Western Michigan University
ScholarWorks at WMU

Dissertations Graduate College

4-2019

A Component Analysis of an Electronic Data Collection Package

Cody Morris
Western Michigan University, morris_33@live.com

Follow this and additional works at: https://scholarworks.wmich.edu/dissertations

Part of the Applied Behavior Analysis Commons

Recommended Citation

Morris, Cody, "A Component Analysis of an Electronic Data Collection Package" (2019). *Dissertations*. 3414.
https://scholarworks.wmich.edu/dissertations/3414

This Dissertation-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Dissertations by an authorized administrator of ScholarWorks at WMU. For more information, please contact wmu-scholarworks@wmich.edu.

A COMPONENT ANALYSIS OF AN ELECTRONIC DATA COLLECTION PACKAGE

by

Cody Morris

A dissertation submitted to the Graduate College
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
Psychology
Western Michigan University
April 2019

Doctoral Committee:

Stephanie Peterson, Ph.D., Chair
Wayne Fuqua, Ph.D.
Heather McGee, Ph.D.
Lloyd Peterson, Ph.D.

Copyright by
Cody Morris
2019

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my appreciation to everyone who supported me in the process of conducting this project and completing my doctoral training. It has been a long and rewarding journey that I could not have completed by myself.

First, I'd like to thank Dr. Stephanie Peterson, my advisor and mentor. Taking research methods with you when I was an undergraduate ended up being a very lucky decision on my part. Ever since you drew my interest to behavior analysis in that class, you've stuck by me facilitating learning opportunities and imparting your wisdom. Your influence on my development as a behavior analyst cannot be overstated.

Thank you to everyone involved with this research project. To my committee members, Dr. Wayne Fuqua, Dr. Heather McGee, and Dr. Lloyd Peterson, thank you for the guidance on this project and involvement in my education. It has been an honor to have worked with and learned from you all over the years. To Kelsey Webster and Daphne Snyder, thank you for your assistance and support on this project.

To the current and former members of the Behavioral Assessment and Treatment Lab, it has been a lot of fun, and I appreciate all of the support along the way. I'm especially thankful to Yannick Schenk for his mentorship throughout my thesis project and Nate VanderWeele and Kelsey Webster for collaborating with me on many projects.

Thank you to everyone that I've worked with on the Psychological Assessment and Treatment Services (PATS) team. It has been the absolute joy of my life to work with you all throughout my time in PATS. I can't tell you how much I've learned from working with you all

Acknowledgements—Continued

over the years and how proud I am of the work we've accomplished. I wish you all the absolute best in your careers and hope that you will take a little piece of PATS with you wherever you go.

I'd also like to thank the administrative staff in the Department of Psychology who have supported me throughout my involvement in the department. Thank you to Amber Hutson, Kaycie James, and Bette Ludwig for all of the help over the years. I am especially grateful to Kim Tembreull who worked tirelessly to ensure the projects I worked on ran smoothly and that I was taken care of.

Finally, I'd like to thank my friends and family who have supported me throughout this journey. To Roxie and Ronnie Swank, I couldn't have done any of this without your continued love and support. To Todd Russell, thank you for the push in this direction! I owe you all so much.

-It takes a village to raise a graduate student.

Cody Morris

A COMPONENT ANALYSIS OF AN ELECTRONIC DATA COLLECTION PACKAGE

Cody Morris, Ph.D.

Western Michigan University, 2019

Data collection is essential to the practice of applied behavior analysis, but human error in collection can lead to inaccuracies. Because inaccuracies in measurement may adversely affect treatment decisions, procedures to increase data collection fidelity are necessary. This is especially important in settings wherein behavior analysts rely on others to report data. Procedures for training and directly supervising data collectors do exist, however, few resources exist for data collectors working with limited supervisor presence. Electronic data collection (EDC) systems are uniquely positioned to help address this need, but little research exists analyzing active components of EDC systems. Therefore, the purpose of this study is to systematically evaluate the individual components of an EDC system on data collection fidelity of caregivers in a home setting in the absence of a supervisor.

TABLE OF CONTENTS

Table of Contents—Continued

APPENDICES

INTRODUCTION

Measuring behavior is a cornerstone of behavior analysis (Baer, Wolf, & Risley, 1968; Sidman,1960/1988; Johnston & Pennypacker, 1993; Cooper, Heron, & Heward, 2007). Johnston and Pennypacker (1993) describe measurement as the process of attaching numbers to events to distinguish them from other events. The numbers derived from measurement are data, which is the primary material used by behavior analysts to evaluate their work (Cooper et al., 2007). As Sidman (1960/1988) stated, the chief criterion in evaluating behavioral intervention is the resulting data. The Behavior Analyst Certification Board's (BACB) *Professional and Ethical Compliance Code for Behavior Analysts* (2014) even lists collecting and graphically displaying data, so that data can be used for decisions and recommendations for behavior change program development as an obligation for behavior analysts.

Because of the reliance of data in behavior analysis, effort has been made to help behavior analysts select effective measurement systems for obtaining accurate data (LeBlanc, Raetz, Sellers, & Carr, 2016; Johnston & Pennypacker, 1993; Cooper et al., 2007). Obtaining accurate data is important because inaccuracies in data could inadvertently lead to incorrect treatment decisions made by behavior analysts (Taber-Doughty & Jasper, 2012; Cooper et al., 2007). In fact, LeBlanc and colleagues (2016) argued that the practice of applied behavior analysis is invalid in the absence of meaningful data.

Despite developing and utilizing quality measurement systems, issues with data accuracy persist. Of the many factors contributing to issues with data accuracy, the biggest threat is human error (i.e., observers not collecting/recording data as it is intended to be collected/recorded; Johnston & Pennypacker, 1993; Cooper et al., 2007). Cooper et al. point to inadequate observer training and unintended influences on the observer as the main sources of human error in measurement. This is exacerbated when behavior analysts rely on other members of the treatment team (e.g., parents, teachers, and direct-care staff) to collect the data, especially in behavioral consultation models where oversight may be limited (Madsen, Peck, & Valdovinos, 2015; Sleeper et al., 2017; Dixon, 2003; Whiting & Dixon, 2012; Reis, Wine, & Brutzman, 2013).

One solution to issues with data collection fidelity is supervisory presence and feedback. A study by Mozingo, Smith, Riordan, Reiss, and Bailey (2006) demonstrated direct-care staff's difficulty in collecting accurate data without supplemental support and effective supervision of the data collection. The supervision provided to direct-care staff in the study consisted of the supervisor observing staff collect data while also collecting data on the client's behaviors during 6-min probes. This data was then used to provide feedback pertaining to agreements and disagreements in the recorded data. After multiple exposures, the feedback was completely faded out, but the supervisor still conducted frequent observations. In a similar study by Reis et al. (2013), participants were asked to collect data while watching training videos. Upon completion of each training video and data collection session, the participants' data was immediately reviewed by a supervisor who provided feedback on correct and incorrect elements. The results of this study showed that participants' data collection accuracy improved following the intervention.

Although the studies by Reis et al. (2013) and Mozingo et al. (2006) indicate that supervisor delivered feedback can improve data collection, the intervention requires that supervisors be present and able to provide feedback across multiple opportunities (Reis et al., 2013). This can be problematic in settings that utilize lone workers due to difficulties in coordinating observations and observer reactivity (Olson & Austin, 2001; Hickman & Geller 2005; Olson, 2018). In applied settings, clinicians responsible for coordinating all facets of treatment may have limited time to focus on observing data collectors and providing them feedback. In addition, the limited time clinicians can observe data collection may lead to an inflated and unrepresentative sample due to observer reactivity (Cooper et al., 2007). So, although clinician feedback/presence may alter data collection while they are present, maintenance of those improvements in the absence of supervisor presence may be challenging. Behavior analysts working in settings that necessitate the use of data collected by others would benefit from resources to ensure data collection fidelity.

An alternative solution to issues with data collection fidelity is electronic data collection (EDC; Morris, 2016; Sleeper et al., 2017; Dixon, 2003). There are two categories of EDC; one can transduce the data (i.e., record the behavior automatically without supplementary input) and the other requires input from an observer. In applied settings focused on assessing and treating problem behaviors, the latter is more common. For the rest of this paper, the term "EDC" will refer to programs that require input from an observer.

Few studies have directly compared EDC to standard paper-and-pencil systems in terms of accuracy or fidelity. A study by Tapp et al., (2006) compared paper-and-pencil data recording to a computer program called, "INTMAN", that was designed to assist data collection. The results of the study showed that the INTMAN program was not only more accurate than the paper-and-pencil program, but also more efficient. However, a study by Tarbox, Wilke, Findel-Pyles, Bergstrom, and Granpeesheh (2010) also compared paper-and-pencil data recording to a computer program and indicated that the electronic system was no more accurate than the paper-and-pencil system and took more time to complete. One limitation of both of these studies is the speed at which technology becomes outdated and the lack of programmed interventions that could affect data collection performance.

