



Toward an Understanding of Data Collection Integrity

Cody Morris¹  · Alissa A. Conway²  · Jessica L. Becraft³  · Biancé J. Ferrucci¹ 

Accepted: 20 February 2022 / Published online: 14 March 2022
© Association for Behavior Analysis International 2022

Abstract

Data collection is an integral part of the practice of behavior analysis because behavior analysts rely on data to inform their clinical decisions. Data collection integrity (DCI) is the degree to which data are collected as planned, and issues with DCI can lead to misinformed clinical decisions. The current study aims to add to the limited research on DCI by evaluating risk factors and interventions that target DCI. An online survey, conducted through Qualtrics™, was completed by a combined total of 232 Board-Certified Behavior Analysts (BCBAs) and Board-Certified Behavior Analysts-Doctoral (BCBA-Ds). Participants answered questions about their demographics, their data collectors, their concerns about data collection, the systems they use to collect data, the training they provide data collectors, and the strategies they use to address data-collection issues. Results indicated that many risk factors related to DCI issues might be prevalent in behavior analytic practice. Recommendations on how to address DCI issues are provided.

Keywords Data collection · Data collection integrity · Treatment integrity · Data accuracy

Introduction

Behavioral data (referred to as data throughout this paper) is the quantification of the measurements of targeted variables relevant to a behavioral analysis (e.g., specific behaviors and stimulus changes). These quantified measurements provide summary information about observed events that enable analyses of those events. The utility of data in the practice of behavior analysis is providing a precise and objective basis for clinical decisions (Cooper et al., 2020, p. 74; Johnston et al., 2020, p. 132). Therefore, data are highly valued within applied behavior analysis and typically form the primary basis for clinical decisions by practicing behavior analysts (Cooper et al., 2020, p. 75; LeBlanc et al., 2016). In fact, using data is not only considered to be the best practice for behavior analysts (Slocum et al., 2014) but it is also required by the Ethics Code for Behavior Analysts (the Code; 2020, p. 12).

To obtain data that can be used to inform clinical decisions, measurement of targeted variables must take place. There are two basic options for measurement within behavior analysis – automatically recorded data that is independently produced by a measurement system or human recorded data that requires observation/input from an observer (Johnston et al., 2020, p. 133). Automatically recorded data typically require the behavior to interact directly with the recording equipment (e.g., pressing a key on a computer or selecting something using a mouse on a computer), whereas human observation requires an observer to witness and record a behavior engage in behavior (e.g., observing the number of times that a client engages in aggression or timing a student complete a transition). Although both measurement options have utility within behavior analysis, practitioners typically rely on human observation due to the free-operant nature of many frequently targeted behaviors (i.e., aggression, property destruction, elopement, self-injurious behavior) in applied settings.

The accuracy and reliability of data obtained through measurement are vital to its trustworthiness (Cooper et al., 2020, p. 102). Any error in measurement could lead to inaccurate data, which in turn could lead to misinformed clinical decisions that affect clients. Thus, careful consideration of the inherent risks involved with measurement is warranted. While the primary inherent risk of automatically recorded

✉ Cody Morris
cody.morris@salve.edu

¹ Department of Psychology, Salve Regina University, 100 Ochre Point Ave, Newport, RI 02840, USA

² NeurAbilities Healthcare, King of Prussia, PA, USA

³ Kennedy Krieger Institute and Johns Hopkins University School of Medicine, Baltimore, MD, USA

data is the accuracy of the recording system itself, the inherent risks of human recorded data include both the accuracy of the recording system *and* the accuracy of the observer (Cooper et al., 2020, p. 106). To address the issues related to the accuracy of the recording system, appropriate data-collection systems must be selected that capture representative dimensions of the targeted behavior. For example, when targeting a behavior such as screaming that occurs for extended periods of time, a system that records duration rather than frequency would help produce more representative and accurate data.¹ Addressing issues related to the accuracy of the observer is much more complex than addressing issues related to the accuracy of the recording system because it involves many more variables, and human error in measurement presents the biggest threat to the accuracy and reliability of data (Cooper et al., 2020, p. 106).

The prevalence of observer/human measurement error and the accuracy of recorded data are associated with data collection integrity (DCI). DCI refers to the degree to which data are recorded as planned in the same way that treatment integrity refers to the degree to which interventions are delivered as planned (Gresham, 1989). To illustrate the similarity and difference between DCI and treatment integrity, consider the following example: A paraprofessional who is asked to implement a differential reinforcement of other behavior (DRO; Zane & Davis, 2013) procedure while collecting data on the intervention would be expected to follow specific steps when implementing the procedure as well as when collecting the data. If the paraprofessional failed to implement the DRO correctly (e.g., provided reinforcement following a targeted behavior), a treatment integrity issue would be present. Similarly, if the paraprofessional failed to collect the data correctly (e.g., recorded an instance of a targeted behavior when none occurred), a DCI issue would be present. Where treatment integrity issues are harmful to the effectiveness of the intervention, DCI issues are harmful to the evaluation of the effectiveness of the intervention. Although DCI and treatment integrity are separate issues, overlap between the two is likely. Going back to the DRO example, if the paraprofessional delivered reinforcement after a target behavior, it seems likely that they would most likely fail to record that instance of target behavior as well.

