ORIGINAL PAPER

Accuracy of Teacher-Collected Descriptive Analysis Data: A Comparison of Narrative and Structured Recording Formats

**Dorothea C. Lerman · Alyson Hovanetz ·
Margaret Strobel · Allison Tetreault**

Abstract Recording the antecedents and consequences of problem behavior for the purposes of conducting descriptive analyses (called "A-B-C recording") can be particularly challenging, given the multiple variables that are commonly present in the natural environment. Nonetheless, psychologists and behavioral consultants often must rely on parents and teachers to collect descriptive data in community settings. Few studies have examined the accuracy of data collected by caregivers or the best way to train people to collect these data. The purpose of this study was to examine the accuracy of teacher-collected data with two commonly used recording formats after the teachers received the type of training typically provided to public school teachers. Thirteen of the sixteen participants reported that they had previously collected A-B-C data in their classrooms as part of the functional assessment process. Participants used narrative and structured A-B-C data forms to collect data while watching videotapes of scripted exchanges between actors. They collected data more accurately when using the structured format compared to the narrative format and indicated a preference for this method of assessment. These findings have important implications for training educators to collect descriptive data.

Keywords A-B-C recording · Data collection · Descriptive analysis · Functional assessment · Problem behavior

Introduction

Descriptive analyses of problem behavior are conducted by observing behavior under naturalistic conditions and determining which environmental events, if any,

D. C. Lerman (✉) · A. Hovanetz · M. Strobel · A. Tetreault
Psychology Department, University of Houston, 2700 Bay Area Blvd., P.O. Box 245,
Clear Lake, Houston, TX 77058, USA
e-mail: lerman@uhcl.edu

are highly correlated with the behavior. In practice, a main goal of this assessment is to generate hypotheses about the function of problem behavior by identifying frequently occurring antecedents and consequences. It is commonly recommended to include this type of analysis, along with information obtained from teachers, parents, and other caregivers, as a component of functional assessments in community settings (Chandler and Dahlquist 2006; O'Neill et al. 1997). Results of functional assessments are then used to develop function-based interventions for problem behavior.

Research findings indicate that the descriptive analysis does not afford the degree of precision in identifying behavioral function that is associated with the functional analysis approach, in which suspected causal variables are directly manipulated (e.g., Lerman and Iwata 1993; Thompson and Iwata 2007). Nonetheless, descriptive analysis can be an important tool for several reasons. First, information obtained through descriptive observations can be used to guide the design of a functional analysis (Hanley et al. 2003). Doing so could permit more efficient functional analyses or refinements in functional analysis conditions that are needed to produce clear results (e.g., Tiger et al. 2006). Second, a descriptive analysis is the best option for conducting a functional assessment when a functional analysis is not possible (e.g., when evaluating low frequency behavior or variables that are not easily controlled, such as peer attention). In such cases, either formal or informal descriptive observations may provide information that is important to treatment development. Third, when relevant environmental events are exceedingly clear, descriptive observations might render a functional analysis unnecessary and, thus, permit more immediate intervention. Finally, the descriptive analysis can provide a baseline of naturally occurring events (e.g., teacher responses to problem behavior) that may be useful for examining intervention outcomes (e.g., results of teacher training).

Ideally, individuals collecting descriptive data in school settings should be highly trained and experienced, preferably school psychologists or behavior specialists who routinely conduct functional assessments. However, in practice, psychologists and specialists often must rely on teachers and other school personnel to conduct these observations in their absence. Teacher-collected data are particularly useful when problem behavior occurs infrequently and, thus, might require more observation time than a school psychologist can provide. In some cases, the behavior may primarily occur in settings or at times that complicate direct observation by the school psychologist or behavior specialist. Teacher-collected data also can provide supplemental information about events surrounding the behavior that occur outside of the psychologist's circumscribed observation periods.

For these reasons, data collected by individuals with limited training and experience can be a crucial component of the functional assessment process. Conducting descriptive observations, however, presents special challenges for teachers, who rarely can devote their full attention to data collection due to competing responsibilities and other distractions. Furthermore, the relevant antecedents and consequences of problem behavior may involve the behavior of the teachers themselves, which may be more difficult to accurately and objectively record than the behavior of someone else. Finally, due to the complexity of

naturalistic environments, observers often must detect multiple events that occur simultaneously either prior to or following behavior.

