management, 29*(1), 69-79. doi:10.1080/01608060802660157

Ruble, D. N., Boggiano, A. K., Feldman, N. S., & Loebl, J. H. (1980). Developmental analysis of the role of social comparison in self-evaluation. *Developmental Psychology*, *16*(2), 105-115. doi:10.1037/0012-1649.16.2.105

Schmiege, S. J., Klein, W. M. P., Bryan, A. D. (2010). The effect of peer comparison information in the context of expert recommendations on risk perceptions and subsequent behavior. *European Journal of Social Psychology, 40*, 746-759. doi:10.1002/ejsp.645

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science, 18*(5), 429–34. doi:10.1111/j.1467-9280.2007.01917.x

Seidner, L. B., Stipek, D. J., & Feshbach, N. D. (1988). A developmental analysis of elementary school-aged children's concepts of pride and embarrassment. *Child Development*, 367-377. doi:10.2307/1130316

Shepperd, J. A. (1995). Remedying motivation and productivity loss in collective settings. *Current Directions in Psychological Science, 4*, 131-140. doi:10.1111/1467-8721.ep10772565

Shier, L., Rae, C., & Austin, J. (2003). Using task clarification, checklists and performance feedback to improve the appearance of a grocery store. *Performance Improvement Quarterly, 16*(2), 26-40. doi:10.1111/j.1937-8327.2003.tb00277.x

Slowiak, J. M. (2014). "How may I help you?" improving telephone customer service in a medical clinic setting. *Journal of Organizational Behavior Management, 34*(1), 39-51. doi:10.1080/01608061.2013.873382

So, Y., Lee, K., & Oah, S. (2013). Relative effects of daily feedback and weekly feedback on customer service behavior at a gas station. *Journal of Organizational Behavior Management, 33*(2), 137-151. doi:10.1080/01608061.2013.785898

Steelman, L. A., & Rutkowski, K. A. (2004). Moderators of employee reactions to negative feedback. *Journal of Managerial Psychology*, *19*(1), 6-18. doi:10.1108/02683940410520637

Suls, J., Gastorf, J., & Lawhon, J. (1978). Social comparison choices for evaluating a sex- and age-related ability. *Personality and Social Psychology Bulletin, 4*(1), 102-105. doi:10.1177/014616727800400121

Tesser, A., & Smith, J. (1980). Some effects of task relevance and friendship on helping: You don't always help the one you like. *Journal of Experimental Social Psychology, 16*, 582-590. doi:10.1016/0022-1031(80)90060-8

Therrien, K., Wilder, D. A., Rodriguez, M., & Wine, B. (2005). Preintervention analysis and improvement of customer greeting in a restaurant. *Journal of Applied Behavior Analysis, 38*(3), 411-415. doi:10.1901/jaba.2005.89-04

Thorndike, A. N., Riis, J., & Levy, D. E. (2016). Social norms and financial incentives to promote employees' healthy food choices: A randomized controlled trial. *Preventive Medicine, 86*, 12–18. doi:10.1016/j.ypmed.2016.01.017

Ueno, T., Inada, R., Saeki, O., & Tsuji, K. (2005). Effectiveness of displaying energy consumption data in residential houses. Analysis on how the residents respond. In *Proceedings, European Council for an Energy-efficient Economy, 6*, 1289-1299.

Vandegrift, D., & Holaday, B. (2012). Competitive behavior: Tests of the N-effect and proximity to a standard. *Journal of Neuroscience, Psychology, and Economics*, *5*(3), 182-192. doi:10.1037/a0029221

Van Houten, R., Hill, S., Parsons, M. (1975). An analysis of a performance feedback system: The effects of timing and feedback, public posting, and praise upon academic performance and peer interaction. *Journal of Applied Behavior Analysis, 4*, 449-457. doi:10.1901/jaba.1975.8-449

Wade, C., & Tavris, C. (2017). *Psychology*. (12th ed.). Upper Saddle River, NJ: Pearson Education, Inc.

Welsch, W. W., Ludwig, C., Radiker, J. E., & Krapfl, J. (1973). Effects of feedback on daily completion of behavior modification projects. *Mental Retardation, 11*, 24-27.

Wheeler, L., & Miyake, K. (1992). Social comparison in everyday life. *Journal of personality and social psychology*, *62*(5), 760-773. doi:10.1037/0022-3514.62.5.760

Whittemore, I. C. (1925). The competitive consciousness. *Journal of Abnormal and Social Psychology, 20*, 17-33. doi:10.1037/h0072083

Wikoff, M. B., Anderson, D. C., & Crowell, C. R. (1982). Behavior management in a factory setting: Increasing work efficiency. *Journal of Organizational Behavior Management, 4*, 97-128. doi:10.1300/J075v04n01_04

Wilhite, H., Hoivik, A., & Olsen, J. G. (1999). Advances in the use of consumption feedback information in energy billing: the experiences of a Norwegian energy utility. In *Proceedings, European Council for an Energy-Efficient Economy*.

