Article

Hammill Institute on Disabilities

Behavioral Disorders
2018, Vol. 44(1) 53–62

DOI: 10.1177/0198742918779793
journals.sagepub.com/home/bhd
SAGE

Brief Report: Measurement of Young Children's Engagement and Problem Behavior in Early Childhood Settings

Debra A. Prykanowski, MA[1], Jose R. Martinez, PhD[1,2], Brian Reichow, PhD[1], Maureen A. Conroy, PhD[1], and Ke Huang, MA[1]

Abstract

Given the importance of engagement and impact of problem behavior on young children's behavioral, social, and academic development, these two behaviors are often observed and measured in early childhood settings. The purpose of this brief report is to describe the findings from an examination of the use of multiple direct observation methods to measure engagement and problem behaviors of five children in early childhood settings across two different activity types. Findings from this study support previous research that young children's engagement and problem behavior differ depending on the type of classroom activity. In addition, the data indicate that behavioral estimates for both engagement and problem behavior obtained using momentary time sampling produced measures of behaviors comparable with those obtained with continuous recording, whereas partial-interval recording consistently overestimated the occurrence of the behaviors compared with continuous recording. Implications of the findings and directions for future research are discussed.

Keywords

measurement, engagement, problem behavior, early childhood, time sampling, interval recording, behavioral observation

Young children's engagement in academic and social activities within early childhood settings contributes to academic achievement as well as appropriate social and emotional development (Hojnoski & Missall, 2010; Williford, Vick Whittaker, Vitiello, & Downer, 2013). Engagement, defined as the degree to which children are attentive and interactive in a consistent manner with academic or social activities (Fredricks, Blumenfeld, & Paris, 2004), is associated with critical developmental constructs (e.g., development of appropriate emotional regulation) that are vital for academic achievement and later life success (Hojnoski & Missall, 2010; Williford et al., 2013). Lack of engagement in early childhood is associated with poor academic performance and development of chronic problem behaviors (Hojnoski & Missall, 2010). Chronic problem behaviors in early childhood have been associated with the development of negative relationships with teachers and peers, which further increases the risk of young children developing behavioral, social, and academic challenges (Hughes & Kwok, 2006).

Engagement and Problem Behavior in Early Childhood Settings

Engagement in a variety of classroom activities within early childhood settings requires children to demonstrate different behaviors and offers diverse opportunities to practice and develop skills (Booren, Downer, & Vitiello, 2012; Rimm-Kaufman, La Paro, Downer, & Pianta, 2005). Although the use of different classroom activities allows teachers to intentionally structure young children's time in the learning environment, it is important to identify how specific situational demands in different activities might be related to young children's engagement (Booren et al., 2012). To that extent, teacher-directed (TD) activities might require more passive engagement (e.g., listening to the teacher) from children whereas child-initiated (CI) activities might require more active and independent engagement in various activities (e.g., engaging with a preferred activity as opposed to listening to teacher instructions). Rimm-Kaufman et al. (2005) found that whereas children were more likely to be engaged during TD activities, aggression was also more likely to occur in TD activities. Alternatively, Booren et al. (2012) found lower engagement in TD activities (e.g., circle time) when compared with CI activities (e.g., centers, playground time). These authors

[1]University of Florida, Gainesville, USA
[2]The University of Kansas, Lawrence, USA

Corresponding Author:
Debra A. Prykanowski, Anita Zucker Center for Excellence in Early Childhood Studies, University of Florida, Norman Hall, P.O. Box 117050, Gainesville, FL 32611-7011, USA.
Email's: dprykanowski@ufl.edu, debprykanowski@gmail.com

hypothesized that the high demands for sustained attention during more structured activities (i.e., TD activities) might contribute to the low levels of engagement.

Although limited research exists on the impact of classroom activity on problem behaviors specific to early childhood settings, there is literature suggesting that activity types might also be an important contextual factor linked to the occurrence of problem behavior (Dunlap & Kern, 1996). Ferro, Foster-Johnson, and Dunlap (1996) found a relation between activities that were highly preferred and low levels of problem behaviors in children with intellectual disabilities. Other research suggests that specific elements of classroom activity that would be found in TD activities such as task difficulty and length of task might contribute to the occurrence of problem behavior (Dunlap & Kern, 1996). Moreover, elements of classroom activity that would be found in CI activities such as preference and choice, indicate an association with low levels of problem behavior (Dunlap & Kern, 1996). Thus, further examination of the impact of classroom activity on the occurrence of problem behavior, especially within early childhood settings, is warranted.