Descriptive Recording Formats

The measurement system itself also may impact the ease and accuracy of teacher-collected descriptive data. Two descriptive analysis formats that have been recommended for use in school settings are narrative A-B-C recordings and structured (checklist) A-B-C recordings (Smith et al. 2007; Lerman and Iwata 1993). For both formats, the observer uses a specially prepared data sheet to document the occurrence of certain antecedents and consequences each time the target behavior is observed. The observer either provides a narrative description of events surrounding the behavior (for narrative A-B-C recording) or selects from a list of events that have been specified prior to the observation (for structured A-B-C recording). Although narrative A-B-C recordings can generate a lot of information, they are more time consuming to complete than structured recording forms. It may be difficult to capture all relevant events when using the narrative format in naturalistic settings. On the other hand, structured A-B-C recordings require the observer to instantly classify an event as one of the specific antecedents or consequences that appear on the data form. This may be challenging when the prescribed event (e.g., "removal of instructions, "delivery of attention") can take a variety of forms in the natural environment. Thus, greater expertise and training may be needed to record accurately using the structured format.

Accuracy of Teacher-Collected Descriptive Data

Despite the benefits and challenges associated with teacher-collected descriptive data, little research has been conducted on the accuracy of these data or the best way to train educators to collect them. In a number of studies, teachers were taught to implement functional assessment procedures (e.g., Grey et al. 2005; Maag and Larsen 2004) or to participate in the assessment process with the assistance of a behavioral consultant (e.g., Kamps et al. 2006; Watson et al. 1999). However, few of these studies examined the accuracy or reliability of teacher-collected descriptive data. In fact, experimenters collected the descriptive data in a majority of these studies.

In a notable exception, Ellingson et al. (2000) evaluated the reliability of teacher-collected data on the antecedents and consequences of problem behavior. The teachers used a check-list format to note the occurrence of behavior, along with the relevant antecedents and consequences. Interobserver agreement was determined by having the experimenters collect data simultaneously but independently during the four 30-min observations. For the three teachers, mean percentages of agreement on the antecedents and consequences recorded for each behavior were 92, 86, and 73%, with a range of 57–100% across observations. This level of agreement between the teachers and the experimenters was sufficient to

produce the same conclusions about the consequence maintaining problem behavior, which was attention for all of the children. Although these results suggest that relatively inexperienced observers can collect descriptive data with a high degree of precision, the list of potential antecedents and consequences that appeared on the observers' recording forms was fairly limited. For example, escape from demands and access to tangibles were not included as potential consequences on the checklist. The circumscribed list may have increased the likelihood of agreement between the observers.

Further research is needed to determine the degree to which educators who have received the type of training typically provided to public school teachers (i.e., lecture-based in-service; National Research Council 2001; Scheuermann et al. 2003; SPeNSE 2002) can collect descriptive data accurately on complex environmental events. Two commonly recommended formats—narrative A-B-C recording and structured A-B-C recording—should be examined in further investigations because each offers advantages and disadvantages for caregivers and practitioners. Narrative recordings are more likely than structured recordings to reveal idiosyncratic sources of control. For example, narrative recordings might indicate that only certain types of instructions set the occasion for problem behavior. Nonetheless, as noted above, narrative recording forms require a substantial amount of time to complete, particularly if the behavior occurs frequently. Narrative descriptions of events also may be more subjective, and they do not lend themselves as easily to quantitative analysis. Relative to the narrative format, structured A-B-C recordings take less time and produce more objective data that can be readily quantified and summarized. Nonetheless, as noted above, observers must immediately categorize naturalistic events as one of the prescribed antecedents and consequences listed on the data form.

Purpose of the Study

The main purpose of the study was to examine the accuracy of descriptive data collected by teachers. As an initial examination, we sought to evaluate teachers' accuracy under "ideal" and highly controlled conditions while still retaining the complexity of events typically operating in naturalistic classroom settings. This was accomplished by having teachers collect data while watching videotape segments of scripted exchanges between actors. Use of the scripted exchanges permitted control over the type and complexity of antecedents and consequences surrounding instances of problem behavior. Control over these events was especially important because another purpose of the study was to compare the teachers' accuracy when using the narrative A-B-C and structured A-B-C recording formats. The conditions were considered "ideal" for data collection purposes because the teacher was in a quiet room with no distractions or other responsibilities. Thus, results should indicate the optimal levels of accuracy that might be expected under more naturalistic conditions. A final purpose of the study was to examine the acceptability of each format to the teachers and their preference for one format over the other.

Method

Participants and Settings

Participants were 13 certified special education teachers and 3 paraprofessionals who were enrolled in a 5-day intensive training program for teachers of students with developmental disabilities. The participants' ages ranged from 25 to 66 years. The teachers had an average of 7.9 years of experience (range, 3–18 years), and the paraprofessionals had an average of 18 years of experience (range, 14–25 years). Twelve participants reported prior experience with narrative A-B-C recording, three participants reported prior experience with structured A-B-C recording, and three participants reported no prior experience with A-B-C recording. The training program covered a variety of topics, including the basic principles of applied behavior analysis, preference assessments, and effective direct teaching (e.g., using prompts and reinforcement). The training was conducted in the library of a public school, and the video scoring sessions were conducted in multiple unused classrooms.