In a more recent study, Morris (2016) compared a paper-and-pencil data collection system to an EDC package with unsupervised direct-care staff. The direct-care staff who participated in the study had been trained to collect data using a paper-and-pencil system and had been using that system on a regular basis as part of their basic job responsibilities. When the EDC package was introduced, the staff were trained to use the electronic system, but no supplementary instruction regarding data collection practices was provided. The results of the study showed that in the absence of supervisor presence and using a paper-and-pencil data collection system, direct-care staff entered data at the time instructed to an average of 8% of the time, entered the correct number of intervals containing problem behavior an average of 52% of the time, and indicated the correct interval in which behavior occurred an average of 40% of the time. When the EDC package was introduced, replacing the paper-and-pencil data collection system, all measures of data collection fidelity increased. During this phase, unsupervised direct-care staff entered data at the time instructed an average of 95% of the time and entered the correct number of intervals as well as indicated the correct interval that the behavior occurred an average of 92% of the time for each.

Although Morris (2016) demonstrated that an EDC system could produce an increase in data collection fidelity in the absence of supervisor presence, there were many limitations to the study. The first limitation of the study was that functional control over the behavior was not demonstrated due to unrelated changes in the experimental environment that prevented a planned reversal. The second limitation of the study was that the EDC system was introduced as a package system. Because all of the components were implemented simultaneously, it is unclear what aspects of the system were actively affecting direct-care staff's data collection behavior.

The EDC package used in the Morris (2016) study consisted of an electronic data sheet, automated prompts, and two forms of programmed feedback. The electronic data sheet was included into the package because it was the basis necessary for the rest of the package. The data sheet included drop-down menus, cells that could be typed into, and basic instructions. The appearance of the data sheet was designed to replicate the paper-and-pencil data collection system that it replaced. The rest of the package (i.e., automated prompts and feedback) were selected based on empirical support for their utility and feasibility of including them into an EDC system.

Automated prompts were included into the EDC package because previous research, including EDC studies, suggested its effectiveness at increasing desired behavior. In a study by Realon, Favell, and McGimsey (1992), a computer program was successfully used to prompt staff interaction with clients while automatically recording data related to the interaction. In addition to Realon et al., other non-EDC studies such as Van Houten and Sullivan (1975) and Van der Mars (1988) have demonstrated the effectiveness of automated prompts on increasing the praise rates of teachers in educational settings. The automated prompts included in the Morris (2016) study consisted of an auditory signal and a descriptive message that appeared on the apparatus.

Automated feedback was also included in the package due to a myriad of empirical studies supporting automated feedback as a tool for improving performance, including technology-based interventions. Studies by Moon and Oah (2013), Goomas (2012a), and Goomas (2012b) all demonstrated the utility of technology-based automated feedback on improving targeted behaviors. Importantly, in a study directly comparing computer generated feedback to human delivered face-to-face feedback by Warrilow (2017), computer generated feedback was shown to be as effective as face-to-face feedback. This indicates that the face-to-face feedback shown to be effective at improving data collection in the Reis et al. (2013) and Mozingo et al. (2006) studies could possibly be replaced with feedback delivered via technology. The type of feedback provided in the Morris (2016) study were two forms of graphic feedback that specified the percent of data entries that were entered on-time or late/early. One form of the graphic feedback provided information regarding averaged performance over an hour-long period and the other form of graphic feedback provided information regarding specific (i.e., interval-by-interval) performance. Both forms of feedback were presented concurrently, directly below the data collection table.

In addition to suggesting that EDC packages can increase data collection fidelity of unsupervised direct-care staff, Morris (2016) showed the utility of timeliness of data entry as a proxy measure of data collection fidelity. Throughout the study, data collection timeliness correlated with the other two measures of data collection fidelity. At the beginning of the study low baseline timeliness data aligned with low scores of the other measures of data collection fidelity. When the intervention package that consisted of interventions only targeting timeliness was introduced, an increase in timeliness occurred along with an increase in the other accuracy measures. When the timeliness of data entry decreased due to extraneous variables, other accuracy measures decreased as well. This covariation suggests that adequate timeliness of data collection aligns with adequate data collection fidelity measures such as correctly indicating the specific interval containing target behavior. This finding is supported by other studies that demonstrate the importance of timeliness in time-sampling procedures (Repp, Roberts, Slack, Repp, & Berkler, 1976; Bijou, Peterson, & Ault, 1968)

The correlation between timeliness of data entry and accuracy shown in the Morris (2016) study is of interest because timeliness is the only measure related to data collection that can be collected as a true value via permanent product. Other related measures of data collection can only be collected with observed measures and would require agreement between multiple observers. Although these measures could be of interest despite only being observed measures, they require at least one and sometimes two observers (for interobserver agreement) and would therefore potentially produce observer reactivity that could confound measurement.

Alternatively, the timeliness measure can be unobtrusively collected by automatically timestamping the data when entered into the system, thus no observer is necessary. Therefore, the results of the Morris (2016) experiment suggest that studies interested in data collection fidelity in the absence of a supervisor could consider using timeliness measures without other measures of data collection fidelity that would require direct observation.

Although package treatments are sometimes necessary, for a treatment to be considered analytic, researchers must identify and isolate active and essential parts of the package to determine their individual utility. Additionally, in identifying the essential components, the treatment may be made more efficient and improve social validity (Ward-Horner & Sturmey, 2010). Therefore, the purpose of this study was to systematically evaluate the individual components of an EDC system similar to the Morris (2016) study on data collection timeliness of caregivers in a home setting in the absence of supervisor presence.

METHOD

Participants and Setting

This study was conducted with three caregivers (i.e., two parents and one direct-care staff) of children who engaged in challenging behaviors and received home-based behavioral consultative services. The caregivers who participated in the study were given the pseudonyms of Christine, Lindsey, and Steve. Christine and Lindsey were both parents of children receiving services for problem behavior. Steve was the direct-care staff of a child receiving services for problem behavior. As part of standard service, the caregivers had already agreed to collect probe data on the children's behavior.

During data collection probes, caregivers were asked to specifically focus on the data collection task and minimize distractions. The exact recording periods used in this study varied by participant but remained consistent for each participant throughout their involvement. When a caregiver agreed to participate in the study, the primary investigator discussed appropriate recording periods with them. Through this conversation, a recording period was mutually agreed on between the participant and the primary investigator. The criteria for the recording period consisted of (a) the child being present and likely to engage in problem behaviors (i.e., the child was home, awake, and had a history of engaging in problem behaviors during said timeframe), (b) the recording period was feasible (i.e., the caregiver could minimize distractions like needing to cook dinner), and (c) someone else was available to assist the child when necessary so the participant could focus solely on collecting data.

Subject Recruitment/ Informed Consent Process

Caregivers of children already receiving consultative services for severe problem behaviors were asked to participate in the study. The caregivers had already agreed to or begun collecting probe data on the child's behavior as part of standard treatment. The inclusion criteria for participation were (a) the individual was the parent or direct-care staff of a child who received consultative services for severe problem behavior, (b) they were expected to continue involvement in services for at least 6 weeks, (c) they had the ability to operate a computer without assistance, (d) their data collection performance required intervention, and (e) they agreed to participate in the study.

Throughout the course of this study two caregivers volunteered to participate but were excluded from participation because they did not meet all inclusionary criteria. The first caregiver was excluded from the study because he was unable to operate a computer without assistance. The second caregiver was excluded from the study because his baseline performance did not require intervention. This was likely due to the fact that he was married to a caregiver who had previously participated in the study and, therefore, likely had more information about the interventions and recording ability of the system.

The recruitment for this study consisted of the primary investigator asking caregivers of clients who already agreed to collect probe data on their child's behaviors to participate in the study. The caregivers were told the data collected would be used in the development of data collection tools for other caregivers and direct-care staff. The primary investigator explained that all data collected would be protected and de-identified so that no personal information would be connected to it. When the caregiver agreed to participate, a consent form was reviewed with them and they were asked to sign it.

Data Collection Task

The data collection task completed by caregivers consisted of 1-min partial-interval data collected in 10-min sessions for a maximum daily duration of 40 min. This data collection method was selected for a number of reasons. The first reason was because partial-interval systems appear to be the most sensitive of the discontinuous measurement systems which are deemed to be the most feasible data collection system to implement (Fiske & Delmolino, 2012). Because of this feasibility and sensitivity of the system, partial-interval systems are the most represented discontinuous measurement method in behavior analytic research (Mudford, Taylor, & Martin, 2009; Sharp & Mudford, 2015). The reason 1-min interval durations were selected was because research suggests that intervals should be as short as possible while still feasible to collect (Fiske & Delmolino, 2012; Wirth, Slaven, & Taylor, 2014; Johnston & Pennypacker, 1993). Given the context of the environment wherein the data were being collected and the data collectors themselves, 1-min intervals were deemed to be the shortest interval feasible. The reason 10-min session durations were selected was because Tiger and colleagues (2013) argued that although longer total duration observations are important for maximizing accuracy when assessing problem behavior, collecting data without interruption for extended lengths of time in applied settings is not feasible. As an alternative, Tiger et al. (2013) utilized 10-min sessions that could be completed multiple times during one observation period. Finally, a total maximum duration of 40 min (i.e., 4 sessions) was set for each day. While research indicates that longer recording sessions produce less measurement error (Fiske & Delmolino, 2012), feasibility of observation duration is equally important. For the sake of the study, it was also important to set parameters for consistency in length of recording period between participants. So, 40-min daily maximums were implemented to prevent large disparities in the amount of data collected per day between participants.