Use of Direct Behavioral Observation to Measure Child Behaviors

The use of direct observations in educational settings has been an essential component for progress monitoring and treatment evaluation in special education research (Lewis, Scott, Wehby, & Wills, 2014). Therefore, it is critical that measurement of target behaviors in these settings produce accurate portrayals (estimates) of the behaviors. Researchers and practitioners need to be aware of and address any potential measurement issues that might influence the outcomes of their observations. If accurate and reliable measurement does not occur, there is potential for inaccurate conclusions to be reached with the collected data (Lane & Ledford, 2014). Moreover, if measured accurately, researchers and practitioners can have more confidence in their findings and thus implement specific interventions based on the observation to improve contexts (Booren et al., 2012).

Many factors, including the type of direct observation method used, the settings in which behaviors are observed, and the topography of the behavior, influence the accuracy and reliability of the measurement of children's behaviors in authentic educational settings. The observation method chosen should consider the dimensions (e.g., frequency or duration) of the behavior being measured to ensure the sample obtained represents an estimate that is "true" to the actual occurrence of the behavior (Lewis et al., 2014). In addition, researchers and practitioners may need to consider the practicality of different observation methods to ensure that the method is feasible given the context (Wood, Hojnoski, Laracy, & Olson, 2015).

Continuous recording (CR) and time sampling recording (i.e., partial-interval recording [PIR], whole-interval recording [WIR], and momentary time sampling [MTS]) are common direct observation methods used in the literature to quantify behavior, monitor children's progress, and evaluate intervention outcomes (Lane & Ledford, 2014; Yoder & Symons, 2010). One benefit of CR is that this measure can capture both the frequency and duration of behaviors, whereas time sampling recording methods only provide an estimate of behaviors. Alternatively, time sampling (e.g., MTS) methods may be more practical for early childhood researchers and practitioners because they do not require continuous observations of the children (Lane & Ledford, 2014; Wood et al., 2015). Previous research has compared estimates of behaviors obtained using time sampling methods to measures of behaviors obtained using CR (for a review, see Lane & Ledford, 2014). For example, Wood and colleagues (2015) compared estimates of young children's engagement in early childhood settings obtained using WIR, PIR, and MTS to measures of the behavior obtained using CR. The researchers found that PIR overestimated the duration of young children's engagement, whereas WIR consistently underestimated the duration of the behavior. Wood and colleagues (2015) concluded that estimates of engagement obtained with MTS were more representative than the estimates obtained with either PIR or WIR when compared with CR. Although not all specific to early childhood, Wood and colleagues' (2015) findings are consistent with the findings of other studies that have used similar procedures to compare estimates of behaviors obtained with time sampling methods to measures obtained with CR (e.g., Devine, Rapp, Testa, Henrickson, & Schnerch, 2011; Mann, Have, Plunkett, & Meisels, 1991). Despite the consistency of the findings, many of these studies used simulated data and researchers continue to suggest the need for additional comparison studies when measuring young children's engagement in natural early childhood settings (e.g., Wood et al., 2015).

To date, the Wood et al. (2015) study is the only study that has compared time sampling methods to CR when measuring young children's engagement in authentic early childhood settings (i.e., classrooms) and, to our knowledge, no study has compared direct observation methods when measuring young children's problem behaviors within these settings. Moreover, although research has shown differences in young children's engagement and problem behaviors across different classroom activities, Wood et al. (2015) only examined children's engagement in TD activities. Therefore, the purpose of this study is twofold: (a) to examine differences in engagement and problem behaviors of young children in early childhood settings across two different classroom activities; and (b) to examine differences between the measures of engagement and problem behaviors obtained using CR and the estimates of those behaviors obtained using time sampling recording methods. The

Table 1. Description of Children Characteristics.

Child	Age (months)	Gender	Race	C-TRF (*T* scores)	
				Total problem behavior	Externalizing problem behavior
1	60	Male	African American	70 (Clinical)	70 (Clinical)
2	63	Male	African American	61 (Borderline)	62 (Borderline)
3	60	Female	African American	54	60 (Borderline)
4	53	Male	African American	49	55
5	52	Male	African American	61 (Borderline)	60 (Borderline)

Note. Scores are reported as standard scores. A standard score ranging from 60 to 63 is considered borderline and ⩾64 is considered clinical. C-TRF = Caregiver-Teacher Rating Form (Achenbach & Rescorla, 2000).

specific research questions addressed in this study were as follows:

Research Question 1: Are there differences in young children's behaviors (i.e., engagement and problem behavior) based on the type of activity (i.e., TD versus CI) observed?