Materials

Two versions of a 15-min videotape were constructed prior to the study. First, the experimenters generated a list of specific types and combinations of antecedents and consequences and then developed two different scripts that incorporated these events within three classroom situations (individual instruction, group instruction, and snack). Next, adult actors were videotaped while they followed the prepared scripts. One actor pretended to be a teacher and the remaining three actors pretended to be students with developmental disabilities. The roles of the actors were the same across all video segments and videos. One student, Jane, engaged in 11 instances of aggression and disruption on each video, defined as hitting others, throwing objects, tearing materials, or pushing materials off furniture. Each instance of problem behavior was separated by a minimum of 1 min to give the teachers a reasonable amount of time to record events surrounding the behavior. The antecedents and consequences that occurred in relation to problem behavior were intended to sample the range of complexity often observed in classroom settings. One or two different antecedents (e.g., demand delivery, tangible removal) and up to three different consequences (e.g., demand removal, attention delivery, tangible delivery) were arranged to occur simultaneously. The specific events shown in the video were designed to represent those that frequently occur in classroom settings. Examples of the antecedents and consequences are shown in Table 1. The same number, type, and combinations of antecedents and consequences occurred in the two videos but in a different order and with different scripted exchanges (e.g., different demand situations, materials, etc.). The videos were designed to indicate multiple control of problem behavior (i.e., behavior maintained by both attention and escape from demands).

Two different A-B-C recording sheets, portions of which are shown in Fig. 1, were developed for the study. The designated target behaviors were pre-printed on

Table 1 Examples of antecedents and consequences shown in the videotapes

Antecedents	Problem behavior	Consequences
A teacher sitting with Jane starts to speak to someone off-camera, ignoring Jane (removal of attention)	Jane hits the teacher	The teacher turns her attention to Jane and delivers a reprimand (delivery of attention)
A teacher tells Jane to finish her work (delivery of an instruction)	Jane pushes the materials off the desk	The teacher delivers a reprimand and then picks up all of the materials. When all of the materials have been picked up, it is time to leave the class for another activity (delivery of attention and escape from demand)
A teacher takes a toy away from Jane while telling her to get back to her desk (removal of a tangible and delivery of an instruction)	Jane hits the teacher	The teacher physically restrains Jane briefly and guides her back to the desk (delivery of attention)

both recording forms. Teachers placed a check in the appropriate space next to the description of the target behaviors on the form to indicate the occurrence of problem behavior. They then specified the antecedents and consequences for each instance through either narrative recordings (i.e., descriptions of events that preceded and followed the behavior) or checks marks. The narrative A-B-C form was printed on both sides of a single sheet of paper so that all 11 instances of the response could be recorded. The structured A-B-C form included instructions for summarizing the data and determining the degree to which each of four possible functions—positive reinforcement (in the form of attention), positive reinforcement (in the form of materials), negative reinforcement (in the form of escape from demands), and sensory stimulation—were supported by the data. Finally, participants were asked to complete rating forms intended to assess their acceptability of the two recording formats, as well as their relative preference for the two formats (see Fig. 2).

Procedures

All participants were first exposed to a 60-min group lecture that was designed to simulate "training as usual" received by many public school teachers (National Research Council 2001; Scheuermann et al. 2003; Study of Personnel Needs in Special Education 2002). The lecture covered the functions of problem behavior, common forms of antecedents and consequences, data collection via narrative and structured A-B-C recording, and A-B-C data interpretation and summary. During the lecture, the experimenter provided the teachers with operational definitions of the events listed on the structured A-B-C recording form, as well as examples of completed narrative and structured A-B-C forms. Immediately following the lecture, participants were escorted to separate classrooms by research assistants. Participants watched the videos on either laptop computers or television monitors. Half of the participants used the narrative recording form to score the first video and the structured form to score the second video. The other half completed the forms in the reverse order while watching the two videos.

Narrative (A-B-C) Recording Form

Behaviors		Antecedents	Consequences
Aggression/Disruption = Hitting others, throwing objects, knocking over furniture, pushing objects off furniture			
Aggression/Disruption = Hitting others, throwing objects, knocking over furniture, pushing objects off furniture			

Structured (A-B-C) recording Form

Behaviors									
Aggression/Disruption = Hitting others, throwing objects, knocking over furniture, pushing objects off furniture									
Antecedents									
Ignored by teacher; someone walked away									
Leisure material/food removed or denied									
Other request denied									
Given instruction/prompt to work									
None									
Consequences									
Attention, response block, told to "stop"									
Access to food/leisure materials/ activities									
Work requirement/instruction delayed/ removed									
Other activity delayed or removed									
Teacher walked away									
None									

Fig. 1 Portions of the narrative A-B-C recording form (*top panel*) and structured A-B-C recording form (*bottom panel*) completed by the participants while watching the videos

Prior to the first video, the participant was told that she would be collecting descriptive data on the aggression and disruption of a student named Jane. The research assistant reviewed the operational definitions of the target behaviors and instructions on completing the forms with the participant and answered any

Acceptability Rating Form

Please rate the narrative A-B-C recording form (or structured A-B-C recording form) along the following dimensions. Please circle the number which best describes your agreement or disagreement with each statement.