Materials

The materials used in this study consisted of a paper-and-pencil data collection system as well as an EDC system. The paper-and-pencil data was collected on a single-page data sheet attached to a clipboard. The EDC system consisted of a Microsoft Excel® spreadsheet made to match the paper-and-pencil data sheet and was housed on a laptop computer. The Excel® based EDC system was created by modifying procedures described by Morris, Deochand, and Peterson (2018).

Paper data sheets completed by caregivers were collected by the primary investigator after each session and stored in a locked filing cabinet in a locked room. Data entered into the computer-based program was temporarily stored on the encrypted and password protected laptop computer during the session. After the session, the primary investigator collected the computer and transferred the data onto an encrypted and password protected external hard drive that was stored in a locked filing cabinet in a locked room.

Design and Analysis

A nonconcurrent multiple baseline across participants design was used to conduct an add-in component analysis of an EDC package with 3 participants. The add-in component analysis consisted of introducing treatment components individually and in combination. By using this model, each component was evaluated for its sufficiency while controlling for additive and multiplicative effects (Ward-Horner & Sturmey, 2010). The sequence of treatment components followed a least-to-most intrusive hierarchy that consisted of baseline- EDC, automated prompts, automated overall session feedback, automated specific interval feedback, and guided selection. All phase changes were staggered across participants.

The design used with one participant (Steve) slightly deviated in that it also included a withdrawal to baseline after each treatment condition. The withdrawal phases were utilized with Steve because he was the first caregiver to participate in the study and an ongoing evaluation of practice effects was desired by the experimenters. Although the nonconcurrent multiple baseline design allows for the assessment of practice effects across participants, evaluation of those variables within an individual's data path is not possible without additional controls like withdrawal of treatment.

Dependent Variables

Time of data entry was the dependent variable of this study and was automatically recorded as a permanent product of the EDC system. This dependent variable was selected because of the correlation between data collection timeliness and other measures of data collection fidelity described earlier, and the benefits of only using time measures (i.e., no observer required to address observer reactivity). For the purposes of this study, timely data was defined as data entered during or at the end of the interval, but no later than 5 sec after the conclusion of the interval. That is, data could be entered throughout the interval as the behavior occurred or within 5 sec of the conclusion of the interval. Although data could be entered throughout the interval, the closing of the interval (i.e., the participant indicating that no other behaviors occurred for the entire interval) could not occur until the 60 sec elapsed or the data would be deemed too early. Therefore, any interval completed before the conclusion of the interval or more than 5 sec after the interval was considered late/early.

The criteria for timely data were selected because equal interval durations are an essential component of successful interval systems (Repp, Roberts, Slack, Repp, & Berkler, 1976; Bijou, Peterson, & Ault, 1968). When data are entered at a time other than that prescribed, the interval lengths become unbalanced, negating the validity of the system. Longer intervals in a time-sampling recording system are known to produce less accurate data (Johnston & Pennypacker, 1993; Cooper et al., 2007; LeBlanc et al., 2016) indicating that an extended latency to data recording after an interval is also likely to produce inaccuracies. So, 5 sec was selected as the allotted time to enter data to accommodate a reasonable time to enter the data into the system following the end of the interval. Early data entry (i.e., closing an interval prior to the end of the recording period) is also likely to lead to inaccuracies because the entire period could not be represented.

Pre-Baseline Training and Assessment

Prior to baseline and treatment conditions, each participant was trained to collect 1-min partial interval data and use the EDC system. The training methodology employed was behavioral skills training (BST) which included instructions, modeling, rehearsal, and feedback. After being trained to collect data using the paper-and-pencil system, each participant was asked to use that method to collect data while being observed. During this time, the primary investigator observed and collected data on the participant’s timeliness and accuracy (i.e., correctly indicating whether a behavior occurred or did not occur during an interval) of data collection. Each participant was required to meet a criterion of collecting on-time and accurate data for at least 90% of intervals in a standard session using the paper-and-pencil system before being allowed to collect data using the EDC system. When the participants met the criteria using the paper-and-pencil system, the procedure was repeated with the EDC system using the same criteria. When the participants collected timely and accurate data for at least 90% of a session using the EDC system, they were permitted to begin collecting data without supervisor presence. The purpose of both probes and criterion requirements was to establish that participants had the ability to collect highly accurate and timely data so that any changes in performance during the experimental tasks could be attributed to the task contingencies and not the skill of the participant.

Experimental Conditions

All baseline and treatment conditions were conducted in the absence of a supervisor/observer during the previously agreed to recording period for each participant. To accomplish this, a laptop computer with a prepared EDC system capable of creating permanent product measures of the dependent variable was given to the caregiver before every recording period. Upon completion of the task, the participant saved the file and stored the computer until the primary investigator retrieved it. The components evaluated in this study were based on Morris (2016) and consisted of a basic EDC system, automated prompts, and two forms of automated feedback.

Baseline- EDC. The data collection task during this condition consisted of a basic EDC table, instructions, a running clock, a button to start and end the observation period, drop-down menus to indicate when behaviors occurred, and a cell for the participant to enter their initials to signify they observed the complete interval (Appendix B). All subsequent conditions were built into this baseline- EDC system. One exception to these components is in Steve's initial baseline phase. During this phase, Steve did not have a running clock present in the EDC system and instead used the running clock on his phone. Later, in Steve's return to baseline conditions, a running clock was imbedded into the EDC system.

Automated prompts. Automated prompts in the form of a reminder statement ("Enter Data Now!") were added to the baseline- EDC system directly next to the running clock (Appendix C). This statement appeared in a highlighter-yellow cell during the final recording period for each interval (e.g., between 1:00 and 1:05). Otherwise the message was not visible, and the cell was grey.

Automated overall session feedback. Automated overall session feedback was programmed to appear at the end of each recording session (i.e., 10-min period). The feedback appeared to the right of the data collection table in two forms (Appendix D). The first form of overall session feedback was narrative feedback that described the percent of intervals entered on-time and the percent of intervals that were late/early. The second form of overall session feedback consisted of the same information but was graphically displayed using a bar graph. The performance data was calculated automatically via the EDC system and consisted of averaging the performance across the 10 intervals recorded.

Automated specific interval feedback. Automated specific interval feedback was programmed to appear after each data entry (i.e., 1-min period). The feedback appeared to the right of the data collection table and appeared in the same two formats as overall session feedback. However, instead of averaging the performance of all entries during the entire session, the performance for each interval was displayed individually and immediately after it was recorded (Appendix E).

Guided selection. Prior to the guided selection phase, the primary investigator interviewed each participant to determine their individual preference of components to be included within the data collection package. During this interview, the primary investigator suggested the continued use of the most effective component previously assessed for each client and asked the client if they'd like any other component included. During this phase, each participant agreed to the continued use of automated specific interval feedback and two participants, Christine and Lindsey, requested the addition of automated prompts to the package.

Phase Change Criteria

Individual phase change decisions were made for each participant based on the trend and stability of their data and staggered across participants. For the purposes of this study, trend was defined as the overall direction of three or more data points and stability was defined as the consistency in which data points fell into the same range of values. Experimental conditions that produced ascending trends or level trends that demonstrated an improvement over baseline were continued for at least as long as criterion level performance (i.e., a trend averaging at or above 80% of intervals entered on time per session) was observed during baseline. This was implemented to assess for novelty effects of the components. Conditions that produced a level trend that was not an improvement over baseline were only continued long enough to determine a trend in performance. Finally, conditions that produced highly variable or decreasing trends in the data were also only continued long enough to determine patterns in performance. These changes were implemented sooner than ascending trends or level trends demonstrating an improvement over baseline because the purpose of this study was to identify components of an EDC system that would improve or maintain data collection timeliness. Therefore, components that did not produce an improvement or maintain criterion level timeliness were of less interest.

Methods of Data Collection

During pre-baseline training and assessment, the primary investigator observed participants collect data while collecting data on the dependent variable and instances of problem behavior demonstrated by the child of interest. This data was collected by hand using a paper-and-pencil system. During experimental conditions, no observers were present, and the dependent variable was generated automatically by the EDC system and saved as a permanent product measure.