DCI, like treatment integrity, is a multifaceted issue that is likely especially problematic in applied settings (Morris & Peterson, 2020). However, compared to treatment integrity, research focused on DCI is limited. Thus, little is known about the prevalence and risk factors associated with DCI. Cooper et al. (2020, p. 106) provided some consideration for DCI when they listed poorly designed measurement systems,

inadequate observer training, and unintended influences on observers as major contributors to human measurement error (i.e., DCI concerns). However, no references to specific research were provided when describing the contributors to DCI concerns other than two citations within the inadequate observer training section. Although the risk factors related to DCI that Cooper et al. describe are logical and help direct clinicians to potential issues, more research is needed to validate (or potentially invalidate) their suggestions.

In addition to the information provided by Cooper et al. (2020), a small number of research studies have targeted DCI, albeit sometimes using different terminology such as *data collection accuracy*. One group of research studies focused on supervision strategies (e.g., Mazingo et al., 2006; Reis et al., 2013), and another group has focused on technology-based strategies such as electronic data collection systems (e.g., Morris & Peterson, 2020; Tapp et al., 2006; Tarbox et al., 2010). Although many of these studies demonstrated that DCI can be problematic when not targeted for intervention, a complete analysis of potential risk factors and interventions related to DCI is absent.

The purpose of this study was to obtain information related to risk factors and interventions targeting DCI by surveying behavior analysts about their data collection practices and concerns. Specifically, this study looked at demographic information, information about primary data collectors, data collector responsibilities, concerns related to DCI, data collection systems, data collector training, and strategies to address DCI issues.

Method

Participants

Board-Certified Behavior Analysts (BCBAs) and Board-Certified Behavior Analysts-Doctoral (BCBA-Ds) who collect or supervise the collection of data focused on problem behavior in applied settings were invited to participate in this study by completing a survey. A total of 379 people began the survey, but 147 of them were excluded from the study because they either did not complete the survey or they did not meet the inclusionary criteria (i.e., have a BCBA or BCBA-D credential and collect or supervise the collection of data focused on problem behavior in applied settings). The total number of participants who completed the survey in its entirety was 232. Of the participants, 210 (90.5%) were BCBAs, and 22 (9.5%) were BCBA-Ds.

Materials

A 68-item survey was created and hosted within Qualtrics™, an online surveying platform. The survey consisted of two

¹ See LeBlanc et al. (2016) for decision tree outlining the steps to selecting appropriate data collection methods.

segments with separate focuses. One segment of the survey focused on information relevant to DCI specifically, and the other segment was part of another study focused on information relevant to parameters of measurement more broadly. Only the questions and responses related to DCI content are reported here. The DCI survey consisted of 37 questions that included multiple question formats (i.e., yes/no questions, Likert Scale ratings, and open-ended questions). The first section of the survey focused on demographic information. The subsequent sections of the survey focused on information about primary data collectors, data collector responsibilities, concerns related to DCI, data collection systems, data collector training, and strategies to address DCI issues.

Procedure

BCBAs and BCBA-Ds were recruited to participate in this study via emails sent through the Teaching Behavior Analysis (TBA) Listserv and the Behavior Analyst Certification Board’s (BACB) mass email service. Initial recruitment emails were sent via the TBA Listserv and BACB’s mass email service within a couple of weeks of one another. Four months later, a second email was sent by the BACB’s mass email service that consisted of the same information sent the first time.

When potential participants opened the link to the Qualtrics™ from the recruitment email, they were directed to a webpage that provided information about the survey as well as information about the process of consenting to participate in the survey. After participants consented to participate by continuing with the survey after the consent page, they were directed to answer the survey questions. Although the DCI segment of the survey included 37 possible questions, many of the questions were conditional based on the participants’ answers to previous questions. For example, questions about the type of electronic data collection system only appeared if the participant indicated that they used electronic data collection systems. Therefore, the specific number of questions varied by participant. The estimated time to complete the survey was 15–20 min.

Results

The results of the survey are organized by category of questions. Each category is summarized below with detailed information about each question and answer shown in the corresponding tables.

Demographic Characteristics

The participants reported information about the primary setting that they provide clinical services, the populations

they serve, the age of the populations they serve, how their clinical services are funded, and how long they have been providing services. See Table 1 for the detailed results for this section.

The results of the questions targeting demographics indicated that a majority of the participants provide home-based (33%), school-based (25%), or out-patient clinic-based (24%) services primarily to clients with Autism Spectrum Disorder (ASD; 93%) and Developmental Disabilities (DDs; 60%). The age range of the clients served by participants varied but primarily consisted of minors as characterized as early intervention (59%), youth (75%), and adolescents (65%). The reported funding for clinical services mostly came from insurance (61%), schools (35%), and private funds (29%). The average length of practice experience across the participants was 6.22 years.