	Strongly Disagree	Disagree	Disagree Slightly	Slightly Agree	Agree	Strongly Agree
1. This would be an acceptable way to assess a child's problem behavior.	1	2	3	4	5	6
2. Most teachers would find this assessment form appropriate to use in the classroom.	1	2	3	4	5	6
3. I would suggest this assessment form to other teachers.	1	2	3	4	5	6
4. I would be willing to use this assessment form in the classroom setting.	1	2	3	4	5	6
5. This assessment form is consistent with those I have used in classroom settings.	1	2	3	4	5	6

****
Assessment of Preference

PLEASE ANSWER THE FOLLOWING: If I was going to choose just one of the assessment forms to use in my classroom, I would probably choose the (circle only one):

Narrative A-B-C   OR   Structured A-B-C

Fig. 2 Acceptability rating form completed by the teachers for each type of assessment after scoring the videos (*top panel*); statement completed by the participants to indicate preference for one format over the other (*bottom panel*)

questions. The participant was then given the relevant scoring sheet and asked to score the video. At the end of the video, the form was removed and the second A-B-C sheet was given to the participant. The participant scored the second video. After the data form was removed, the participant was asked to complete the assessment rating forms.

Data Analysis

The accuracy of the structured and narrative A-B-C recordings was calculated for each participant, as described in the sections below.

Accuracy of Structured A-B-C Scoring

The teacher-completed structured forms were compared to a "gold standard" data form that was constructed on the basis of the videotape scripts. To verify that the scripted antecedents and consequences were portrayed by the actors in the manner intended, an expert in functional assessment who was unaffiliated with the study

completed the structured A-B-C forms while watching the videos. The expert had a doctorate degree and more than 10 years of experience conducting functional assessments. The expert's data records were compared to the gold standard forms, and 100% agreement was obtained for one video. For the other video, a disagreement was obtained involving one of the consequences (specifically, the gold standard form indicated that two different consequent events occurred simultaneously and the expert scored just one event), resulting in 97% occurrence agreement and 100% nonoccurrence agreement (see below for description of accuracy calculations).

When calculating level of agreement between the teachers' forms and the gold standard form, events on the data form that were related to the same behavioral function (i.e., positive reinforcement [attention], positive reinforcement [materials], negative reinforcement [escape], or automatic reinforcement) were combined; however, antecedents and consequences were analyzed separately. For example, three functionally similar consequences for behavior maintained by negative reinforcement (escape) were "work requirements/instruction delayed or removed," "other activity delayed or removed," and "teacher walked away." Thus, a teacher's scoring was considered correct when any of those events was checked on her data sheet if one of these events was scored on the gold standard data record for a particular instance of problem behavior. Across 11 instances of problem behavior, the teacher had a total of 66 opportunities to agree or disagree on the antecedents and consequences of behavior. That is, a total of 66 antecedents and consequences could either be scored by the teacher or left unscored across the 11 instances of problem behavior. (The event "None" was not included as an opportunity, as it was considered a default category. None of the teachers incorrectly scored "None" as an antecedent or consequence).

Three different types of agreement were calculated for the antecedents and consequences (separately and combined). First, the percentage of agreement on occurrences of antecedents and consequences was calculated by totaling the number of events scored on the gold standard form that was also scored by the teacher and dividing the sum by the total number of scored events on the gold standard form (total occurrence was 30). The result reflected the percentage of events that was scored by the teacher. However, this measure does not take into account events on the data sheet that were accurately disregarded by the teacher. Thus, the percentage of agreement on nonoccurrences of antecedents and consequences was calculated by totaling the number of unscored events on the gold standard form that was also unscored by the teacher and dividing the sum by the total number of unscored events on the gold standard form (total nonoccurrence was 36). Both occurrence and nonoccurrence agreement provide an overall measure of accuracy but do not indicate the extent to the events surrounding particular instances of problem behavior were recorded accurately. Thus, a third measure was included. The percentage of agreement on the antecedents and consequences for each instance of problem behavior was determined by totaling the number of instances in which the teacher scored all of the same antecedents and consequences as that shown on the gold standard form and dividing by the total number of instances of problem behavior.