Interobserver Agreement

The dependent variable of this study was transduced by the data collection task and stored as a permanent product. The data reported by the task was automatically calculated and reported so scoring of the results was not necessary either. Therefore, no interobserver agreement was necessary for this study because human observation and/or measurement of the dependent variable was not necessary.

Treatment Fidelity

To help prevent technological errors within the EDC system, the primary investigator checked each system before and after use. During this check, a checklist was used that reviewed each essential function of the system (Appendix F). In addition, the participants were asked if the program performed as expected following each session. The protocol for a malfunctioning system was to fix the problem before allowing a participant to use the system or, if caught after it was used, evaluate whether the data were confounded. However, this protocol was never used because no technological issues were observed or reported throughout the study.

Social Validity

Social validity was measured via a questionnaire at the conclusion of each participant's involvement in the study. Items on the questionnaire pertained to the participant's prior experience with Microsoft Excel®, malfunctions encountered while using the program, the usability of the program, preferred features, perceived influence of preferred features, perceived effectiveness of features, and willingness to continue using a similar system (Appendix G). In addition to the questionnaire, unprompted comments from the caregivers pertaining to the EDC systems were noted throughout the course of the study.

RESULTS

Prior to baseline and treatment conditions, all participants met the criteria for completing pre-baseline training and assessment. Figure 1 shows the percentage of intervals entered on-time per session for each participant across baseline and treatment conditions. Visual analysis reveals individual differences in responding by participant but similar overall outcomes. Christine's data are shown in the top panel of the figure and demonstrate an initial increase in performance during baseline followed by a steep decrease. On average, the percent of intervals entered on-time per session by Christine during baseline was 53%. When automated prompts were added to Christine's EDC system, performance, again, initially increased but decreased steeply after a few sessions. Christine's average performance with automated prompts was 40% of intervals entered on-time per session. When automated prompts were replaced with automated overall session feedback, performance initially increased yet again but decreased to a lesser extent than previously observed. Christine's average performance with automated overall session feedback was 72% of intervals entered on-time per session. When automated overall session feedback was replaced with automated specific interval feedback, Christine's performance increased and stabilized at criterion levels. Christine's average performance with automated specific interval feedback was 89% of intervals entered on-time per session. Finally, during the guided selection condition in which Christine's EDC system included automated prompts and automated specific interval feedback, Christine's performance started low, but quickly improved and stabilized above criterion levels. Her average performance with automated prompts + automated specific interval feedback was 91% of intervals entered on-time per session. At the conclusion of Christine's involvement, her average performance had increased 38% over baseline while using an EDC system with prompts + specific interval feedback.

Lindsey's data, shown in the middle panel of Figure 1, show patterns consistent with Christine's data. However, one major difference in Lindsey's performance was that her baseline performance initially maintained above criterion levels before decreasing. Although Linsey entered data on-time in an average of 78% intervals across her entire baseline phase, the average of her last four data points was 58%. These data may be explained by a novelty effect that wore off over time. When automated prompts were introduced to Lindsey's EDC system, a very small increase was observed, and data stabilized. Lindsey's average performance during the automated prompting condition was 65% of intervals entered on-time. Following the automated prompting condition, automated overall session feedback was introduced. Lindsey's performance during this phase initially improved slightly but began a descending trend after a few sessions. Lindsey's average performance with automated overall session feedback maintained at 65% of intervals entered on-time. When automated specific interval feedback was introduced to replace automated overall session feedback, Lindsey's performance was variable but showed an overall improvement. Lindsey's average performance with automated specific interval feedback was 71% of intervals entered on-time. During the guided selection phase, automated prompts and automated specific interval feedback were used in combination in Lindsey's EDC system. Lindsey's average performance during this phase was 83% of intervals entered on-time. Although 83% is only a modest improvement over Lindsey's overall average baseline of 78%, it was a 25% improvement over the average of the last four data points of baseline (i.e., 58%). In addition to improving average performance, the guided selection phase appeared to improve the stability of performance over time whereas the baseline condition decreased over time.

Steve's data, shown in the bottom panel of Figure 1, demonstrate markedly different patterns as compared to Christine and Lindsey. During Steve's baseline condition, his performance was immediately low and stable. The average number of intervals entered on-time per session during this condition was 14%. When the automated prompting phase was introduced, Steve's performance initially decreased but rapidly increased and maintained at elevated levels. His total average performance during this phase was 50% of intervals entered on-time and his average performance across his last four sessions was 78%. During Steve's brief withdrawal phase prior to introducing automated overall session feedback, his data showed a decreasing trend. When the automated overall session feedback was introduced, his performance increased and stabilized above criterion levels. His average performance during this phase was 95% of intervals entered on-time. Following the automated overall session feedback phase, a brief withdrawal was again implemented that produced a decrease in performance. Following this withdrawal, automated specific interval feedback was implemented. During the automated specific interval feedback phase, Steve entered 100% of intervals on-time across all four sessions. When this condition was withdrawn, performance slightly decreased but maintained above criterion levels. During the guided selection phase, the only feature included in Steve's basic EDC system was automated specific interval feedback. During this phase, Steve's performance maintained at an average of 97% of intervals on-time per session. Although Steve's performance improved significantly during all treatment phases, his performance was best during the automated specific interval feedback and guided selection phases. During these two phases Steve's average performance was improved by 86% and 83% over baseline, respectively.

In addition to primary evaluation of the dependent variable of this study, secondary analyses were conducted by reviewing the permanent product measures related to timeliness of data entry (see Table 1). The first variable that was evaluated using this method was the percent of intervals entered late/early that contained problem behavior. This variable was of interest because previous research suggested that the occurrence of problem behavior during the recording period could be disruptive for data collection fidelity (Mozingo et al., 2006; Morris,

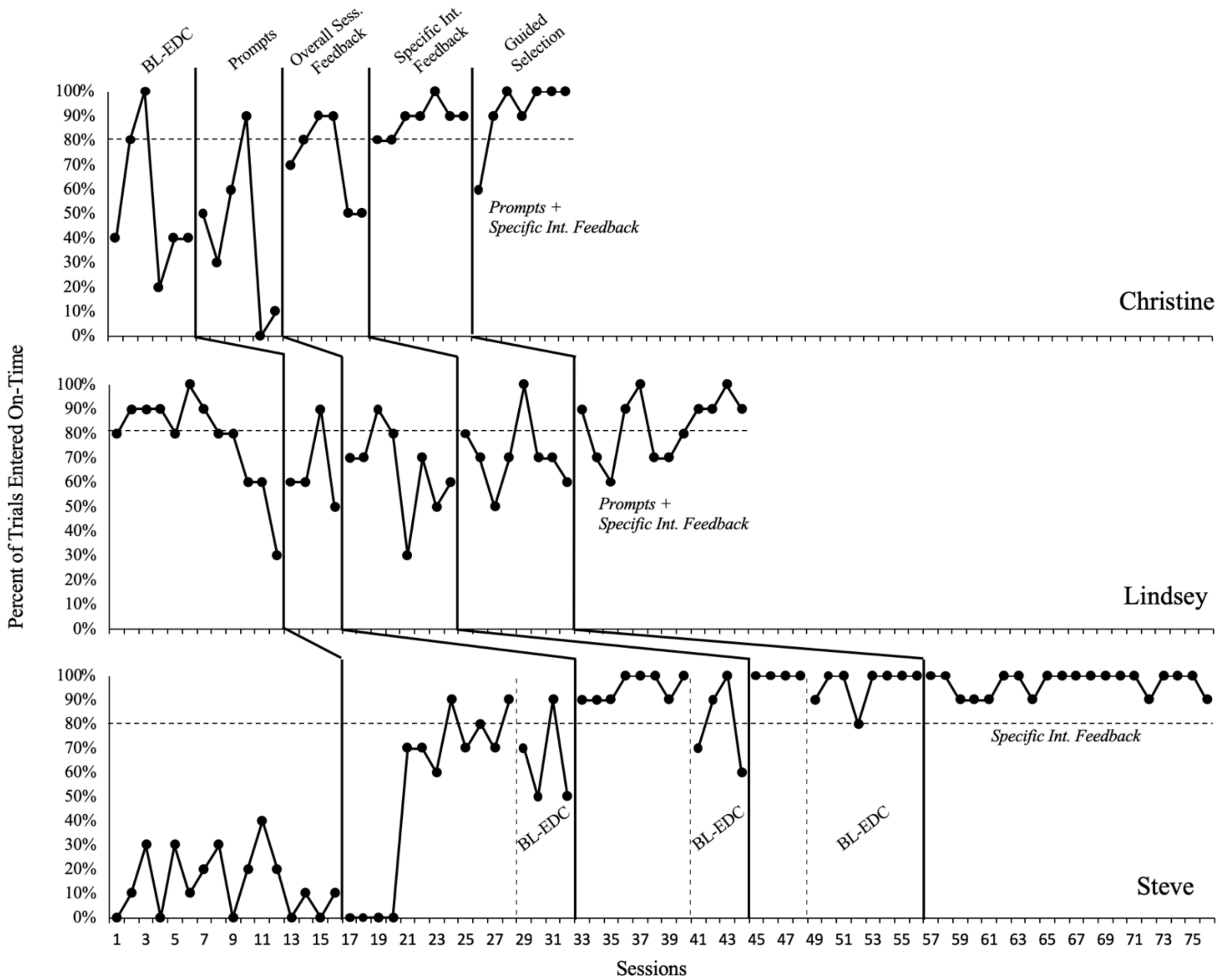


Figure 1. Percentage of On-Time Data Entries per Session. The horizontal dotted line within each panel represents the criterion level for performance.