Wilk, L. A., & Redmon, W. K. (1997). The effects of feedback and goal setting on the productivity and satisfaction of university admissions staff. *Journal of Organizational Behavior Management, 18*(1), 45-68. doi:10.1300/J075v18n01_04

Williams, J. H., & Geller, E. S. (2000). Behavior-based intervention for occupational safety: Critical impact of social comparison feedback. *Journal of Safety Research*, *31*(3), 135-142. doi:10.1016/S0022-4375(00)00030-X

Wood, J. V. (1996). What is social comparison and how should we study it? *Personality and Social Psychology Bulletin*, *22*(5), 520–537. doi:10.1177/0146167296225009

Appendix A
End-of-Shift Tasks for Therapists

	Task	Definition
1	Remove all materials from top of workspace/ desk	Place materials (e.g., Stimuli, timers, reinforcers, clipboard) and client items inside the desk.
2	Sanitize the workspace/ desk	Use at least one fresh Lysol wipe to clean the top surface until entire area is visibly wet. Let air dry.
3	Remove all materials from top of the client's supply tower	Place materials (e.g., Stimuli, timers, reinforcers, clipboard) and client items inside the supply tower.
4	Sanitize the client's supply tower	Use at least one fresh Lysol wipe to clean the top and front of the supply tower until entire area is visibly wet. Let air dry.
5	Sterilize toys in workspace	Use the Lysol spray to disinfectant toys within 4 feet of the workspace.
6	Put toys away	Return toys within 4 feet of the workspace to the designated bins and/ or the supply tower.
7	Sanitize the client chair	Use at least one fresh Lysol wipe to clean the client's chair until the entire seat is visibly wet. Let air dry.
8	Sanitize the therapist chair	Use at least one fresh Lysol wipe to clean the therapist's chair until the entire seat is visibly wet. Let air dry.
9	Stack or push-in chairs	Stack therapist and client chair against the wall at the end of the day. Push-in chairs at desk after AM shift.
10	Wash client's hands	Help client wash hands with soap and water.
11	Wash hands	Use soap and water to clean hands.

End-of-Shift Checklist		
	Task	**+ or -**
1	Remove all materials from top of workspace/ desk	
2	Sanitize the workspace/ desk	
3	Remove all materials from top of the client's supply tower	
4	Sanitize the client's supply tower	
5	Sterilize toys in workspace	
6	Put toys away	
7	Sanitize the client chair	
8	Sanitize the therapist chair	
9	Stack or push-in chairs	
10	Wash client's hands	
11	Wash hands	

Appendix B
Informed Consent

Title: The Effect of Feedback on Task Completion in a Human-Service Setting
Principal Investigator: Curtis Phillabaum
Faculty Advisor: Rachael Tilka, Ph.D.

Introduction / Purpose: Our research team wants to study the effectiveness of different strategies to enhance the appearance and cleanliness of the facility. Specifically, the researcher will implement graphic and/or verbal feedback to attempt to enhance the cleaning behaviors of staff.

Procedures and Benefits: If you agree to participate, researchers will evaluate the completion of end-of-shift tasks and introduce strategies to attempt to enhance performance. The researcher will graph your performance relative to the group (you will be given a code name to maintain anonymity) and/or provide verbal feedback on your task completion. Data collected during this study could potentially be published and used by others in the field.

Potential Risks of Participating: Your participation does not involve any risks other than what you would encounter in everyday life. Participants will not be compensated for participation, but will be paid their typical hourly pay during the course of the study.

Potential Benefits of Participating: Participation in this study could help assist with clarifying tasks that should be performed and could also help you to increase performance with the completion of specific tasks in your workplace.

Confidentiality: Identifiable information will not be collected and names of participants will be changed to keep confidentiality. Digitized data will be stored on password protected laptops. Data will be destroyed after seven years or at the written request of the participant.

Participants' Rights: Participation is voluntary. Refusal to participate will involve no penalty or loss of benefits to which you are otherwise entitled. You may discontinue participation at any time without penalty. Participation or non-participation will not affect your standing with the company. Your individual-level results are for research purposes only.

Contact: Any questions regarding this study can be directed to Curtis Phillabaum at 443-846-4317 or through email at cphillabaum2016@my.fit.edu.

Whom to contact about your rights as a research participant in the study: This study was approved by Florida Institute of Technology's IRB. The current IRB chair may be contacted for questions about the rights of people who take part in research. Her contact information is as follows:

Dr. Jignya Patel, IRB Chairperson
150 West University Blvd.
Melbourne, FL 32901
Email: FIT_IRM@fit.edu
Phone: 321-674-7347

Your signature below indicates that you agree to participate in this study and that you have read and understood all information provided above.