Two teachers (P 4 and P 12) scored just 10 of the 11 instances of problem behavior when using the structured A-B-C form. Due to the difficulty in identifying the specific response that was missed, levels of agreement were re-calculated 11 times for these participants to determine the accuracy of scoring with the omission of each instance of problem behavior on the gold standard form. In other words, we determined the level of agreement assuming that the first instance was missed, the second instance was missed, etc. The highest level of accuracy obtained via the 11 calculations was then included in the analysis for the participant. Given the purpose of the study, this less stringent approach seemed more reasonable than selecting the lowest level of accuracy.

Accuracy of Narrative A-B-C Scoring

A method for evaluating the accuracy of the narrative A-B-C recordings required special consideration. A "gold standard" narrative recording would have been difficult to generate given the numerous possible wording variations that could accurately specify the relevant antecedents and consequences of behavior. Furthermore, those interpreting and analyzing the written narratives (i.e., the experimenters) would have an advantage not typical in practice because the events were already known. Thus, 16 functional assessment "experts" were recruited to interpret the narrative recordings. This approach more closely approximated the type of analysis that occurs during routine clinical practice. All experts held a doctorate in behavior analysis, had a minimum of 8 years of experience conducting functional assessments, and were a co-author or primary author on at least three peer-reviewed papers that contained functional assessment data.

Each expert evaluated one teacher's set of narrative recordings that had been transcribed on an electronic data sheet and sent to the expert via e-mail. The experts were asked to examine the narrative recordings and to separately code each antecedent and consequent event into functional categories similar to those provided on the structured assessment. For instance, if a teacher noted the antecedent, "told the child to complete a worksheet" the expert would have coded this as an antecedent related to escape maintained problem behavior. Similarly if the teacher noted the consequence "walked away after being hit" the expert would have coded this as a consequence related to the delivery of escape. The experts selected the function(s) by placing a check mark next to one or more of the functions that appeared on the electronic data sheet below the narrative recordings for each instance of behavior (separately for antecedents and consequences). The experts returned the scored narrative recordings to the first author via e-mail. All of the accuracy calculations described above were then completed for each participant on the basis of the expert's analysis.

Three teachers missed one or two of the 11 instances of problem behavior when using the narrative A-B-C form (P 2 missed 2 instances, P 8 missed 1 instance, and P 12 missed 2 instances). As with the structured A-B-C forms, data for P 8 were re-calculated to determine the accuracy of scoring with the omission of each instance of problem behavior on the gold standard form. The highest level of agreement was then included in the analysis. A similar approach was used for P 2 and P 12, except

that the data were re-calculated to determine the accuracy of scoring with the omission of two instances of problem behavior on the gold standard form. All possible combinations of omitted pairs were included in the calculations, and the highest level of agreement was included in the analysis.

Results

The mean percentages of agreement (i.e., accuracy) of the recorded antecedents and consequences for participants using the narrative and structured formats are shown in the top panel of Fig. 3. Accuracy for occurrences, nonoccurrences, and instances of problem behavior were averaged across participants. The highest levels of accuracy were associated with nonoccurrences (mean, 91.5%), indicating that participants rarely recorded events that did not occur. Participants were moderately accurate for occurrences of antecedents and consequences (mean, 64.5%), with slightly higher levels of accuracy associated with the structured recording (mean, 69%) than the narrative format (mean, 60%). On the other hand, the percentage of problem behavior in which participants accurately scored all antecedents and consequences was relatively low (mean, 27%), with slightly higher levels of accuracy for the structured format (mean, 32%) than for the narrative format (mean, 22%). The bottom panel of Fig. 3 shows the same accuracy data but for antecedents and consequences separately. These data indicate that participants were somewhat
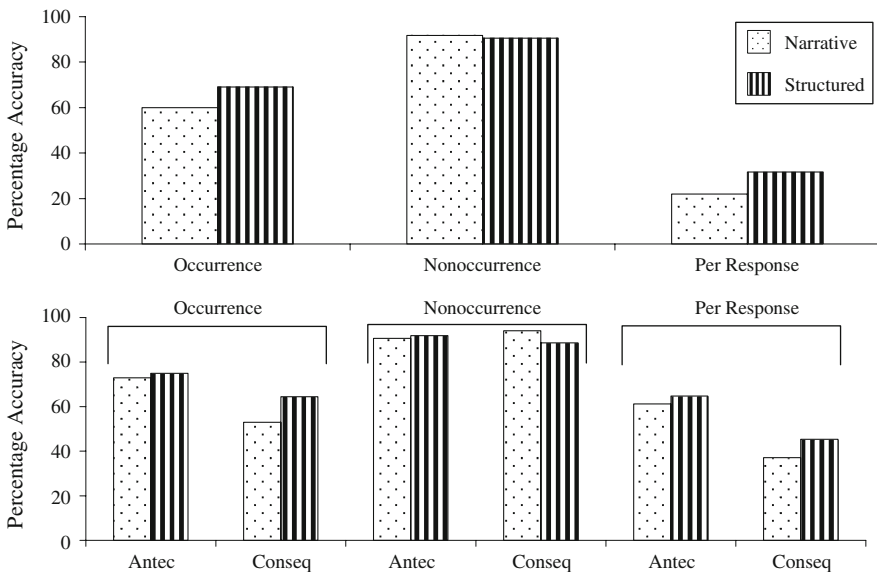