2016). To determine the prevalence of late/early entries preceded by problem behavior, all intervals that were not entered on-time and also contained problem behavior were counted and divided by all intervals that were not entered on-time. The results of the analysis ranged by participant. Christine's data indicated that 43% of her late/early intervals included problem behavior, Lindsey's data indicated that 42% of her late/early intervals contained problem behavior, and Steve's data indicated that 17% of his late/early entries contained target behavior.

This means that 57%, 58%, and 83% of late/early data entries did not include problem behavior for each participant, respectively.

To further evaluate the effect of problem behavior on the timeliness of data entry, another analysis was conducted to calculate the percent of intervals that contained problem behavior and were entered on-time. To conduct this calculation all intervals that contained problem behavior and were also entered on-time were counted and divided by all intervals that contained target behavior. The results showed that Christine entered 36% of intervals on-time that contained problem behavior, Lindsey entered 69% of intervals on-time that contained problem behavior, and Steve entered 47% of intervals on-time that contained problem behavior.

Table 1

Summary of the Effect of Problem Behavior on Data Collection Timeliness

Participants	Of Intervals Entered Early/Late		Of Intervals that Contained Problem Behavior	
	% Containing PB	% Not Containing PB	% Early/Late	% On-Time
Christine	43%	57%	64%	36%
Lindsay	42%	58%	31%	69%
Steve	17%	83%	53%	47%

Another secondary analysis that was conducted in the course of this project evaluated changes in type and severity of late/early data entries. This variable was evaluated to assess if, in addition to increasing the total number of on-time data entries, successful treatment would decrease the severity of late/early data entries. To conduct this evaluation, the average late entry and average early entry were calculated for each participant during baseline and treatment conditions that produced stable and criterion level performance. During baseline, Christine's data indicated that 54% of her late/early data entries were late and 46% of those entries were early. When she entered late data, it was late by an average of 18 sec and when she entered early data, it was early by an average of 96 sec. Lindsey's baseline data indicated that 76% of her late/early data entries were late and 24% of those entries were early. When she entered late data, it was late by an average of 33 sec and when she entered early data, it was early by an average of 2 sec. Steve's baseline data indicated that 49% of his late/early data entries were late and 51% of those entries were early. When he entered late data, it was late by an average of 21 sec and when he entered early data, it was early by an average of 8 sec.

During Christine's treatment conditions that produced criterion level performance (i.e., automated specific interval feedback and guided selection), data indicated that 100% of her late/early data entries were late and 0% of those entries were early. When she entered late data, it was late by an average of 11 sec and she did not enter any intervals early. Lindsey's treatment conditions that produced criterion level performance (i.e., guided selection) data indicated that 74% of her late/early data entries were late and 26% of those entries were early. When she entered late data, it was late by an average of 16 sec and when she entered early data, it was early by an average of 2 sec. Steve's treatment conditions that produced criterion level performance (i.e., automated overall session feedback, automated specific interval feedback, and guided selection) data indicated that 90% of his late/early data entries were late and 10% of those entries were early. When he entered late data, it was late by an average of 12 sec and when he entered early data, it was early by an average of 6 sec.

The results of the social validity questionnaire indicated that the participants ranged in prior experience with Microsoft Excel® from never having used it before to reporting a moderate amount of experience with it prior. Christine reported that she had never used Excel® prior to the study while Lindsey and Steve both reported moderate experience levels. When asked what types of tasks they had previously engaged in on Excel®, Lindsey reported data input only and Steve reported data input, graph creation, and table creation. When all participants were asked if the task ever malfunctioned, no issues were reported related to system error or design. On a 7-point Likert scale asking if the task was easy to use (7 being very much, 4 being neutral, and 1 being not at all), Christine and Steve selected 7 while Lindsey selected 5. When asked what features were their favorite, Christine indicated automated prompts, Lindsey indicated automated prompts and automated specific interval feedback in that order, and Steve indicated automated specific interval feedback. When given a 7-point Likert scale asking if their preferred components helped produce better data (7 being very much, 4 being neutral, and 1 being not at all), all participants selected 7. When asked what features helped produce the best data, Christine reported automated prompts while Lindsey and Steve both reported automated prompts and automated specific interval feedback. Finally, when given a 7-point Likert scale asking if they'd be happy to continue using the data collection system (7 being very much, 4 being neutral, and 1 being not at all), Lindsey and Steve selected 7 and Christine selected 6.

DISCUSSION

The purpose of this study was to systematically evaluate the individual components of an EDC system similar to Morris (2016), which evaluated data collection timeliness of caregivers in a home setting in the absence of supervisor presence. Toward this end, a component analysis methodology described by Ward-Horner and Sturmey (2010) was used to evaluate automated prompts, automated overall session feedback, and automated specific interval feedback alone and in combination on data collection timeliness. Timeliness was selected as the primary dependent variable of this study because previous research had demonstrated the importance of timeliness on other data collection fidelity measures (Morris, 2016) and timeliness as a dependent variable did not require observers that could have produced observer reactivity.

The results of the study indicated that each intervention assessed improved data collection timeliness over baseline with at least some participants by varying degrees. However, automated prompts alone did not improve timeliness for all participants, nor did they produce any sustained criterion level performance for any participant. During automated overall session feedback, all participants' performance improved over baseline, but only Steve's performance met criterion levels. During automated specific interval feedback, all participants' performance again improved, but Lindsey's performance still did not maintain at criterion levels. Finally, during the guided selection condition, all participants met and maintained criterion levels of on-time data entry. For Christine and Lindsey, the guided selection condition included automated prompts + automated specific interval feedback. For Steve, the guided selection condition consisted of only automated specific interval feedback.

Lindsey's criterion level performance during a phase consisting of automated prompts + automated specific interval feedback and no other phases indicates those components may be necessary to use in combination to obtain criterion level performance from some caregivers. However, her previous best performance during the automated specific interval feedback condition suggests that it may be the more important component of the two. Christine and Steve's criterion level performance during the automated specific interval feedback condition prior to the guided selection condition also support automated specific interval feedback as the most effective isolated intervention assessed in this study.

Given the effectiveness of automated specific interval feedback during this study, further analysis of the intervention's function may be beneficial. At the onset of the study and throughout each participant's involvement, the primary investigator told participants that the data collected in the study would not affect them in anyway. However, as described by Warrilow (2017), it is possible that participants created verbal rules that could affect performance. Interestingly, during this study, Steve's performance improved the most of any participant during conditions that included feedback. This may have been due to the fact that Steve was a paid direct-care staff, and he may have created private rules regarding his employment status when data were presented revealing that his data entry was being timestamped and reviewed by the investigator. This could explain the reason that Steve's data maintained at criterion levels during the final withdrawal to baseline conditions after Steve was aware that data were being timestamped.

Another contingency that could have been affecting the participants' performance was the potential desire to impress or please the primary investigator, who was the clinician serving their children. Throughout the course of the study, all participants repeatedly sought feedback from the primary investigator about their performance. The primary investigator never provided performance feedback to any participant, but it was clear that participants were interested in the investigator's perception of their performance. If participants desired to impress the primary investigator, the feedback could have taken on stronger reinforcing or punishing properties.

Finally, it is possible that the graphic feedback being presented in the form of a red bar within the bar graph could have been aversive to the participants. Red is commonly associated with negativity/correction in western cultures, so despite no particular contingency being set around the presentation of red in the feedback, it may have been perceived as aversive by the participants. In fact, throughout the study unsolicited comments were noted and one of those comments was made pertaining to the red graphic feedback in the automated specific interval feedback phase. During that phase, one participant reported that when she saw the red graphic feedback, she "got her butt in gear." Although the red graphic feedback was displayed during both feedback conditions, the feedback did not appear until the end of the session during automated overall session feedback, so it was likely not as salient during the task.