Research Question 2: How do estimates of young children's engagement and problem behaviors obtained using different time sampling methods (i.e., PIR and MTS) compare with measures of those behaviors obtained using CR?

Method

Participants

Four males and one female between the ages of 4 and 5 years, systematically identified as at risk for developing emotional and behavioral disorders (EBD), participated in this study. All children were enrolled in Head Start classrooms in the southeastern United States and had previously passed a screening test indicating that they did not have any developmental or intellectual disabilities. The children were part of a larger study examining the efficacy of a Tier 2 classroom-based intervention designed to ameliorate problem behavior of young children at risk for EBD. Children were identified as at risk for EBD using the *Early Screening Project* (Feil, Severson, & Walker, 1998), which is a teacher report designed to measure the frequency and intensity of young children's internalizing and externalizing behaviors. Scores from the *Caregiver-Teacher Reporting Form* (C-TRF; Achenbach & Rescorla, 2000) were used to report information about the children's levels of externalizing problem behaviors. A summary of the children's descriptive information, as well as their C-TRF scores prior to the larger intervention study, is shown in Table 1.

Data Source

As part of the larger study, children were video recorded in their classrooms across two types of activities (i.e., CI and TD). Video recording occurred 1 day a week over the course of 5 weeks. Due to the varying nature of early childhood classrooms and schedule of activities, observations were different lengths in each type of activity. To ensure comparisons across children were conducted based on similar total lengths of observations and that video recordings were chosen from across all possible observation days, video clips were randomly selected until each child had a total range of observations lasting a total of 2,500 to 2,800 s selected. Randomization was achieved by recording the duration of each video clip and assigning the clip a number. A random number generator was used to select the video clips to include in the study until the desired duration was achieved for each child.

A total of 43 video clips across the two target activities and children were used to code the target behaviors using CR. The number of video clips per child ranged from three to six for TD activities (M=4.0) and three to six for CI activities (M=4.6). Each video clip averaged 600 s (range = 320 s–1,235 s). The frequency and duration of the target behaviors were recorded using CR from the videos using Lily Collector (Tapp, 2010), a computerized system designed to collect behavioral data. Once coded, data were transferred from the Lily file to an Excel® spreadsheet to conduct a second-by-second analysis.

Data Collection

Operational definitions of the target activities and behaviors are provided in Table 2. Data were collected on the two target behaviors: engagement and problem behavior. In addition, two codes, nonengagement and no opportunity, were used to ensure a second-by-second record of the full video was created. The second-by-second record was necessary to divide the observation into intervals to calculate PIR and MTS measures. These codes were used if the child demonstrated a behavior that did not fit the definition for the target behaviors, or if due to video recording the behaviors could not be accurately observed (e.g., a child left the video frame). The average duration per occurrence across children, behaviors, and type of activity is provided in Table 3.

Table 2. Operational Definitions of Activities and Behaviors.

Activity/behavior	Operational definition	Example
Activity		
TD activity	Any activity led and directed by a teacher in which a child and at least one other child took part. TD activities have common expectations for engagement (e.g., classroom rules, remain seated, look at the speaker) for all participating children, and children have little to no choice in the activity.	TD activities were either whole group or small group and included, but were not limited to, circle time, story time, tabletop art activity, and games directed by teachers.
CI activity	Any activity where a child is free to direct his or her own engagement. Teachers may participate in the activity, but the child decides how the activity should occur.	CI activities may have been with other peers, adults, or alone. These included, but were not limited to, center time, free play, and recess.
Target behavior		
Engagement (adapted from Wood, Hojnoski, Laracy, & Olson, 2015)	A child exhibiting motor, verbal, and/or attending behaviors that are consistent with all expectations of the ongoing activity and promote participation.	Verbally responding to a teacher's questions, complying with a teacher's request, orienting toward a teacher/peer/ongoing activity, making eye contact with a teacher/peer/ongoing activity, manipulating and/or playing appropriately with a toy, and initiating and/or responding during a social interaction with a peer or adult.
Problem behavior (adapted from Conroy et al., 2014)	A child exhibiting any behavior that is an act of disruption (i.e., motor/verbal behavior that disrupts/interrupts or has the potential to disrupt/interrupt classroom instruction), defiance (i.e., motor/verbal behavior that actively challenges or passively ignores teacher requests/demands), or aggression (i.e., motor/verbal behavior aimed at causing harm, pain, or injury to others).	Purposely ignoring a teacher's request, not following classroom rules, and hitting a peer.
Additional codes		
No opportunity	A child could not be seen or the behaviors of a child could not be accurately coded due to the position of the child in the video.	Child goes to the restroom or child leaves the video frame.
Nonengagement	Coded when a child demonstrated behaviors that were not consistent with the above definitions of engagement, problem behavior, or no opportunity.	Child is turned away from the activity, but does not demonstrate a behavior consistent with the problem behavior definition.