Fig. 3 The mean percentages of accuracy of the recorded antecedents and consequences for participants using the narrative and structured formats. Accuracy was calculated on the basis of overall occurrences of antecedents and consequences, nonoccurrences of antecedents and consequences, and percentage of problem behavior

more accurate when scoring the antecedents of behavior than when scoring the consequences.

The accuracy of individual participants for occurrences of antecedents and consequences (top panel) and instances of problem behavior (bottom panel) is displayed in Fig. 4. Accuracy when using the narrative format ranged from 47 to 86% for occurrences and from 9 to 80% for percentage of problem behavior across participants. When participants used the structured format, accuracy ranged from 59 to 94% for occurrences and from 27 to 90% for percentage of problem behavior. The differences in accuracy between the narrative and structured formats that appeared in the averaged results (see Fig. 3) were obtained for 9 of the 16 participants. Of the seven participants who showed similar or higher levels of accuracy when using the narrative format compared to the structured format, five had used the structured form to score a video prior to using the narrative form (P4, P6, P7, P10, P14, and P16). Thus, immediate prior experience with the structured form may have improved the narrative recordings of the participants.
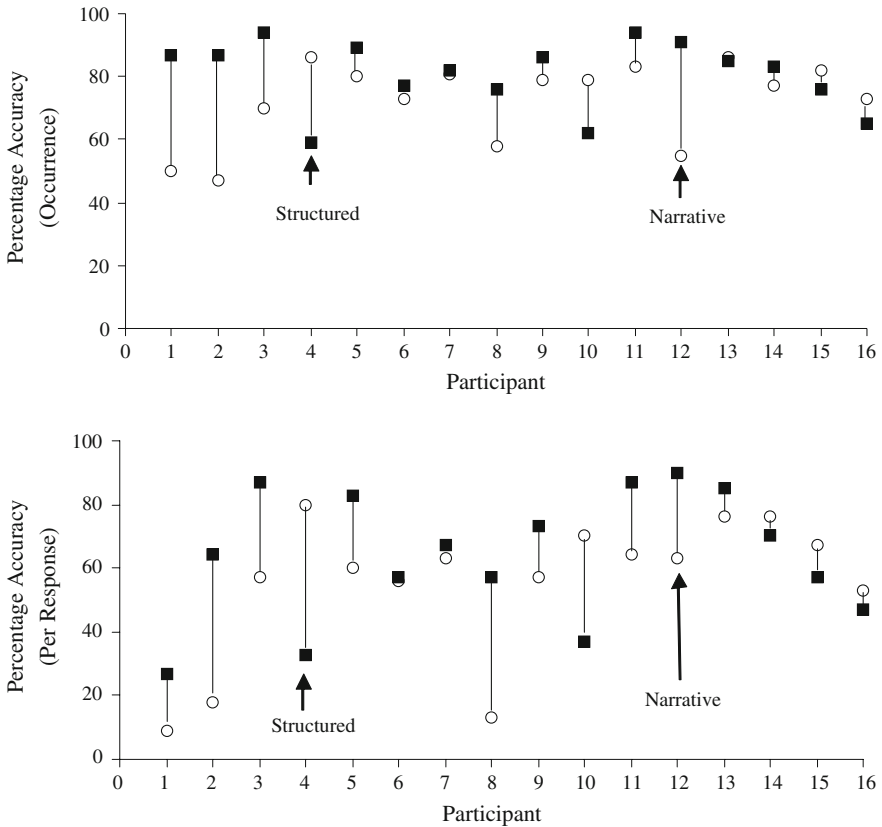

Fig. 4 The accuracy of individual participants using the narrative and structured formats when calculated on the basis of overall occurrences of antecedents and consequences (*top panel*) and percentage of problem behavior (*bottom panel*)

Table 2 Mean number and range of omission errors associated with the two recording