In addition to increasing the overall timeliness of data entry, some of the components evaluated in this study decreased the severity of late/early data entry. For Christine and Steve, early data entries made up nearly 50% of data entry errors during baseline. During baseline, Steve's early entries were entered an average of 8 sec. early and Christine's early entries were entered an average of 90 sec. early. After further investigation into Christine's early data entries, it appeared that her high average was likely the result of entering data in chunks (i.e., entering data back-to-back with little to no time between) rather than following the interval recording instructions. When treatment conditions were implemented that produced criterion level performance, Christine no longer entered data early and Steve reduced his early data entries to only 10% of his data entry error with a 2 sec. average reduction in the amount early. Lindsey's early data entry had very little change throughout the study but only consisted of 24% of her errors with an average of 2 sec. early during baseline. The reduction in early data collection during this study is significant considering the improbability of early data being accurate. If data are entered late, at least data collectors have an opportunity to remember the events correctly whereas entering data before events occur is merely guessing.

The severity of late data entries was decreased for all participants from baseline to treatment conditions that produced criterion level performance. Christine's average late data entry decreased by 6 sec., Lindsey's average late data entry decreased by 17 sec., and Steve's average late data entry decreased by 8 sec. Taken together, the data showing a reduction in early and late data entries suggest that in addition to decreasing the total number of errors, the successful treatment conditions decreased the severity of the error. This was most notable in Christine's early data entry errors and Lindsey's late data entry errors.

Despite observing overall improvements in performance during the automated specific interval feedback and guided selection conditions, variability in the performance persisted. One reason this variability may have occurred was the lack of programed consequences for performance. As previously stated, it is possible that participants created their own private rules, but there were no contingencies arranged by the experimenters. Had incentives been associated with criterion level performance, it is possible that performance would have maintained more stability during evaluation. It is also possible that performance was variable due to the disruptive nature of problem behavior. The results of the secondary analysis evaluating the effect of problem behavior on timeliness suggested that problem behavior was present during less than 50% of the late/early data entries. Of the intervals that did contain problem behavior, data were entered on-time in between 36-69% of the intervals. This suggests that while problem behavior was associated with fewer than half of the data collection errors, problem behavior can cause some variability in performance when it does occur. It is also important to note that during this study, other caregivers were available to assist the child so that the participants could focus solely on data collection. As a result, problem behavior should have been less disruptive for participants of this study than for caregivers who are responsible for data collection and care simultaneously.

The results of the social validity questionnaire suggest that EDC systems could be a long-term intervention. Participants reported a range in prior experience with Microsoft Excel® suggesting that EDC systems may be suitable for anyone able to operate a computer without assistance. When asked to state which features were preferred, two of the three participants indicated automated prompts and two of the three participants indicated specific interval feedback. In follow-up questions asking which features helped produce the best data, all participants stated their preferred features, with the exception of one participant (Steve) who also included automated prompts. This is consistent with unsolicited comments throughout the course of the study by participants. For instance, after automated prompts were removed from Lindsey's EDC system following that phase, she repeatedly asked that they be replaced and stated, "the prompts are a must have." For Lindsey, when automated prompts were returned and combined with automated specific interval feedback, performance improved and maintained at criterion levels. Taken together with the outcome of the study, it appears that participants tended to prefer components that produced better data collection timeliness.

This study included many limitations. One major limitation is that although less than 50% of the intervals that were entered late/early contained target behavior, there was no measure of severity of behavior during the intervals that did. Severity may explain the large variation (i.e., 36-69%) of on-time data entry when problem behavior did occur. Future studies evaluating data collection timeliness should further evaluate the relation between severity of problem behavior and timeliness. Another limitation of the study is that outside influences on the behavior were not controlled. For instance, while the caregivers were asked to minimize distractions and focus on data collection during the recording period, they might receive a phone call or may be distracted by another caregiver or child asking them a question. Future studies could attempt to control for outside influences by evaluating data collection in controlled environments. Perhaps the biggest limitation of this study was that other measures of data collection fidelity beyond timeliness were not measured. The decision to not collect these measures during this study was made to prevent observer reactivity, but other studies could utilize video recording or other unobtrusive methods to collect both timeliness and other fidelity measures.

Future researchers interested in data collection fidelity should continue to evaluate the relation between timeliness and accuracy with and without EDC systems. Procedures for building EDC systems to fit individual contexts are described by Morris, Deochand, and Peterson (2018) and can be used by researchers and practitioners to build free programs that allow for timestamping and automated features like prompts and feedback. By using this and similar methods, researchers can build EDC systems to evaluate data collection across settings, systems, populations, and client behaviors. Additionally, future studies could attempt to identify methods for automatically recording aspects of behavior via wearable technologies or other emerging technology and reduce the need for human data collectors.

In conclusion, this study demonstrated the effectiveness of an EDC system utilizing automated specific interval feedback in combination with automated prompts on improving data collection timeliness. This is significant because timeliness of data collection has been shown to correlate with other measures of data collection fidelity that are essential to producing accurate data. As previously stated, if the data used by behavior analysts to inform their treatment decisions are not accurate, inadvertent errors in decision making are likely. Therefore, given the outcome of this study, researchers or practitioners using caregivers to collect data should consider using EDC systems with automated specific interval feedback + automated prompts to help ensure quality data. Although this was a contrived study conducted with caregivers in home settings, the findings may also be useful to researchers or practitioners working in other settings that require data collection from individuals who are not behavior analysts. For instance, behavior technicians collecting data in an autism center may also benefit from automated specific interval feedback + automated prompts when collecting data.

REFERENCES

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. http://dx.doi.org.libproxy.library.wmich.edu/10.1300/J075v21n01_02

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, *1*, 91–97. http://dx.doi.org.libproxy.library.wmich.edu/10.1901/jaba.1968.1-91

Behavior Analyst Certification Board. (2014). Professional and ethical compliance code for behavior analysts. Retrieved from http://bacb.com/wp-content/uploads/2016/03/160321-compliancecode-english.pdf

Bijou, S. W., Peterson, R. F., & Ault, M. H. (1968). A method to integrate descriptive and experimental field studies at the level of data and empirical concepts. *Journal of Applied Behavior Analysis*, *1*(2), 175-191.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied Behavior Analysis* (2nd ed). Columbus, Ohio: Merrill.

Dixon, M. R. (2003). Creating a portable data-collection system with Microsoft embedded visual tools for the pocket PC. *Journal of Applied Behavior Analysis, 36*(2), 271–284. http://dx.doi.org.libproxy.library.wmich.edu/10.1901/jaba.2003.36-271

Fiske, K., & Delmolino, L. (2012). Use of discontinuous methods of data collection in behavioral intervention: Guidelines for practitioners. *Behavior Analysis in Practice*, *5*(2), 77-81.

Goomas, D. T. (2012a). Immediate feedback on accuracy and performance: the effects of wireless technology on food safety tracking at a distribution center. *Journal of Organizational Behavior Management, 32*(4), 320-328. http://dx.doi.org.libproxy.library.wmich.edu/10.1080/01608061.2012.729391

Goomas, D. T. (2012b). The impact of wireless technology on order selection audits at an auto parts distribution center. *Journal of Organizational Behavior Management*, *32*(2), 131-139. http://dx.doi.org.libproxy.library.wmich.edu/10.1080/01608061.2012.676505

Hickman, J. S., & Geller, E. S. (2005). Self-management to increase safe driving among short-haul truck drivers. *Journal of Organizational Behavior Management*, *23*(4), 1-20.

Johnston, J. M. & Pennypacker, H. S. (1993). *Strategies and tactics of behavioral research*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

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

LeBlanc, L. A., Raetz, P. B., Sellers, T. P., & Carr, J. E. (2016). A proposed model for selecting measurement procedures for the assessment and treatment of problem behavior. *Behavior Analysis in Practice*, *9*(1), 77-83.

Madsen, E. K., Peck, J. A., & Valdovinos, M. G. (2015). A review of research on direct-care staff data collection regarding the severity and function of challenging behavior in individuals with intellectual and developmental disabilities. *Journal of Intellectual Disabilities*. DOI: 10.1177/1744629515612328

Moon, K., & Oah, S. (2013). A comparison of the effects of feedback and prompts on safe sitting posture: Utilizing an automated observation and feedback system. *Journal of Organizational Behavior Management*, *33*(2), 152-162.

Morris, C. (2016). *The effects of electronic data collection, immediate graphic feedback, and automated scheduled prompts on data collection adherence* (Master's thesis). Retrieved from http://scholarworks.wmich.edu/masters_theses/728.

Morris, C. A., Deochand, N., & Peterson, S. M. (2018). Using Microsoft Excel® to Build a Customized Partial-Interval Data Collection System. *Behavior Analysis in Practice*, 1-13.

Mozingo, D. B., Smith, T., Riordan, M. R., Reiss, M. L., & Bailey, J. S. (2006). Enhancing frequency recording by developmental disabilities treatment staff. *Journal of applied behavior analysis*, *39*(2), 253-256.

Mudford, O. C., Taylor, S. A., & Martin, N. T. (2009). Continuous recording and interobserver agreement algorithms reported in the Journal of Applied Behavior Analysis (1995–2005). *Journal of Applied Behavior Analysis*, *42*(1), 165-169.