Participant's Full Name _____ Date_____
Participant's Signature _____

Appendix C
Data Sheet

Date: **Shift: AM / PM** **Observer Initials:** **Participant Code:**

Condition: BL / FB **IOA: Y / N** **FB Assessor:**

	Task	1 or 0	Comments
1	Remove all materials from top of workspace/ desk		
2	Sanitize the workspace/ desk		
3	Remove all materials from top of the client's supply tower		
4	Sanitize the client's supply tower		
5	Sterilize toys in workspace		
6	Put toys away		
7	Sanitize the client chair		
8	Sanitize the therapist chair		
9	Stack or push-in chairs		
10	Wash client's hands		
11	Wash hands		
	Total correct		**Percentage of completion = / 11 =**

Appendix D
Company Memo

Memo: All staff are encouraged to ensure that classrooms are clean and in order following each shift with a client and that clients are prepared to reunite with caretakers. Please do your best to ensure the following tasks are getting done at the end of your shift with a client.

	Task	**Definition**
1	Remove all materials from top of workspace/ desk	Place materials (e.g., Stimuli, timers, reinforcers, clipboard) and client items inside the desk.
2	Sanitize the workspace/ desk	Use at least one fresh Lysol wipe to clean the top surface until entire area is visibly wet. Let air dry.
3	Remove all materials from top of the client's supply tower	Place materials (e.g., Stimuli, timers, reinforcers, clipboard) and client items inside the supply tower.
4	Sanitize the client's supply tower	Use at least one fresh Lysol wipe to clean the top and front of the supply tower until entire area is visibly wet. Let air dry.
5	Sterilize toys in workspace	Use the Lysol spray to disinfectant toys within 4 feet of the workspace.
6	Put toys away	Return toys within 4 feet of the workspace to the designated bins and/ or the supply tower.
7	Sanitize the client chair	Use at least one fresh Lysol wipe to clean the client's chair until the entire seat is visibly wet. Let air dry.
8	Sanitize the therapist chair	Use at least one fresh Lysol wipe to clean the therapist's chair until the entire seat is visibly wet. Let air dry.
9	Stack or push-in chairs	Stack therapist and client chair against the wall at the end of the day. Push-in chairs at desk after AM shift.
10	Wash client's hands	Help client wash hands with soap and water.
11	Wash hands	Use soap and water to clean hands.

Appendix E
Graphic and Verbal Feedback Script for Lower Performers

1. Greet the participant.

2. Percentage of task completion.

Percentage
"Overall you completed __ out of 11 tasks, or __%."

3. Discuss the performance in relation to the other participants.

Social Comparison
"This graph shows your performance in relation to some of your coworkers. For example, [Insert highest Performer] completed __ % of the tasks."

4. Deliver a value statement. Select one statement from below and vary it for each delivery.

	Value Statements: "By completing the tasks…
1	The center is clean and organized.
2	You are helping other staff with the room duties.
3	Clients and staff are less likely to get sick.
4	You can be proud of your *contribution*.

5. Questions for the researcher.
    a. Refer participant to the posted memo when necessary.

Questions
"Do you have any questions for me?"

6. Reminders for the participant.
    a. The researcher will deliver ONE reminder every time.

Reminders: "Thanks for your help. Please remember…	
1	You can utilize the last 10-minutes of each session to engage in tasks.
2	Check in with the researcher following your next shift.
3	Do not discuss any details about the study with anyone.

Appendix F
Evaluative Feedback Script for High Performers

1. Greet the participant.

2. Performance. State whether the performance maintained, improved, or decreased:

Performance: "Since last session your performance…	
1	Maintained
2	Improved
3	Decreased

3. Delivery of a praise statement.
    a. If performance decreased, do not include a praise statement.
    b. If performance maintained or improved, select one statement from below and vary it for each delivery. Italicized words can be substituted with other descriptive words.
        i. Highlighted statements are for **HPs** who complete at least 9 tasks (82%).

Praise Statement	
1	Thanks for completing the tasks.
2	*Good* job completing the tasks.
3	Keep up the *great* work, [Name].
4	You completed more than 80% of the tasks!
5	I wish others would perform as well as you!
6	[Name], you are a *Rock Star!*

4. Delivery of positive feedback. Review the completed data sheet with the participant and list ALL the tasks that were performed well.

Positive Feedback: "Let's take a look at what you performed well…	
1	"Here's what you completed: __."
2	"You accurately __."
3	"You remembered to __."