Event	True instances	Structured format	Narrative format
Antecedents	12	2.1 (0–5)	2.7 (0–6)
Ignored by someone/someone walked away	2	.5 (0–2)	.3 (0–1)
Leisure material/food removed or denied	2	.7 (0–2)	.7 (0–1)
Instruction/prompt to work provided	8	.9 (0–3)	1.6 (0–6)
Consequences	18	3.5 (0–8)	8.1 (3–11)
Attention	9	1.6 (0–6)	3.8 (0–7)
Access to materials	3	.9 (0–3)	1.3 (0–3)
Work removed or delayed	6	.9 (0–4)	2.9 (0–6)

Additional details about the participants' errors are shown in Tables 2 and 3. Errors of omission (i.e., the participant did not score an event that occurred) and errors of commission (i.e., the participant scored an event that did not occur) were examined across the various antecedents and consequences. The tables show the mean number (and ranges) of these errors, along with the number of true instances, for each of the antecedents and consequences associated with the three social functions (attention, tangibles, and escape). Results are displayed separately for the two recording formats. More omission errors (Table 2) generally occurred with the narrative recording form than with the structured form for both antecedents and consequences, but the differences are most apparent for the antecedent related to escape ("instruction/prompt to work provided") and for the consequences related to attention and escape. A similar pattern occurred for commission errors (Table 3). These findings, however, could be a function of differences in the overall number of true instances among these events.

Mean acceptability ratings were slightly higher for the structured format (24.3) than for the narrative format (21.5) across participants. Twelve of the educators not only rated the structured format higher than the narrative format but also indicated a preference for the structured format (indicated that they would use the structured

Table 3 Mean number and range of commission errors associated with the two recording formats

Event	True instances	Structured assessment	Narrative assessment
Antecedents	12	1.0 (0–2)	1.9 (0–1)
Ignored by someone/someone walked away	2	.1 (0–1)	.4 (0–1)
Leisure material/food removed or denied	2	.5 (0–1)	.4 (0–3)
Instruction/prompt to work provided	8	.4 (0–1)	1.1 (0–5)
Consequences	18	.8 (0–3)	1.3 (0–4)
Attention	9	.1 (0–1)	.2 (0–1)
Access to materials	3	.3 (0–3)	.2 (0–1)
Work removed or delayed	6	.5 (0–2)	.9 (0–2)

format if they could only choose one.). Of the four participants who preferred the narrative format, two showed higher accuracy with the narrative format than with the structured format, and the other two showed no difference in accuracy.

Discussion

In this study, certified special education teachers and paraprofessionals were asked to collect descriptive analysis data using both narrative A-B-C and structured A-B-C recordings under relatively "ideal" and highly controlled conditions. The scripted exchanges were designed to reflect the level of complexity that is often present in naturalistic settings, such as classrooms. The structured A-B-C format was associated with slightly higher levels of accuracy, produced higher acceptability ratings, and was preferred by the majority of teachers when compared to the narrative A-B-C format. Nonetheless, results showed only modest levels of accuracy with either format even though the teachers were in a quiet room with no other responsibilities or distractions while collecting data. The participants scored the videos immediately after receiving an in-service on descriptive analysis recording that was designed to simulate "training as usual" in public schools. Furthermore, 13 of the 16 participants had reported prior experience with A-B-C recording. These finding suggest that psychologists and behavior specialists should use caution when relying on teacher-collected data, even among teachers who report prior experience with A-B-C recording.

Implications for Research and Practice

Although the study should be considered preliminary, results have a number of implications for research and practice. First, the teachers in this study would have benefited from further training and experience. Although 13 of the 16 participants indicated prior experience with either narrative or structured A-B-C recording, the extent of this experience and the type of training received prior to the study are unknown. (It should be noted that participants reporting prior experience did not have higher levels of accuracy than those reporting no prior experience.) All of the participants received a 60-min lecture that described common antecedents and consequences of problem behavior and instructions on collecting, summarizing, and interpreting descriptive analysis data. This in-service was intended to reflect the typical amount of training that a school psychologist might provide to a staff member or teacher. Further research is needed to determine the type and extent of training that is sufficient to produce accurate descriptive analysis data. For now, it is recommended that psychologists and behavior specialists provide more extensive training before asking educators to collect descriptive analysis data. Training should include practice with feedback, followed by periodic reliability checks and additional training if needed.

Second, conclusions about the accuracy of teacher-collected descriptive data may depend on the method used to calculate accuracy. When accuracy was based on a

percentage of overall occurrences of antecedents and consequences, the teachers showed moderate levels of accuracy across the two formats. Accuracy based on nonoccurrences revealed near perfect data collection. However, their accuracy dropped significantly when the percentage of problem behavior with accurately recorded antecedents and consequences was calculated. Such low levels of accuracy may lead to incorrect conclusions about possible functions of problem behavior. Thus, it is recommended that school psychologists and behavior specialists use the most conservative measure of accuracy (based on percentage of problem behavior) when evaluating the accuracy of teacher-collected data. However, further research is needed to identify the minimum level of accuracy needed to correctly identify function from descriptive data.