Note. TD = teacher-directed; CI = child-initiated.

Continuous Recording. Coding of each target behavior began at its onset by pressing the corresponding code key of the target behavior immediately when the behavior began. The target behaviors were mutually exclusive when using CR as only one target behavior could be exhibited at a time. Using the Excel file of the second-by-second record for CR per video clip, frequency was defined as the number of times a behavior onset was coded. The duration of each instance was determined by calculating the number of seconds between a coded behavior's onset and offset.

Time sampling. Based on previous literature (Lane & Ledford, 2014; Powell, Martindale, & Kulp, 1975), we chose to compare PIR and MTS using a 10-s interval with CR. Time sampling files were created by dividing the second-by-second record of each CR recording into 10 s intervals. For PIR, the first author examined each 10-s interval and recorded the first instance of each coded behavior within those 10 s. All subsequent occurrences of the same behavior within the 10-s interval were not coded as to capture only one instance of each target behavior per interval. The PIR system is not mutually exclusive as more than one behavior could be coded per interval. For MTS, we examined the 10th second of each interval and recorded only the behavior that was occurring at that moment. Due to the nature of using MTS coding, the behavioral codes were mutually exclusive; thus, only engagement or problem behavior could be observed at the moment of observation.

Table 3. Duration per Occurrence of Behavior Across Activities per Child.

		Teacher-directed					Child-initiated				
Child	Bx	Total duration of video coded[a]	No. of video clips	Frequency	Duration[a] *M* (*SD*)	Range[a]	Total duration of video coded[a]	No. of video clips	Frequency	Duration[a] *M* (*SD*)	Range[a]
1	ENG	2,874	4	56	19.1 (16.3)	1.0–65.0	2,796	6	52	38.6 (50.6)	2.0–226.0
	PB			48	32.5 (38.9)	2.0–190.0			28	14.1 (19.6)	2.0–98.0
2	ENG	2,780	5	41	49.0 (66.6)	3.0–394.0	2,798	4	39	64.3 (74.3)	1.0–381.0
	PB			12	7.3 (7.7)	2.0–27.0			11	5.5 (3.8)	1.0–14.0
3	ENG	2,737	4	55	30.2 (49.7)	1.0–332.0	2,714	6	56	38.0 (40.0)	1.0–169.0
	PB			48	17.4 (22.0)	1.0–114.0			26	13.4 (19.0)	1.0–78.0
4	ENG	2,710	3	47	46.3 (61.6)	1.0–221.0	2,525	3	36	71.3 (82.0)	2.0–303.0
	PB			33	11.5 (13.3)	1.0–63.0			20	17.5 (23.6)	1.0–76.0
5	ENG	2,525	4	49	27.2 (31.8)	1.0–197.0	2,515	4	36	57.7 (54.5)	2.0–186.0
	PB			30	30.4 (36.5)	2.0–133.0			22	17.18 (23.6)	3.0–106.0
M (*SD*)	ENG		4	49.6 (6.1)	34.36 (12.8)			4.6	43.8 (9.5)	53.98 (15.1)	
	PB			34.2 (14.9)	19.8 (11.2)				21.4 (6.6)	13.5 (4.8)	

Note. Bx = behavior; ENG = engagement; PB = problem behavior.
[a]Measured in seconds.

Table 4. Percentage of Difference Between Engagement and Problem Behavior Across Activities.

	Engagement			Problem behavior						
	CR (% duration)			CR (% duration)			No opportunity (CR % duration)		Nonengagement (CR % duration)	
Child	TD	CI	% difference	TD	CI	% difference	TD	CI	TD	CI
1	35.66	72.78	51.00	49.06	14.09	71.28	5.92	12.63	9.36	0.50
2	72.30	89.96	19.63	3.20	2.14	33.13	20.65	7.15	3.85	0.75
3	60.76	78.33	22.43	30.51	12.86	57.85	6.10	7.70	2.63	1.11
4	80.26	86.72	7.45	14.02	12.11	13.62	1.00	0.86	4.72	0.17
5	52.83	82.54	35.99	36.12	15.03	58.39	3.80	8.63	7.25	0.00
M	60.36	82.07	26.45	26.58	11.25	57.67	7.49	7.39	5.56	0.51
SD	17.36	6.79	16.69	18.14	5.21	23.15	7.64	4.24	2.72	0.45

Note. CR = continuous recording; TD = teacher-directed; CI = child-initiated.