Although the total number of participants that completed the survey was somewhat low (232), the demographic characteristics of the participants seem to reflect the typical characteristics of the field (BACB, 2021; Jones et al., 2020). For example, the results of the survey indicated that most of

Table 1 Participant demographic information

Question Category	N	Percentage
Primary Setting		
Home-Based	77	33%
School	59	25%
Out-Patient Clinic	55	24%
Residential Program	20	9%
Other	21	9%
Population(s) Served		
Autism Spectrum Disorder	216	93%
Developmental Disabilities	139	60%
Mental Illness	54	23%
Other	22	9%
Age of Population(s) Served		
Early Intervention	137	59%
Youth	175	75%
Adolescents	151	65%
Adults	65	28%
Geriatrics	13	6%
How are your clinical services funded?		
Federal Grants	14	6%
State Grants	32	14%
Research Grants	3	1%
Insurance	142	61%
Private Funds	67	29%
Schools	82	35%
Other	28	12%
Question Category	Average	
Length of Experience Practicing	6.22 Years	

the participants provided clinical services to minors with ASD and/or DDs. Thus, despite the low total of participants, the sample size obtained appears to be representative of the field. Furthermore, the average of 6.22 years of experience across the participants suggests that they were experienced practitioners with the clinical perspective to provide an accurate depiction of DCI.

Primary Data Collector Information

The participants answered five questions about the primary data collectors for their data collection system. See Table 2 for the detailed results for this section.

The first question asked the participants to specify who the primary data collector is for their clients. The participants indicated that most data collectors are either Registered Behavior Technicians (RBTs; 47%) or other direct-care staff (40%). Combined, RBTs and direct-care staff made up 87% of the primary data collectors, while the remaining 13% consisted of teachers, parents, BCBAs, and others. Therefore, one important consideration for interpreting the rest of the results and/or studying DCI more broadly is that direct-care staff (RBTs included) are likely the most represented and relevant group.

The next four questions about the data collectors focused on the number of data collectors typically involved in a client's data collection responsibilities. The participants indicated that they utilize an average of 2.0 data collectors per client per day and an average of 5.4 data collectors per client per 3-month span of time. When asked what percent of data collection involves a secondary observer, the participants indicated an average of 14.1%. However, the median and the mode for the percent of data collection with a secondary observer was only 10%. Taken together, these data suggest that multiple independent data collectors are tasked with serving as the primary data collector for a single client with minimal support provided via a secondary data collector.

Data Collector Responsibilities

The participants answered three questions about the data collectors' responsibilities. See Table 3 for the detailed results for this section.

The participants indicated that data collectors have responsibilities beyond data collection that include implementing treatment for the client whose data is being collected (97%), caregiver tasks (35%), implementing treatment of other clients (27%), collecting data for other clients (25%), and teaching a classroom of students (17%). In fact, less than 1% of the participants indicated that data collectors had no other responsibilities. When asked how many other clients a data collector provides treatment and collects data for, the participants indicated an average of 2.97 and 2.63, respectively. Thus, it is abundantly clear that data collection is not the sole responsibility of data collectors, which creates competing contingencies that could negatively affect DCI. For example, a direct care staff who is tasked with providing services and collecting data for two clients might not be physically capable of executing all of their responsibilities simultaneously. Therefore, the direct care staff might need to choose between implementing treatment or collecting data. If they fail to implement the treatment, the client's treatment progress could be harmed. If they fail to collect data, the analysis of the client's treatment progress could be harmed.

Concerns about Data Collection Integrity

The participants answered six questions about the concerns about data collection integrity. See Table 4 for the detailed results for this section.

When asked if they ever doubted the accuracy of the reported data, 76% of the participants selected "yes." Similarly, 72% of the participants indicated that they had doubted the reliability of reported data, and 85% of the participants

Table 2 Primary data collector information

Question Category	<i>n</i>	Percentage
Who is the primary data collector for your clients?		
Registered Behavior Technicians (RBTs)	109	47%
Direct-Care Staff	93	40%
Teachers	14	6%
Parents	5	2%
BCBAs	2	1%
Others	9	4%
Question Category		Average
Throughout a typical day, about how many different data recorders serve as the primary data collector for a single client?		2.0
Over a typical 3-month span of time, approximately how many different staff collect data for one client?		5.4
Approximately what percentage of data collection includes a secondary observer collecting data?		14.1%

Table 3 Data collector responsibilities

Question Category	<i>n</i>	Percentage
What other responsibilities do the primary data collectors have while collecting data?		
Implement treatment for the client whose data is being collected	224	97%
Caregiving tasks	81	35%
Implement treatment for other clients	62	27%
Collect data on other clients	57	25%
Teaching a classroom of students	39	17%
None	1	<1%
Question Category		Average
While collecting data for one client, about how many other clients does the staff typically provide treatment for?		2.97
While collecting data for one client, about how many other clients does the staff typically collect data for?		2.63

indicated that data collectors have failed to complete data collection. Therefore, at least three concerns seem prominent amongst the participants of this survey. The first concern is the accuracy of the data being reported. Accuracy in this context refers to the degree to which an observed or reported value matches the true value (Cooper et al., 2020, p. 102). The second concern is the reliability of the data being reported. Reliability in this context is similar to accuracy but distinctly refers to the consistency of reported measurements (Cooper et al., 2020, p. 102). For example, a data collector who consistently records hugging as a form of physical aggression would achieve high reliability of their data, but the data would not be accurate. Both accuracy and reliability

are important in obtaining trustworthy data. Doubting the accuracy of data means that the participant believes that the data they review may not reflect the actual events that transpired, and doubting the reliability of the data means that they believe that the data are not being recorded consistently. The third prominent concern identified in this section is the failure to complete data collection. Meaning, participants reported that their staff do not always complete their assigned data collection. Thus, in addition to considering issues with the accuracy and reliability of data collection, it is also important to focus on ensuring the task is completed.