Olson, R., & Austin, J. (2001). Behavior-based safety and working alone: The effects of a self-monitoring package on the safe performance of bus operators. *Journal of Organizational Behavior Management*, *21*(3), 5-43.

Olson, R., Hess, J., Parker, K., Thompson, S., Rameshbabu, A., Luther Rhoten, K., & Marino, M. (2018). From research-to-practice: an adaptation and dissemination of the compass program for home care workers. *International Journal of Environmental Research and Public Health*, *15*(12), 27-77.

Realon, R. E., Favell, J. E., & McGimsey, J. F. (1992). Computer prompts to improve social interactions and data collection. *Mental retardation*, *30*(1), 23.

Reis, M. H., Wine, B., & Brutzman, B. (2013). Enhancing the accuracy of low-frequency behavior data collection by direct-care staff. *Behavioral Interventions*, *28*(4), 344-352.

Repp, A. C., Roberts, D. M., Slack, D. J., Repp, C. F., & Berkler, M. S. (1976). A comparison of frequency, interval, and time-sampling methods of data collection. *Journal of Applied Behavior Analysis*, *9*(4), 501-508.

Sharp, R. A., & Mudford, O. C. (2015). Distribution of reported durations of behavior in applied behavioral research. *Behavioral Interventions*, *30*(4), 352-363.

Sidman, M. (1960/1988). Tactics of scientific research: Evaluating experimental data in psychology. New York: Basic Books/ Boston, Massachusetts. Authors Cooperative (reprinted).

Sleeper, J. D., LeBlanc, L. A., Mueller, J., Valentino, A. L., Fazzio, D., & Raetz, P. B. (2017). The Effects of Electronic Data Collection on the Percentage of Current Clinician Graphs and Organizational Return on Investment. *Journal of Organizational Behavior Management*, 1-13. DOI: 10.1080/01608061.2016.1267065

Taber-Doughty, T., & Jasper, A. D. (2012). Does latency in recording data make a difference? Confirming the accuracy of teachers' data. *Focus on Autism and Other Developmental Disabilities, 27*(3), 168-176.

Tapp, J., Ticha, R., Kryzer, E., Gustafson, M., Gunnar, M. R., & Symons, F. J. (2006). Comparing observational software with paper and pencil for time-sampled data: A field test of interval manager (INTMAN). *Behavior Research Methods*. 38 (1), 165-169.

Tarbox, J., Wilke, A. E., Findel-Pyles, R. S., Bergstrom, R. M., & Granpeesheh, D. (2010). A comparison of electronic to traditional pen-and-paper data collection in discrete trial training for children with autism. *Research in Autism Spectrum Disorders*, *4*(1), 65-75. http://dx.doi.org/10.1016/j.rasd.2009.07.008

Tiger, J. H., Miller, S. J., Mevers, J. L., Mintz, J. C., Scheithauer, M. C., & Alvarez, J. (2013). On the representativeness of behavior observation samples in classrooms. *Journal of applied behavior analysis*, *46*(2), 424-435.

Van der Mars, H. (1988). The effects of audio-cueing on selected teaching behaviors of an experienced elementary physical education specialist. *Journal of Teaching in Physical Education*, *8*(1), 64-72.

Van Houten, R., & Sullivan, K. (1975). Effects of an audio cueing system on the rate of teacher praise. *Journal of Applied Behavior Analysis*, *8*, 197-201. http://dx.doi.org.libproxy.library.wmich.edu/10.1901/jaba.1975.8-197

Ward-Horner, J., & Sturmey, P. (2010). Component analyses using single-subject experimental designs: A review. *Journal of applied behavior analysis*, *43*(4), 685-704.

Warrilow, G. D. (2017). *The Effects of Feedback Modality on Performance* (Master's thesis, Western Michigan University, Kalamazoo, MI).

Whiting, S. W., & Dixon, M. R. (2012). Creating an I-phone application for collecting continuous ABC data. *Journal of Applied Behavior Analysis,* 45(3), 643-56.

Wirth, O., Slaven, J., & Taylor, M. A. (2014). Interval sampling methods and measurement error: A computer simulation. *Journal of Applied Behavior Analysis*, *47*(1), 83-100.

Appendix A

HSIRB Approval

WESTERN MICHIGAN UNIVERSITY

Institutional Review Board
FWA00007042
IRB00000254

Date: November 8, 2018

To: Stephanie Peterson, Principal Investigator
Cody Morris, Student Investigator for dissertation
Kelsey Webster, Student Investigator

From: Amy Naugle, Ph.D., Chair Amy Naugle

Re: IRB Project Number 18-11-06

This letter will serve as confirmation that your research project titled "A Component Analysis of an Electronic Data Collection Package" has been **approved** under the **exempt** category of review by the Western Michigan University Institutional Review Board (IRB). The conditions and duration of this approval are specified in the policies of Western Michigan University. You may now begin to implement the research as described in the application.

Please note: This research may **only** be conducted exactly in the form it was approved. You must seek specific board approval for any changes to this project (e.g., ***you must request a post-approval change to enroll subjects beyond the number stated in your application under "Number of subjects you want to complete the study"***). Failure to obtain approval for changes will result in a protocol deviation. In addition, if there are any unanticipated adverse reactions or unanticipated events associated with the conduct of this research, you should immediately suspend the project and contact the Chair of the IRB for consultation.

Reapproval of the project is required if it extends beyond the termination date stated below.

The Board wishes you success in the pursuit of your research goals.

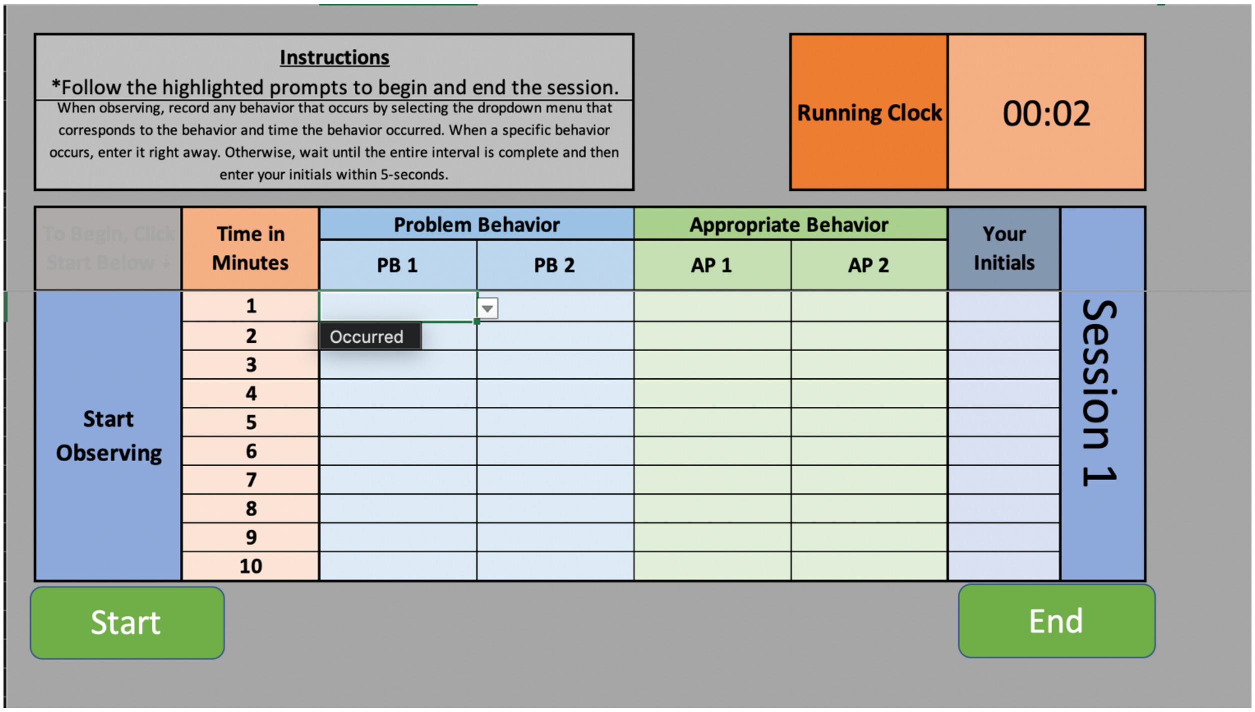
Approval Termination: **November 7, 2019**

Office of the Vice President for Research
Research Compliance Office
1903 W. Michigan Ave., Kalamazoo, MI 49008-5456
PHONE: (269) 387-8293 FAX: (269) 387-8276
WEBSITE: wmich.edu/research/compliance/hsirb

CAMPUS SITE: Room 251 W. Walwood Hall

Appendix B

Baseline- EDC System



Instructions

*Follow the highlighted prompts to begin and end the session.

