

Assessment of Prosocial and Disruptive Behaviors in a Head Start Classroom: A Pilot Study of a Behavioral Observation System for Preschoolers

Christopher Campbell, Elaine K. Martin, Sarah Davenport, Brittany K. Liebsack, Glen J. Veed, Mary Fran Flood, and David J. Hansen
University of Nebraska-Lincoln

Introduction

Recent national statistics by the U.S. Department of Health and Human Services (2001) indicate that 10-20% of preschool children experience mild to moderate behavior problems. Relatedly, empirical research suggests that early emotional and behavioral problems can negatively impact a preschooler's acquisition of important skills relevant to future school readiness and adjustment (Knitzer, 2003), with greater risk for children from low-income, highly stressed families who are exposed to poverty, malnutrition, and community violence (Campbell, 1997; Garbarino, 1995). Further, research has demonstrated that children who do not complete high school are significantly more likely as adults to display a multitude of behaviors that are destructive to themselves and others including substance abuse, delinquency, and crime (e.g., Hawkins & Lishner, 1987; Hinshaw, 1992). Thus, evidence-based assessments are critical to the identification of early childhood problems in order to guide subsequent interventions so that future problems can be diminished (Lopez et al., 2000).

For many years, rating scales have been used as the primary assessment tool for children with behavior problems (Eyberg & Pincus, 1999). While behavior rating systems have multiple advantages (e.g., low cost, ease of administration), these instruments may be subject to biasing factors, such as the respondent's global impressions of the child and respondent's mood at the time of administration (Edelbrock, 1988; Patterson, 1982). Direct observation is an alternative, more objective solution to evaluate classroom behaviors and has been found to be the most ecologically valid means of identifying children who exhibit disruptive behaviors in the classroom (Barkley, 1991). However, observational assessment is not limited to challenging behaviors and has been consistently cited as the major tool for assessing a variety of behavioral, socioemotional, and learning needs of young children across natural contexts (e.g., Boehm & Sandberg, 1982; Cohen, Stern, & Balaban, 1997; Volpe & McConaughy, 2005). Common examples of observed behaviors include interactions with peers and adults, participation in activities, play engagement, motor behaviors, emotional states, and attention (Brassard & Boehm, 2007). According to Brassard and Boehm (2007), observation is the means of validating information collected by other assessment tools, and crucial in the identification of appropriate interventions.

The purpose of the current study was to establish empirical support for the psychometric properties of a 35-item direct observation assessment measure for preschool children. More specifically, the inter-rater reliability of the coding system, internal consistency of the proposed scales, as well as the temporal stability of the scales was evaluated. In addition, convergent validity was assessed by examining the correlation of the observational system with another teacher rating scale.

Method

Participants

Participants were 27 preschool children (15 male and 12 female) enrolled in three different preschool classrooms at a Midwestern Head Start Center. The children ranged in age from 3 to 5 years ($M = 4.64, SD = 0.58$). The majority of the children were European American (63%), 26% were African American, 7% Hispanic, and 4% Native American. The majority of the children (74.1%) were currently raised by single mothers (biological), and the average yearly household income for participants was \$17,800 ($SD = \$13,600$).

Measures

Campbell and Martin Behavioral Observation of Preschoolers System (CAMBOPS-35). The CAMBOPS-35 was originally created to evaluate the effectiveness of consultation services delivered to a Midwestern Head Start Child Development Center. The coding system is a 15-minute observation period separated into 30-second intervals (25-second observation interval and a five-second recording interval) to capture 35 prosocial and disruptive behaviors across five factors in two domains (Figure 1).

Behavior Assessment System for Children-Teacher Rating Scales-Preschool (BASC-2 - TRS-P; Reynolds & Kamphaus, 2004). The TRS-P is a 100-item teacher-report measure used to assess both adaptive and problem behaviors in a preschool setting. Respondents rate descriptions of child behaviors on a four-point scale of frequency, ranging from *Never to Almost Always*. Items are divided into four composite scores and 11 subscales. The TRS-P has been shown to have strong reliability and validity (Reynolds & Kamphaus, 2004).

Procedure

A team of nine undergraduate students, trained to reliability, completed independent observations of the children's behaviors in the Head Start classrooms. On average, each child was observed twice each week (total of 30 minutes) for 8 weeks during classroom activities and free play, both in the classroom and outdoors. Once coders were trained to reliability, inter-observer reliabilities were calculated across the 2-month observation period to catch observer drift. When coding was complete, frequencies of children's prosocial and disruptive observed behaviors were calculated.