Three follow-up questions in this section provided more context about DCI concerns. The first question asked if data

Table 4 Concerns about data collection integrity

Question Category	<i>n</i>	Percentage
Do you ever doubt the accuracy of any of the reported data?		
Yes	176	76%
No	56	24%
Do you ever doubt the reliability of any of the reported data?		
Yes	167	72%
No	65	28%
Do data collectors ever fail to complete data collection?		
Yes	197	85%
No	35	15%
Do data collectors ever fill in data before the events occur (early completion)?		
Yes	24	10%
No	208	90%
Do data collectors ever fill in data sheets after they were supposed to?		
Yes	158	68%
No	74	32%
If yes to the previous question, when do the data collectors fill in the late data sheets?		
By the end of the session	54	34%
By the end of the hour	12	8%
By the end of the day	69	44%
By the end of the week	20	13%
By the end of the month	3	2%

collectors ever fill data in before the events they were supposed to be observing, the second question asked if data collectors fill data in after they were supposed to, and the third question asked for more information about late data entry. When asked if data collectors record data before or after they are supposed to, 10% of the participants indicated that data collectors had recorded data early, and 68% of the participants indicated that data collectors recorded data after they were supposed to. Of the participants who indicated that data collectors record data late, 34% said that data is entered by the end of the session, 8% said that it was entered by the end of the hour, 44% said it was entered by the end of the day, 13% said it was entered by the end of the week, and 2% said it was entered by the end of the month.

The information about late/early data entry provided by the participants indicated that late data entry is much more common than early data entry. Although the low reported rate of issues with early data collection is good because early data entry is more problematic than late data entry, late data entry can still lead to issues with the accuracy and reliability of the data (see Morris & Peterson, 2020). When the participants of this survey were asked about when late data entry was completed, the most common answer was by the end of the day. Although data entered at the end of the day may be most convenient for the staff due to reduced competing contingencies, research has demonstrated that accuracy and reliability of data decrease as the latency to data entry increases (Jasper & Taber-Doughty, 2015; Taber-Doughty & Jasper, 2012). In fact, Jasper and Taber-Doughty (2015) specifically compared data collected immediately after a behavior, at the conclusion of the lesson, and at the end of the day. The results of that study demonstrated better accuracy and reliability when data were collected immediately or at the end of the lesson as compared to at the end of the day.

Data Collection Systems/Arrangement

Information about the data collection systems used was separated into two categories. The first category focused on the general data collection system, and the second category focused on the nuances of paper and electronic data collection systems. See Tables 5 and 6 for the detailed results of each category.

When asked about their general system for data collection, 59% of the participants said that their funding sources require specific data reporting, and 66% said that they used a prescribed data collection system that was provided through their agency/workplace. Eighty-one percent of the participants said that they use similar data collection methods or sheets for all of their clients. When asked to clarify how similar their data collection was across clients, about half of the participants (56%) indicated that only the format was similar. When asked what variables influenced their selection of the data collection system, the participants provided

a wide range of options, with the most common being the use of electronic systems (38%) and considerations for data displays (33%). Therefore, the largest influence on the selection of a data collection method appears to be the methods that are being used for other clients.

One of the most interesting findings within the data collection system/arrangement section was that a majority of the participants (70%) rated their satisfaction with their current data collection system as either extremely satisfied (20%) or somewhat satisfied (50%). The rest of the participants were either neither satisfied nor dissatisfied (9%), somewhat dissatisfied (19%), or extremely dissatisfied (3%). Considering that the participants had reported in previous sections that they were concerned about the accuracy (76%) and reliability (72%) of the data produced by their data collection systems and that 85% of the participants reported that data collectors sometimes fail to collect data, it is unclear why so many participants would report satisfaction with their data collection system. One potential explanation for the reported satisfaction despite the noted concerns of the data collection system is that participants attribute errors in data collection to the data collector rather than the system itself. However, that information was not specifically targeted within the current survey.

The last seven questions in the data collection systems/arrangements content area of the survey focused on the modality of the data collection system (i.e., paper vs. electronic data collection systems). Nearly half (45%) of the participants indicated that they use electronic data collection systems, while 79% reported using paper data collection systems. Thus, despite the increasing availability of electronic systems and guidance on using them (Dixon 2003; Morris & Peterson, 2020; Sleeper et al., 2017), a majority of the participants of this study continue to use paper data collection. Furthermore, 25% of the participants reported using both electronic and paper collection systems, which makes up 45% of the participants who reported using electronic data collection systems. The overlap of electronic and paper systems suggests that the use of electronic data collection systems alone (without the addition of paper data collection) is not yet a common practice within behavior analysis.