Third, results of the study suggest that teachers may generally prefer the structured A-B-C format and that they may record data more accurately when using it. However, another important outcome that cannot be determined from the study is whether teachers would be more likely to collect data consistently when using one format versus the other. Further research should address this question because it may be advantageous to use the format that is associated with the most consistent data collection.

Fourth, results tentatively indicate that prior training and experience with structured A-B-C recording can improve the accuracy of narrative A-B-C recording. Five of the seven participants who showed similar or higher levels of accuracy when using the narrative format compared to the structured format had used the structured form to score a video prior to using the narrative form. Individuals who have familiarized themselves with a list of potentially relevant antecedents and consequences or who have acquired the skills needed to instantly classify events may find it easier to determine what events to record or to record events more objectively when using narrative A-B-C recording. Thus, school psychologists who would prefer to receive narrative recordings may find it beneficial to provide teachers and caregivers with training and practice in both formats.

Finally, the extent to which these findings have generality to data collected in actual classrooms should be investigated in further research. The video simulations contained scenarios that the experimenters had commonly observed in classrooms and included both simple and complex relations between problem behavior and environmental events. Nonetheless, the generality of the results may be limited to settings with similarly complex relations, to behaviors that occur approximately once per minute, and to teachers who are not collecting data on their own reactions to problem behavior.

References

Chandler, L. K., & Dahlquist, C. M. (2006). *Functional assessment: Strategies to prevent and remediate challenging behavior in school settings* (2nd ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Ellingson, S. A., Miltenberger, R. G., Stricker, J., Galensky, T. L., & Garlinghouse, M. (2000). Functional assessment and intervention for challenging behaviors in the classroom by general classroom teachers. *Journal of Positive Behavior Interventions, 2*, 85–97. doi:10.1177/109830070000200202.

Grey, I. M., Honan, R., McClean, B., & Daly, M. (2005). Evaluating the effectiveness of teacher training in applied behavioural analysis. *Journal of Intellectual Disabilities, 9*, 209–227. doi:10.1177/1744629505056695.

Hanley, G. P., Iwata, B. A., & McCord, B. E. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis, 36*, 147–185. doi:10.1901/jaba.2003.36-147.

Kamps, D., Wendland, M., & Culpepper, M. (2006). Active teacher participation in functional behavior assessment for students with emotional and behavioral disorders risks in general education classrooms. *Behavioral Disorders, 31*, 128–146.

Lerman, D. C., & Iwata, B. A. (1993). Descriptive and experimental analysis of variables maintaining self-injurious behavior. *Journal of Applied Behavior Analysis, 26*, 293–319. doi:10.1901/jaba.1993.26-293.

Maag, J. W., & Larsen, P. J. (2004). Training a general education teacher to apply functional assessment. *Education & Treatment of Children, 27*, 26–36.

National Research Council. (2001). *Educating children with autism*. Washington, DC: National Academy Press.

O'Neill, R. E., Horner, R. H., Albin, R. W., Sprague, J. R., Storey, K., & Newton, J. S. (1997). *Functional assessment and program development for problem behavior: A practical handbook*. Pacific Grove, CA: Brooks/Cole Publishing Company.

Scheuermann, B., Webber, J., Boutot, A., & Goodwin, M. (2003). Problems with personnel preparation in autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities, 18*, 197–206. doi:10.1177/10883576030180030801.

Smith, R. G., Vollmer, T. R., & St. Peter Pipkin, C. (2007). Functional approaches to assessment and treatment of problem behavior in persons with autism and related disabilities. In P. Sturmey & A. Fitzer (Eds.), *Autism spectrum disorders: Applied behavior analysis, evidence, and practice* (pp. 187–234). Austin, TX: PRO-ED.

Study of Personnel Needs in Special Education. (2002). *A high quality teacher for every classroom*. Retrieved from www.spense.org/results.html. Accessed 28 October 2008.

Thompson, R. H., & Iwata, B. A. (2007). A comparison of outcomes from descriptive and functional analyses of problem behavior. *Journal of Applied Behavior Analysis, 40*, 333–338. doi:10.1901/jaba.2007.56-06.

Tiger, J. H., Hanley, G. P., & Bessett, K. K. (2006). Incorporating descriptive assessment results into the design of a functional analysis: A case example involving a preschooler's hand mouthing. *Education & Treatment of Children, 29*, 107–124.

Watson, T. S., Ray, K. P., Sterling Turner, H., & Logan, P. (1999). Teacher-implemented functional analysis and treatment: A method for linking assessment to intervention. *School Psychology Review, 28*, 292–302.