When observing, record any behavior that occurs by selecting the dropdown menu that corresponds to the behavior and time the behavior occurred. When a specific behavior occurs, enter it right away. Otherwise, wait until the entire interval is complete and then enter your initials within 5-seconds.

Running Clock	00:02

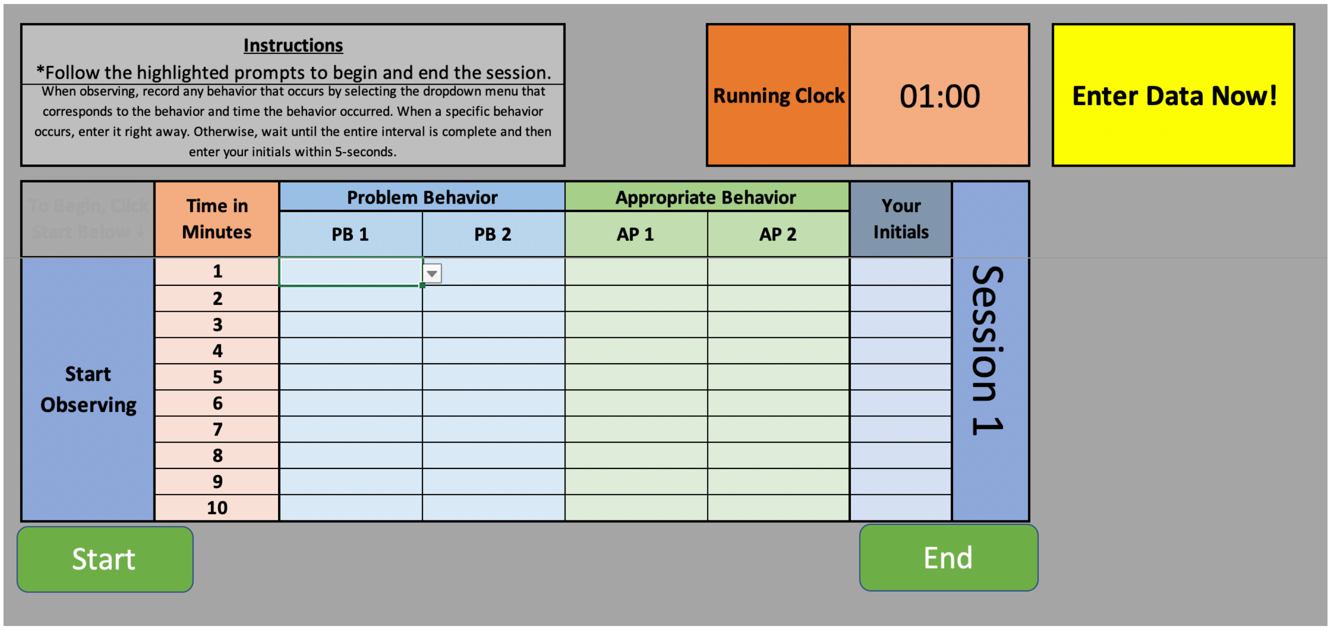
To Begin, Click Start Below ↓	Time in Minutes	Problem Behavior		Appropriate Behavior		Your Initials	
		PB 1	PB 2	AP 1	AP 2		
Start Observing	1						Session 1
	2	Occurred					
	3						
	4						
	5						
	6						
	7						
	8						
	9						
	10						

Start

End

Appendix C

Automated Prompts


Instructions
*Follow the highlighted prompts to begin and end the session.
When observing, record any behavior that occurs by selecting the dropdown menu that corresponds to the behavior and time the behavior occurred. When a specific behavior occurs, enter it right away. Otherwise, wait until the entire interval is complete and then enter your initials within 5-seconds.
Running Clock
01:00
Enter Data Now!
Time in Minutes
Problem Behavior
PB 1
PB 2
Appropriate Behavior
AP 1
AP 2
Your Initials
Start Observing
1
2
3
4
5
6
7
8
9
10
Session 1
Start
End

Appendix D

Automated Overall Session Feedback

Instructions

*Follow the highlighted prompts to begin and end the session.

When observing, record any behavior that occurs by selecting the dropdown menu that corresponds to the behavior and time the behavior occurred. When a specific behavior occurs, enter it right away. Otherwise, wait until the entire interval is complete and then enter your initials within 5-seconds.

Running Clock	01:00

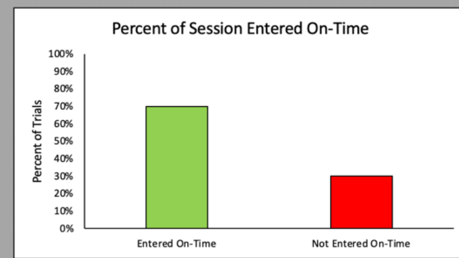
To Begin, Click Start Below	Time in Minutes	Problem Behavior		Appropriate Behavior		Your Initials	
		PB 1	PB 2	AP 1	AP 2		
Start Observing	1	Did Not Occur	Did Not Occur	Did Not Occur	Did Not Occur	CM	Session 1
	2	Did Not Occur	Occurred	Did Not Occur	Did Not Occur	CM	
	3	Did Not Occur	Did Not Occur	Did Not Occur	Did Not Occur	CM	
	4	Occurred	Did Not Occur	Did Not Occur	Did Not Occur	CM	
	5	Did Not Occur	Did Not Occur	Did Not Occur	Occurred	CM	
	6	Did Not Occur	Did Not Occur	Occurred	Did Not Occur	CM	
	7	Did Not Occur	Occurred	Did Not Occur	Did Not Occur	CM	
	8	Did Not Occur	Occurred	Did Not Occur	Did Not Occur	CM	
	9	Occurred	Did Not Occur	Did Not Occur	Did Not Occur	CM	
	10	Did Not Occur	Did Not Occur	Did Not Occur	Did Not Occur	CM	

Start

To Finish, Click "End" -->

End

Percent of Trials Entered On-Time	Percent of Trials Not Entered On-Time
70%	30%



Appendix E

Automated Specific Interval Feedback

Instructions

*Follow the highlighted prompts to begin and end the session.

When observing, record any behavior that occurs by selecting the dropdown menu that corresponds to the behavior and time the behavior occurred. When a specific behavior occurs, enter it right away. Otherwise, wait until the entire interval is complete and then enter your initials within 5-seconds.

Running Clock	01:00

To Begin, Click Start Below ↓	Time in Minutes	Problem Behavior		Appropriate Behavior		Your Initials	
		PB 1	PB 2	AP 1	AP 2		
Start Observing	1	Did Not Occur	Did Not Occur	Did Not Occur	Did Not Occur	CM	Session 1
	2	Did Not Occur	Occurred	Did Not Occur	Did Not Occur	CM	
	3	Did Not Occur	Did Not Occur	Did Not Occur	Did Not Occur	CM	
	4	Occurred	Did Not Occur	Did Not Occur	Did Not Occur	CM	
	5	Did Not Occur	Did Not Occur	Did Not Occur	Occurred	CM	
	6	Did Not Occur	Did Not Occur	Occurred	Did Not Occur	CM	
	7	Did Not Occur	Occurred	Did Not Occur	Did Not Occur	CM	
	8	Did Not Occur	Occurred	Did Not Occur	Did Not Occur	CM	
	9	Occurred	Did Not Occur	Did Not Occur	Did Not Occur	CM	
	10	Did Not Occur	Did Not Occur	Did Not Occur	Did Not Occur	CM	

Start

To Finish, Click "End" -->

End

On-Time Vs Not On-Time	
1	On-Time
2	On-Time
3	On-Time
4	Not On-Time
5	On-Time
6	On-Time
7	Not On-Time
8	On-Time
9	On-Time
10	Not On-Time

Appendix F

Treatment Fidelity Checklist

Treatment Integrity Checklist

Prior to Session

- Are all formulas enabled?
- Are the appropriate features present?
- Is the task locked?
- Is the data sheet hidden?
- All sessions ready?
- Does the sheet open with Session 1?
- Does the macro to begin the session work?

After Session

- Did any formula malfunction?
- Was anything tampered with?
- Is there evidence that features worked appropriately?

Appendix G

Social Validity Questionnaire

Social Validity Questionnaire

Date: ______________ Participant: _________

How much prior experience do you have using Microsoft Excel?

1	2	3	4	5	6	7
None at All			Neutral			Very Much

Did the task ever malfunction while in use?

Was the program easy to use?

1	2	3	4	5	6	7
Not at All			Neutral			Very Much

What was your favorite feature?

Prompts Overall Feedback Specific Trial Feedback None

Do you think the features helped produce better data?

1	2	3	4	5	6	7
Not at All			Neutral			Very Much

What features helped the most?

Prompts Overall Feedback Specific Trial Feedback None

If you had to continue collecting data, would happy to use this system?

1	2	3	4	5	6	7
Not at All			Neutral			Very Much