After the participants indicated their use of electronic and paper data collection systems, they provided specific information about their systems. Participants who reported using electronic data collection systems were asked to specify the type of electronic system they use, which produced varied responses. The two most common systems for electronic data collection were Catalyst (22%) and Central Reach (22%), but nearly 40% of the participants specified the use of unique electronic data collection systems. Participants who reported using paper data collection systems were asked to specify the components of their system. A majority of the participants reported using a one-page data collection sheet (60%), printing in greyscale (86%), and including instructions (84%) and

Table 5 General data collection systems/arrangement

Question Category	<i>n</i>	Percentage
Are there specific data reporting requirements through the funding source(s)?		
Yes	137	59%
No	95	41%
Do you have a specified data collection sheet or measurement tracking system prescribed by your agency/workplace?		
Yes	152	66%
No	80	34%
Do you use a similar data collection method or datasheet for all of your clients?		
Yes	187	81%
No	45	19%
If yes to the previous question, how similar are the data collection methods across clients?		
Exact Replication	15	8%
Minimal Changes	19	10%
Same Format	105	56%
Completely Individualized	49	26%
Did any of the following influence the selection of the data collection system?		
Available templates	40	17%
Considerations for data displays	76	33%
Electronic data collection systems	88	38%
Research Publications	34	15%
Published decision trees	10	4%
Funding Requirements	36	16%
Treatment Manual	19	8%
None	52	22%
Other	35	15%
Are you satisfied with your current data collection system?		
Extremely Satisfied	47	20%
Somewhat Satisfied	115	50%
Neither Satisfied nor Dissatisfied	20	9%
Somewhat Dissatisfied	43	19%
Extremely Dissatisfied	7	3%

operational definitions (79%) on the datasheet. Although the presenting data collection systems in greyscale that include instructions and definitions were common amongst the participants of this study, no published research directly evaluates the utility of any of those components.

Data Collector Training

The participants answered four questions about the concerns about data collection integrity. See Table 7 for the detailed results for this section.

Almost all participants (99%) indicated that their data collectors are trained on the data collection system they use. While the inclusion of the individual components of behavioral skills training (BST) ranged between 71% and 93% across components, the inclusion of all of the components combined together was only 58% – meaning that the participants of this study appeared to use components of

evidence-based training consistently but did not consistently provide the full package. Therefore, the effectiveness of the training procedures used by the participants of this study may not have been as effective as a complete BST training model (Ward-Horner & Sturmey, 2012).

Many (84%) of the participants of this study also reported providing refresher trainings to the data collectors. When asked when those refresher trainings were provided, a majority of the participants (63%) reported doing so “as needed.” Thus, refresher trainings appear to most commonly occur when issues arise, and intervention is needed rather than proactively refreshing the training to prevent issues from arising.

Addressing Data Collection Integrity Issues

The final section of the survey presented two questions about addressing data collection integrity issues. See Table 8 for the detailed results for this section.

Table 6 Paper versus electronic data collection systems

Question Category	<i>n</i>	Percentage
Do you use electronic data collection systems?		
Yes	104	45%
No	128	55%
If yes to the previous question, what type of electronic data collection system do you use?		
Catalyst	23	22%
Central Reach	23	22%
Google Applications	9	9%
ReThink	7	7%
Other	42	40%
Do you use paper data collection systems?		
Yes	183	79%
No	49	21%
If yes to using paper data collection systems, how many sheets of paper are necessary for daily data collection?		
A Single Sheet	110	60%
Multiple Sheets	73	40%
If yes to using paper data collection systems, what color are sheets printed as?		
Color	25	14%
Greyscale	158	86%
If yes to using paper data collection systems, are instructions provided on the datasheet?		
Yes	153	84%
No	30	16%
If yes to using paper data collection systems, are operational definitions of target behaviors provided on the datasheet?		
Yes	144	79%
No	39	21%
Uses both electronic and paper data collection systems.	57	25%

Table 7 Data collector training

Question Category	<i>n</i>	Percentage
Do the data collectors receive any training on the data collection systems they use?		
Yes	229	99%
No	3	1%
If yes to data collectors being trained, what components are included in the training?		
Written instruction	166	72%
Spoken instruction	213	93%
Modeling	212	93%
Rehearsal	163	71%
Practice and Feedback	204	89%
All	132	58%
If yes to data collectors being trained, are refresher trainings provided?		
Yes	192	84%
No	37	16%
If yes to refresher trainings, approximately how often do you do refresher trainings?		
Once per Week	9	5%
Once a Month	27	14%
Once Every Couple of Months	27	14%
Annually	9	5%
As Needed	120	63%

Table 8 Addressing data collection integrity issues

Question Category	<i>n</i>	Percentage
Are any interventions used to <i>maintain</i> appropriate data collection?		
Feedback	208	90%
Monitoring	207	89%
Goal Setting	62	27%
Incentives	31	13%
Other	16	7%
None	7	3%
Are any interventions used to <i>improve</i> data collection?		
Feedback	203	88%
Monitoring	190	82%
Goal setting	74	32%
Incentives	32	14%
Other	11	5%
None	15	6%