Interobserver Agreement

Data collectors were doctoral students in special education with previous experience in behavioral observation coding. Prior to coding videos, a second data collector (i.e., fifth author) was trained to agreement in coding the target behaviors using the CR procedures. A criterion of at least 80% agreement on all target behaviors across three criterion coded standard videos was required for independent coding. Criterion coded videos were developed by the first and second author and disagreements were discussed to reach a consensus. Interobserver agreement (IOA) was calculated for 30% of randomly selected observations. IOA for CR was calculated using a point-by-point time window agreement check (Yoder & Symons, 2010). A second-by-second analysis of the session was conducted. An agreement was defined as the identification of the onset and offset of a target behavior by the two data collectors within 2 s of each other. IOA was calculated by dividing agreements by the sum of agreements and disagreements and multiplying by 100. Mean overall IOA was 96% (range = 83%–100%). Mean IOA for engagement was 97% (range = 75%–100%), 87% (range = 60%–100%) for problem behavior, 81% (range = 71%–100%) for nonengagement, and 90% (range = 75%–100%) for no opportunity. The first author and the fourth author, who collected the data, discussed disagreements to reach a consensus. In addition, IOA was calculated for PIR and MTS measures. A second data collector counted the number of intervals where a behavior occurred (i.e., for MTS, the behavior that was recorded on the 10th second and for PIR, the behaviors recording within the 10-s interval). These data were compared with the first authors' calculations. Interrater agreement for 30% of the PIR and MTS calculations were both 100%.

Data Analysis

The percentage of each target behavior was calculated for each observation method. For CR, percentage of behavior was calculated by dividing the number of seconds the behavior was recorded by the total number of seconds in the observation and multiplying by 100. For PIR and MTS, percentage of intervals for each target behavior were calculated by dividing the number of intervals the behavior was recorded by the total number of intervals in the observation and multiplying by 100. For the purpose of this study, the measures of behaviors obtained with CR were considered the most accurate measurements of the duration of the behaviors and thus, the data obtained with the time sampling methods were compared with the data of behaviors obtained with CR.

Data analyses were conducted for the first research question by examining the occurrences of behaviors obtained using CR in each activity. To determine the extent to which behaviors recorded in CI activities differed from those in TD activities, the percentage difference of the occurrences obtained in CI relative to the measure obtained in TD between the two activities was calculated for both behaviors (Alvero, Struss, & Rappaport, 2008; Wood et al., 2015). For each participant, the percentage of each behavior in TD activities was subtracted from the percentage of behavior in CI activities and divided by the larger value of the two and multiplied by 100. Mean percentage difference across children was calculated. Descriptive comparisons of the percentage differences were conducted following the completion of the calculations. To answer the second research question, the extent to which estimates of each target behavior obtained using PIR and MTS differed (i.e., measurement error) from CR was calculated by subtracting the estimates of each target behavior using PIR and MTS from the duration percentage obtained using CR (Wood et al., 2015).

Results

A comparison of engagement and problem behavior measured using CR for each focal child in each type of activity is shown in Table 4. Across all focal children, the mean

Table 5. Extent to Which Measurements of Engagement and Problem Behavior Using Time-Sampling Methods Differed From CR Within Two Early Childhood Activities.

	Teacher-directed			Child-initiated		
Child	CR (%)	PIR (difference from CR)	MTS (difference from CR)	CR (%)	PIR (difference from CR)	MTS (difference from CR)
Engagement						
1	35.66	12.94	0.00	72.78	11.70	−1.30
2	72.30	7.2	−0.36	89.96	7.89	−0.35
3	60.76	13.41	−0.24	78.33	13.06	0.70
4	80.26	11.56	1.52	86.72	5.08	0.39
5	52.83	12.91	−1.44	82.54	10.23	−0.61
M	60.36	11.60	−0.10	82.07	9.59	−0.23
SD	17.36	2.56	1.06	6.79	3.17	0.80
Problem behavior						
1	49.06	10.38	−0.46	14.09	10.10	1.07
2	3.20	3.27	−0.32	2.14	2.88	0.01
3	30.51	13.4	−0.01	12.86	7.74	0.62
4	14.02	6.43	−1.75	12.11	6.25	0.39
5	36.12	8.9	−0.87	15.03	7.86	1.44
M	26.58	8.48	−0.68	11.25	6.97	0.71
SD	18.13	3.85	0.67	5.21	2.67	0.56