5. Delivery of corrective feedback for less-than-perfect performances. Review the completed data sheet and identify ONE incomplete task & how to improve it for next time.

    a. If 100% completion, acknowledge that no corrective action is currently necessary and to keep up the great work.

Corrective Feedback: "As you can see, there is room for improvement…			
	Step 1: Incomplete Task		**Step 2: How to Improve**
1	"We could not count __."	1	"Next time __."
2	"We did not see you __."	2	"Make sure to __."
3	"You missed __."	3	"Give yourself time to __."

6. Percentage of task completion.

Percentage
"Overall you completed __ out of 11 tasks, or __%."

7. Deliver a value statement. Select one statement from below and vary it for each delivery. The highlighted statements are exclusive to **HPs** who complete at least 9 tasks (82%).

Value Statements: "By completing the tasks…	
1	The center is clean and organized.
2	You are helping other staff with the room duties.
3	Clients and staff are less likely to get sick.
4	You can be proud of your *contribution*.
5	You are setting a great example for your coworkers!
6	You are killing hundreds, maybe thousands of germs!

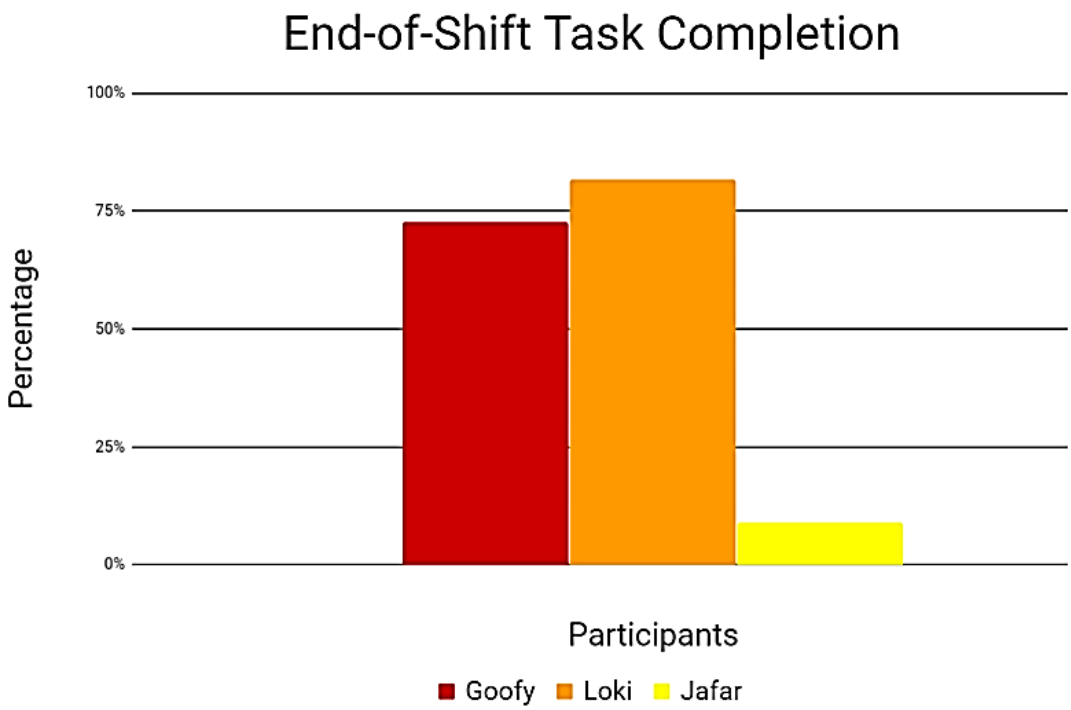
8. Questions for the researcher.
    a. Refer participant to the posted memo when necessary.

Questions
"Do you have any questions for me?"

9. Reminders for the participant.
    a. The researcher will deliver ONE reminder every time.

Reminders: "Thanks for your help. Please remember…	
1	You can utilize the last 10-minutes of each session to engage in tasks.
2	Check in with the researcher following your next shift.
3	Do not discuss any details about the study.

Appendix G
Social Comparison Bar Graph



Appendix H
Social Validity Questionnaire

Complete the following survey. Your responses will be kept confidential and anonymous. Please, answer truthfully. On a scale of 1-5, circle the number that best indicates your response

1. Prior to receiving feedback, how many times per week did you complete 9 or more of the cleaning tasks?

1	2	3	4	5
Rarely		3 days		Daily

2. After receiving feedback, how many times per week did you complete 9 or more of the cleaning tasks?

1	2	3	4	5
Rarely		3 days		Daily

3. I found the social comparison graphs beneficial.

1	2	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly agree

4. I believe this program should be made available to all the employees.

1	2	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly agree

5. I enjoyed receiving the feedback.

1	2	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly agree