When asked if interventions are used with data collectors to *maintain* appropriate data collection, 90% of participants indicated that feedback was used, 89% indicated that monitoring was used, 27% indicated that goal setting was used, 13% indicated that incentives were used, 7% indicated that other strategies were used, and 3% indicated that no strategies were used. Finally, when asked if interventions are used to *improve* appropriate data collection, 88% of participants indicated that feedback was used, 82% indicated that monitoring was used, 32% indicated that goal setting was used, 14% indicated that incentives were used, 5% indicated that other strategies were used, and 6% indicated that no strategies were used. However, data from the primary data collector section suggest that only 14.1% of data collection includes a secondary observer collecting data. Therefore, it is unclear if supervisors commonly collect data on the client's behavior to assess the accuracy and reliability of their data collectors. If supervisors do not collect data on the client to compare to the data collector's data, it is unclear what the monitoring and feedback would be focused on.

Discussion

Cooper et al. (2020, p. 106) listed three major contributors of human measurement error (DCI issues) that consisted of poorly designed measurement systems, inadequate observer training, and unintended influences on observers. However, little research was cited to support their assertions. The current study provides an interesting contribution with preliminary support of Cooper et al.'s three major risk factors.

Poorly Designed Measurement Systems

Poorly designed measurements systems are the most difficult risk factor to evaluate with the data obtained through this study, and perhaps in general, because of the lack of research on measurement design effects on data collectors. Specifically, behavior analytic research is lacking on best-practice measurement system designs, which precludes the comparison of reported practices to an unambiguous standard. For example, little-to-no research is available that evaluates the effects of aesthetic features of data collection systems (e.g., color, arrangement, etc.), the effects of response effort (e.g., the amount and type of data being collected), and the data collection requirements of funding and organizations. Therefore, the evaluation of this risk factor through the current survey data cannot be conclusive but still provides interesting considerations.

The sections of the survey that were most relevant to poorly designed measurement systems were those related to paper and electronic data collection systems, the reported concerns about DCI, and general data collection systems/arrangement. The participants of the survey reported mixed use of electronic data collection systems as compared to or in addition to paper data collection systems. Although previous studies have compared electronic data collection systems to paper data collection systems (e.g., Sleeper et al., 2017; Tarbox et al., 2010), there is not enough data to conclude the superiority of either arrangement. Thus, the use of either system cannot be labeled as a poorly designed system. Instead, what may be more interesting are the specific components of the electronic and paper data collection systems. For example, a majority of participants reported using single-page, greyscale paper data collection systems with instructions and operational definitions included. Researchers interested in measurement systems design could consider evaluating these components to identify any major flaws with the commonly used arrangements. Similarly, researchers could compare specific components of common electronic data collection systems to identify any potential design flaws. However, until more research is conducted on this topic, the data produced in this section of the survey is merely descriptive of the general state and provides no evaluation of the relative prevalence of poorly designed systems.

Despite the limited evaluation on how specific components of the data collection system may contribute to the poor design of a measurement system, potentially useful information related to measurement system design was produced in other sections of the survey. For example, a majority of the participants reported satisfaction with their data collection systems, which indicates that the participants were unaware of any major design flaw in their measurement system. However, a majority of the participants also reported concerns about accuracy, reliability, and issues completing

the data collection requirements. Although issues with the accuracy, reliability, and completeness of data collection could be caused by issues other than the design of the system, it appears that the data collection systems that the participants were satisfied with at the very least did not prevent issues with DCI. Therefore, researchers interested in data collection system design could consider research examining the effects of measurement design on the accuracy, reliability, and completeness of data, as well as the intersection of design satisfaction of the data collectors.

The final survey section related to measurement design was focused on general data collection systems/arrangements. Interestingly, the results of the survey indicated that data collection (measurement) designers are influenced by funding source and organization requirements, as well as the data collection systems used with their other clients. Again, while no data are available to judge the impact of these influences, this information can help direct the efforts of researchers interested in impacting data collection design issues. With the knowledge of funding, organizational, and other client data collection system influences on design, advocating for effective data collection systems at the funding and organizational level may be the most efficient means of ensuring the use of effective data collection systems.

Inadequate Observer Training

Two sections of the survey pertained to data collector (observer) training. The first applicable section identified the primary type of data collector used. The results of the survey indicated that RBTs and other direct-care staff were the most common data collectors. Thus, much of the focus on data collector training should account for the nuances of that group. Specifically, the use of RBTs and direct-care staff add complications to training due to their common turnover (DiGennaro Reed & Henley, 2015) and competing responsibilities (discussed more in the next section). The results of this study indicated that an average of 2.0 data collectors are used per day, and an average of 5.4 data collectors are used across a 3-month period for a single client. This means that data collector training is needed across many individuals for any given client, which may be time and resource-consuming.

Nearly all of the participants of this study reported providing some level of training to their data collectors. When asked about the specific arrangement of the trainings, a majority of the participants reported including at least one component of BST within their training. However, only 58% of the participants reported using all of the components of BST. Thus, one important conclusion from this study is that the most effective and supported training strategy is not commonly used for data collection responsibilities. Similarly, when asked about refresher trainings, a majority of

the participants reported conducting them, but mostly only doing so “as needed.” Although specific information was not gathered to identify when supervisors decide that a refresher is needed, the results suggest that a proactive approach to refresher trainings is not likely.