Note. CR = continuous recording; PIR = partial-interval recording, MTS = momentary time sampling.

percentage of engagement was 60.36% (*SD* = 17.36%) in TD activities and 82.07% (*SD* = 6.79%) in CI activities. The mean percentage of problem behavior was 26.58% (*SD* = 5.21%) in TD activities and 11.25% (*SD* = 5.21%) in CI activities. The percent of difference was examined to determine the difference in measured behaviors in CI activities relative to TD. On average, engagement was 27.30% (*SD* = 16.69%) higher when children were in CI activities compared with TD activities. Moreover, problem behavior was, on average, 46.85% (*SD* = 23.15%) lower in CI activities when compared with TD activities. The higher levels of engagement in CI activities and lower problem behavior during TD activities were found using the time sampling methods as well. These results are consistent with previous studies (Booren et al., 2012; Ferro et al., 1996; Rimm-Kaufman et al., 2005) that found differences in children's behaviors depending on the type of classroom activity they are engaged in, and that children tend to demonstrate more engagement during activities where they are free to direct their attention to materials of their choice. Moreover, similar to findings of Ferro et al. (1996), children in this current study demonstrated fewer occurrences of problem behavior in activities where they could engage with materials of high preference (i.e., CI activities) within an early childhood setting.

Although the percentage of observed engagement across the five focal children in these activities had some variability (range = 35.7%–80.3% for TD activities; range = 72.8%–90.0% for CI activities), conclusions about how the time sampling methods compare with CR can be made. Table 5 shows the mean error calculated from estimates obtained from PIR and MTS compared with CR. In TD activities, the intervals of engagement measured by PIR consistently overestimated engagement measured with CR for all focal children (*M* = 11.60%; *SD* = 2.56%), whereas the difference in percentage for engagement obtained with MTS was much closer to results obtained through CR (*M* = –0.10%; *SD* = 1.06%). When using PIR during CI activities, the mean difference in percentages for measures of engagement from CR was 9.59% (*SD* = 3.17%), whereas using MTS the mean difference in percentages was –0.23% (*SD* = 0.80%). Similar to TD activities, PIR consistently overestimated children's engagement, but MTS produced estimates much closer to those obtained using CR.

Similar results for comparisons of estimates of problem behavior were found across observation methods and activities. Overall, the five focal children engaged in variable durations of problem behavior (range = 3.20%–49.06% for TD activities; range = 2.14%–15.03% for CI activities). In TD activities, PIR resulted in a mean difference in percentage of problem behavior from CR of 8.48% (*SD* = 3.85%), whereas MTS resulted in a mean difference in percentage of –0.68% (*SD* = 0.67%). When estimating problem behavior in CI activities, PIR estimates resulted in a mean difference in percentage from CR of 6.97% (*SD* = 2.67%), whereas MTS resulted in a mean difference in percentage of 0.71% (*SD* = 0.56%). Consistent with the results of the measurement of engagement, PIR always overestimated when compared with CR. Session by session data show that although MTS might over- or underestimate across sessions, MTS

produced estimates that were similar (within 1%–2%) of the measure obtained using CR.

Discussion

This study replicates and extends the current measurement literature of early childhood settings in several ways. First, this study replicated the findings of Wood et al. (2015) as well as the findings of other measurement studies (Devine et al., 2011; Mann et al., 1991) that compared time sampling methods with CR. Findings were consistent in that estimates of the duration of the behaviors obtained with MTS were comparable to the measures obtained with CR, whereas those estimates obtained with PIR consistently overestimated the duration of the behaviors. In the current study, PIR overestimated the duration of both behaviors in each type of activity and thus, the overestimation does not seem to be dependent on the type of behavior or activity type. Although Lane and Ledford (2014) concluded that MTS was appropriate for low- or high-rate behaviors of long durations, this study showed that the two methods performed the similarly even across variable durations of behaviors. For example, although problem behavior measures for one focal child ranged from 49% in TD activities to 14% in CI activities, estimates of the duration of both behaviors obtained with MTS were comparable to the measures of both behaviors obtained with CR. Similarly, the estimates of the duration of both behaviors obtained with PIR overestimated the measures of the behaviors obtained with CR.