Unintended Influences on Observers

Like inadequate observer training, it is important first to consider the specific type of data collector to understand the unintended influences. As previously stated, the most common type of data collector reported in this survey were RBTs/direct-care staff. Therefore, the information reported in the survey related to unintended influences on observers is most relevant to RBTs/direct-care staff but could be true for other types of data collectors to a lesser or greater extent.

The primary risk factor identified in this survey that could produce unintended influences on the observer was the competing responsibilities of the data collectors. Specifically, over 99% of the participants reported that their data collectors had at least one other responsibility while collecting data. The most common competing responsibility identified in this survey was implementing treatment for the client whose data is being collected. Other prominent competing responsibilities included caregiving tasks, implementing treatment for other clients (an average of 2.97 at a time), and collecting data for other clients (an average of 2.63 at a time). Given the response effort required to complete these competing tasks, it seems evident that they could distract the data collector from their data collection and create issues with accuracy, reliability, and completeness of data.

Risk Factor Conclusions

Viewing the results of this survey through the lens of the risk factors proposed by Cooper et al. (2020) provides interesting insight into which risk factors may be more commonly present in applied settings. Although information related to each risk factor was obtained in the course of the survey, conclusions about the validity and prevalence of each risk factor are not possible without more research. Specifically, the lack of research on design features of data collection systems renders the information obtained in the study related to poorly designed measurement systems inconclusive. Aside from the lack of information to evaluate the design features, the survey was able to provide information that aligns with and supports the risk factors related to inadequate observer training and unintended influences on observers. The most common issues related to both categories are that the participants reported not using every component of BST when training data collectors and assigning their data collectors multiple responsibilities likely to produce unintended influences, such as distractions and competing responsibilities.

Thus, another finding of this study is that most of the risk factors of human error in measurement described by Cooper et al. seem to be prevalent within typical data collection arrangements.

Intervention Considerations

Given the prevalence of issues related to the risk factors for DCI issues and the concerns about accuracy, reliability, and completeness of data identified in this survey, consideration of intervention strategies to address DCI issues is warranted. To begin with, practitioners concerned about DCI should review the risk factors described by Cooper et al. (2020, p. 106) and arrange systems and procedures to prevent and remediate issues related to each.

Practitioners should consider the design of the data collection system. Although more research is needed to validate specific components of data collection systems, some aspects of the design of data collection systems have been supported through research. For example, Morris and Peterson (2020) demonstrated that a basic electronic data collection system did not improve data collection without the inclusion of behavioral interventions in the form of programmed prompts and automated feedback. Therefore, careful selection of components of data collection systems should take place, considering the best information available. One potentially critical component of effective data collection systems design is the efficiency and usability of the system. While research on specific design features to improve efficiency and usability are lacking, one design approach involves the consideration of micro- versus macro-data collection arrangements.

Micro-data are data that are focused on and collected during specific units of time, repeatedly over time. Micro-data collection is what is most commonly reported in behavior analytic research – whether researchers use continuous or discontinuous measurement systems. Alternatively, macro-data refers to data that are focused on and collected during larger sections of time. For example, when collecting data on self-injurious behavior, a micro-data collection system would target every instance (or interval) of self-injurious behavior over time, while a macro-data collection system might rely on a baseline and treatment probe rather than continuously collecting data. Micro-data systems have inherent clinical benefits – namely, that they allow for continuous assessment and data-based decision-making. Macro-data systems, however, are less effortful and maybe more feasible than micro-data collection in some circumstances. For example, macro-data collection may be a more realistic data collection system for parents or other caregivers who are asked to collect data on top of their other responsibilities. However, research is needed to evaluate the sensitivity of macro-data collections systems to detect treatment effects,

as well as their correspondence with micro-systems. In addition, although we presume that macro-data collection may be easier for data collectors as it is less effortful than micro data collection, it is important to evaluate data collectors' reliability with and preference for this type of system.

Practitioners should also consider utilizing evidence-based strategies for training data collectors. Many of the participants of this study reported using some components of evidence-based training strategies but failed to incorporate the entire evidence-based training package. To ensure the best possible training, complete use of the most supported strategies should be employed. Additionally, refresher trainings should be proactively arranged to avoid deterioration of DCI and the issues that would result from declining DCI (i.e., inaccurate data). Supplemental strategies such as effective monitoring, feedback, and other behavioral interventions could be used to complement trainings.

Finally, practitioners should attempt to minimize unintended influences on data collectors. Although the complete removal of distractors and competing responsibilities is probably not possible in most situations, reducing them as much as possible will lessen unintended influences on data collectors. For example, the participants of this study reported that their data collectors are typically responsible for the data collection of more than one client at a time while also being responsible for delivering those clients' treatment. If reducing the number of clients served by the data collector is not possible, careful consideration and coordination of what they are expected to do with their clients and data collection would be necessary to ensure that their expectations are feasible.

General Conclusion

The results of this study, while preliminary, provide helpful information about DCI within the practice of applied behavior analysis. Specifically, the survey produced descriptive information about potential risk factors related to DCI, as well as considerations for interventions to prevent/address DCI issues. However, the limitations of this study hinder decisive conclusions about the variables involved with DCI issues. Thus, more research is needed on the topic of DCI.

This study contained multiple limitations. The first limitation of this study was that the survey only focused on one aspect of data collection – that which focused on problem behaviors in applied settings. Therefore, other data collection arrangements such as skill acquisition were not represented in this paper. The second limitation of this study was that the survey targeted the supervisors of the data collection system instead of the data collectors themselves. Focusing exclusively on the supervisors of the data collection may have produced biased reporting in areas that reflect the quality of

the supervision. For example, nearly all of participants (99%) reported training the data collectors, which could be inflated. Surveying data collectors, in addition to, or instead of, the supervisors could have produced more insightful information about data collection including the quality and type of training/supervision received for data collection. The third limitation of this study was the limited number of participants (232). Although the length of the survey likely deterred some participation for this study, the demographic information of the participants suggests the small sample included in this study is still representative of the field.

Future research on DCI should focus on further understanding the variables related to DCI issues. This research could be done through surveys of data collectors or the supervisors of other types of data collection. In addition to survey studies, observational research collecting descriptive data on baseline data collection practices would help create a clearer understanding of the current state of data collection within treatment programs. Finally, researchers should also go beyond collecting information about the current state of data collection by continuing to develop effective and efficient interventions for improving DCI. For example, experimental research could evaluate the utility of various system design components, strategies for training data collectors, and strategies to reduce unintended influences on data collectors.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s40617-022-00684-x>.

Acknowledgements We thank Neil Deochand for his feedback during the creation of the survey.

Declarations

Research Involving Human Participants and/or Animals This project was deemed exempt by an Institutional Review Board.

Informed Consent Informed consent was obtained from all individual participants included in the study.

Conflicts of Interest/Competing Interests Not Applicable.

References

- Behavior Analyst Certification Board (2020). *Ethics Code for Behavior Analysts*. <https://www.bacb.com/wp-content/uploads/2020/11/Ethics-Code-for-Behavior-Analysts-2102010.pdf>
- Behavior Analyst Certification Board (2021, January 5). *BACB Certificiant Data*. <https://www.bacb.com/bacb-certificiant-data/>
- Cooper, J. O., Heron, T. E., & Heward, W. L. (2020). *Applied behavior analysis*. Pearson.
- DiGennaro Reed, F., & Henley, A. J. (2015). A survey of staff training and performance management: The good, the bad, and the ugly. *Behavior Analysis in Practice*, 8(1), 16–26. <https://doi.org/10.1007/s40617-015-0044-5>
- 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>
- Gresham, F. M. (1989). Assessment of treatment integrity in school consultation and prereferral intervention. *School Psychology Review*, 18(1), 37–50. <https://doi.org/10.1080/02796015.1989.12085399>
- Jasper, A. D., & Taber-Doughty, T. (2015). Special Educators and Data Recording: What's Delayed Recording Got to Do With It? *Focus on Autism and Other Developmental Disabilities*, 30(3), 143–153. <https://doi.org/10.1177/1088357614547809>
- Johnston, J. M., Pennypacker, H. S., & Green, G. (2020). *Strategies and tactics of behavioral research and practice*. Routledge.
- Jones, S. H., St. Peter, C. C., & Ruckle, M. M. (2020). Reporting of demographic variables in the Journal of Applied Behavior Analysis. *Journal of Applied Behavior Analysis*, 53(3), 1304–1315. <https://doi.org/10.1002/jaba.722>
- 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. <https://doi.org/10.1007/s40617-015-0063-2>
- Morris, C., & Peterson, S. M. (2020). A Component Analysis of an Electronic Data Collection Package. *Journal of Organizational Behavior Management*, 40(3-4), 210–232. <https://doi.org/10.1080/01608061.2020.1771505>
- 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. <https://doi.org/10.1901/jaba.2006.55-05>
- 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. <https://doi.org/10.1002/bin.1371>
- 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*, 37(1), 83–95. <https://doi.org/10.1080/01608061.2016.1267065>
- Slocum, T. A., Detrich, R., Wilczynski, S. M., Spencer, T. D., Lewis, T., & Wolfe, K. (2014). The evidence-based practice of applied behavior analysis. *The Behavior Analyst*, 37(1), 41–56. <https://doi.org/10.1007/s40614-014-0005-2>
- 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. <https://doi.org/10.1177/1088357612451121>
- 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. <https://doi.org/10.3758/BF03192763>
- 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. <https://doi.org/10.1016/j.rasd.2009.07.008>
- Ward-Horner, J., & Sturmey, P. (2012). Component analysis of behavior skills training in functional analysis. *Behavioral Interventions*, 27(2), 75–92. <https://doi.org/10.1002/bin.1339>
- Zane, T., & Davis, C. (2013). Differential Reinforcement Procedures of Other Behavior (DRO). In F. R. Volkmar (Ed.), *Encyclopedia of Autism Spectrum Disorders*. Springer. https://doi.org/10.1007/978-1-4419-1698-3_1903

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.