Second, this study extended the Wood et al. (2015) study by measuring both engagement and problem behavior across two different early childhood activities. Mann et al. (1991) found that short-duration behaviors were inaccurately measured using time sampling methods. Examination of the session-by-session data showed that although the duration of each instance of the target behaviors varied, MTS continued to produce near similar estimates of the duration of shorter behaviors (e.g., problem behavior; TD: M = 19.8 s [11.2]; CI: M = 13.5 s [4.8]) when compared with actual duration as measured through CR in most of the children. This study also supports that certain time sampling methods such as PIR may produce an inflated estimate of occurrence of target behaviors, giving inaccurate information about the duration of engagement or problem behavior demonstrated by a young child. The implication is that if PIR is used to estimate the duration of these behaviors, more research is needed to determine adjustments that should be made for the error (Lane & Ledford, 2014; Yoder, Ledford, Harbison, & Tapp, 2017). Regardless of the observation method used, the current study supports previous findings that indicate that young children exhibit different levels engagement and problem behaviors depending on the type of activity in which they are participating. This finding indicates that further research could be conducted on the aspects of classroom activity that impact the occurrence of these target behaviors.

Implications for Research and Practice

Although interval recording methods might be more convenient to conduct, especially in natural classroom settings, researchers should note that inflated estimates of a behavior might be obtained when using PIR. This study has shown that across two common behaviors observed in early childhood settings (i.e., engagement and problem behavior), researchers interested in using PIR should be aware that the estimates they obtain will almost certainly be higher than the actual occurrence of the behaviors. For behaviors that are targeted for decrease (e.g., problem behaviors), an overestimate of the behavior might not be as problematic because any reduction of this behavior would be considered a positive result. Conversely, estimates of behaviors targeted for increase (e.g., engagement) might present an inaccurate picture of the actual occurrence of the behaviors. As such, the use of MTS as a behavioral observation method is recommended, as the estimates of behaviors obtained with it are much closer to the actual occurrence of the behaviors. Regardless, those using interval recording systems should acknowledge that the measure of behavior reported is an estimate, and not an exact account of the actual occurrence of behavior.

Limitations

It is important to acknowledge this study's limitations. First, the definition of engagement broadly encompassed both active and passive engagement and thus, differentiation could not be made between these two types of engagement. Future research may provide separate definitions for each type of engagement. Furthermore, the definition of each type of activity used in this study was based on a larger study where a TD activity required the presence of a peer, whereas the CI activity did not. The presence of a peer may influence both the occurrence of engagement and problem behavior exhibited by a child. To that end, a review of the videos included in this analysis showed that with the exception of one video clip, all CI videos included the focal child and at least one peer. Therefore, for both TD and CI activities, children spent the overwhelming majority of time in activities with peers. Second, the current study examined two common early childhood activities but not the full array of activities that occur with regular frequency in early childhood settings. By not examining other types of activities important differences and distinctions that might help place the current findings in a broader context may have been missed. Third, the definitions of the interval coding systems might contribute to the differences in the results. More specifically, as PIR

is not mutually exclusive (i.e., more than one behavior can be coded per interval) this has the potential to artificially increase the likelihood that PIR would provide an overestimate of occurrence. Next, as the observations were recorded over a short period of time (i.e., 5 weeks), we are uncertain if additional observations for a longer period of time would have resulted in the same findings, thus the generalizability of the outcomes should be taken with caution. Finally, the observation method used in this study only accounted for the behaviors of the focal children. Research has shown that teachers' behaviors could impact levels of children's engagement (e.g., McWilliam, Scarborough, & Kim, 2003) and thus, future research may incorporate teachers' behaviors into the analyses.

Conclusion

This study supports that there are observable levels of differences in engagement and problem behavior for young children across different early childhood activities. The differences were notable in that higher estimates of duration of engagement and lower estimates of duration of problem behavior were obtained in CI activities when compared with TD activities consistently across all children. This study also supports previous research (Devine et al., 2011; Mann et al., 1991; Wood et al., 2015) that found PIR consistently overestimates occurrences of behavior, whereas MTS produces estimates closer to CR. Moreover, this study supports the conclusions of the Wood et al. (2015) study that MTS might be a viable option for obtaining estimates of engagement and problem behaviors in early childhood settings.

Acknowledgments

The authors would like to thank and acknowledge contributions from the following individuals in the completion of this study: Dr. James Haile, Dr. Mark Lewis, Dr. Patricia Snyder, Dr. Timothy Vollmer, and Dr. Brenna Wood.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported partially by funding from the U.S. Department of Education, Institute of Education Sciences (R305A40487). The opinions expressed are those of the authors, not the funding agencies.

References

Achenbach, T. M., & Rescorla, L. A. (2000). *CBCL/1, 5-5 & C-TRF/1, 5-5 profiles*. Burlington, VT: Research Center for Children, Youth, and Families.

Alvero, A., Struss, K., & Rappaport, E. (2008). Measuring safety performance. *Journal of Organizational Behavior Management*, *27*, 1–28.

Booren, L. M., Downer, J. T., & Vitiello, V. E. (2012). Observations of children's interactions with teachers, peers, and tasks across preschool classroom activity settings. *Early Education & Development*, *23*, 517–538.

Conroy, M. A., Sutherland, K. S., Algina, J. J., Wilson, R. E., Martinez, J. R., & Whalon, K. J. (2014). Measuring teacher implementation of the BEST in CLASS intervention program and corollary child outcomes. *Journal of Emotional and Behavioral Disorders*, *23*, 144–155.

Devine, S. L., Rapp, J. T., Testa, J. R., Henrickson, M. L., & Schnerch, G. (2011). Detecting changes in simulated events using partial-interval recording and momentary time sampling III: Evaluating sensitivity as a function of session length. *Behavioral Interventions*, *26*, 103–124.

Dunlap, G., & Kern, L. (1996). Modifying instructional activities to promote desirable behavior: A conceptual and practical framework. *School Psychology Quarterly*, *11*, 297–312.

Feil, E. G., Severson, H. H., & Walker, H. M. (1998). Screening for emotional and behavioral delays: The early screening project. *Journal of Early Intervention*, *21*, 252–266.

Ferro, J., Foster-Johnson, L., & Dunlap, G. (1996). Relation between curricular activities and problem behaviors of students with mental retardation. *American Journal on Mental Retardation*, *101*, 184–194.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*, 59–109.

Hojnoski, R. L., & Missall, K. N. (2010). Social development in preschool classrooms: Promoting engagement, competence, and school readiness. In M. R. Shinn & H. M. Walker (Eds.), *Interventions for achievement and behavior problems in a three-tier model including RTI* (pp. 703–728). Bethesda, MD: National Association of School Psychologists.

Hughes, J. N., & Kwok, O. M. (2006). Classroom engagement mediates the effect of teacher–student support on elementary students' peer acceptance: A prospective analysis. *Journal of School Psychology*, *43*, 465–480.

Lane, J. D., & Ledford, J. R. (2014). Using interval-based systems to measure behavior in early childhood special education and early intervention. *Topics in Early Childhood Special Education*, *34*, 83–93.

Lewis, T. J., Scott, T. M., Wehby, J. H., & Wills, H. P. (2014). Direct observation of teacher and student behavior in school settings: Trends, issues and future directions. *Behavioral Disorders*, *39*, 190–200.

Mann, J., Have, T., Plunkett, J. W., & Meisels, S. J. (1991). Time sampling: A methodological critique. *Child Development*, *62*, 227–241.

McWilliam, R. A., Scarborough, A. A., & Kim, H. (2003). Adult interactions and child engagement. *Early Education and Development*, *14*, 7–28.

Powell, J., Martindale, A., & Kulp, S. (1975). An evaluation of time-sample measures of behavior. *Journal of Applied Behavior Analysis*, *8*, 463–469.

Rimm-Kaufman, S. E., La Paro, K. M., Downer, J. T., & Pianta, R. C. (2005). The contribution of classroom setting and

quality of instruction to children's behavior in kindergarten classrooms. *The Elementary School Journal*, *105*, 377–394.

Tapp, J. (2010). *Lily data collector*. Nashville, TN: Vanderbilt Kennedy Center.

Williford, A. P., Vick Whittaker, J. E., Vitiello, V. E., & Downer, J. T. (2013). Children's engagement within the preschool classroom and their development of self-regulation. *Early Education and Development*, *24*, 162–187.

Wood, B. K., Hojnoski, R. L., Laracy, S. D., & Olson, C. L. (2015). Comparison of observational methods and their relation to ratings of engagement in young children. *Topics in Early Childhood Special Education*, *35*, 211–222.

Yoder, P., & Symons, F. (2010). *Observational measurement of behavior*. New York, NY: Springer.

Yoder, P. J., Ledford, J. R., Harbison, A. L., & Tapp, J. T. (2017). Partial-interval estimation of count: Uncorrected and Poisson-corrected error levels. *Journal of Early Intervention*, *40*, 39–51.