Table 1. Inter-Observer Reliability of Live Coding of the CAMBOPS-35 Categories in Classroom and Outdoor Situations

Category	Pearson Correlation
Cooperation with Adults	
Follows Instructions from Adult	.76**
Engaged in Activities	.74**
Active Participation	.96**
Talks to Adults Appropriately	.91**
Interacting with Adults	.85**
Peer Interactions	
Actively Playing with Peers	.90**
Talks to Peers	.91**
Shares with Peers	.91**
Invites Peers to Play	.60**
Waits Their Turn	.98**
Imitation of Peers	.80**
Independent & Self-Regulating Behaviors	
Independent Time - Work/Play	.91**
Stimulus or Laughter	.88**
Cries/Whines	.99**
Apologizes for Behavior	.70**
Self-Soothing Behavior	.56**
Challenging Behaviors	
Noncompliance/Defiance	.79**
Completes Consequences	.95**
Disrupts Established Activities	.56**
Aggression Toward Peers	.95**
Aggression Toward Adults	1.00**
Spreads Activities	.98**
Throws, Kicks, or Hits	.41*

* $p < .05$; ** $p < .01$. Note: The table does not include the 12 behaviors that did not occur during the 37 interobserver sessions.

Table 2. Temporal Stability of the CAMBOPS-35 Categories

Weekly Scales	Week #1	Week #3	Week #5	Week #7
Cooperation with Adults				
Week #1	---	---	---	---
Week #3	.32*	---	---	---
Week #5	-.21	-.02	---	---
Week #7	-.09	.17	.01	---
Peer Interactions				
Week #1	---	---	---	---
Week #3	.18	---	---	---
Week #5	.13	.07	---	---
Week #7	.08	-.19	-.19	---
Independent & Self-Regulating Behaviors				
Week #1	---	---	---	---
Week #3	.05	---	---	---
Week #5	.09	-.03	---	---
Week #7	-.32*	-.18	-.35	---
Challenging Behaviors				
Week #1	---	---	---	---
Week #3	-.08	---	---	---
Week #5	-.08	.05	---	---
Week #7	.09	-.14	-.04	---
Atypical Behaviors				
Week #1	---	---	---	---
Week #3	-.04	---	---	---
Week #5	-.03	-.05	---	---
Week #7	-.03	-.04	-.03	---

* $p < .10$; ** $p < .05$; *** $p < .01$.

Figure 1. Proposed Model of the CAMBOPS-35 Coding System

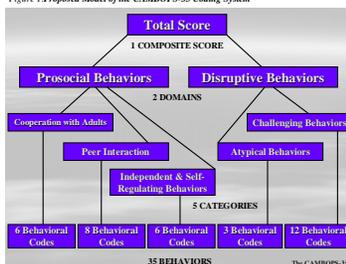


Table 4. Convergent Validity of the CAMBOPS-35 Scales with the Behavior Assessment System for Children - 2nd Edition (BASC-2)

Category	Cooperation with Adults	Peer Interactions	Independent & Self-Regulating Behaviors	Challenging Behaviors	Atypical Behaviors
TRS-P Composite Scales					
Adaptive Skills	-.09	.48*	-.02	-.32*	-.42*
Behavioral Symptoms	-.25	-.13	-.28	.63**	.11
Externalizing	-.35*	-.12	-.23	.75**	-.06
Internalizing	-.29	-.14	-.35*	-.22	.11
TRS-P Primary Scales					
Adaptability	-.09	.04	.04	-.37	-.15
Aggression	-.10	-.27	-.60**	-.16	-.16
Anxiety	-.24	-.10	-.40*	-.22	-.07
Attention Problems	.05	-.25	-.26	.54**	.21
Atypically	-.03	-.43*	-.06	.19	.40*
Depression	.13	.10	-.41*	.18	-.15
Functional Communication	-.04	.63**	-.12	-.11	-.49*
Hyperactivity	.28	.13	-.18	.75**	.06
Social Skills	-.08	.80**	-.00	-.37	-.43*
Somatization	-.44*	-.26	-.07	-.39*	-.33*
Withdrawal	.10	-.13	.06	.12	.07

* $p < .10$; ** $p < .05$; *** $p < .01$.

Results

The inter-observer reliability estimates for the CAMBOPS-35 are presented in Table 1. Pearson correlations ranged from .41 (Throws, Kicks, or Hits) to 1.00 (Aggression Toward Adults - which includes both verbal and physical aggression). Over an 8-week period, all of the 27 participants were observed approximately twice each week ($M = 16.04, SD = 3.71$) resulting in approximately 400 total observations. In addition, all 27 participants were assessed at least once during 37 interobserver sessions conducted throughout the study. All of the 35 CAMBOPS codes were recorded during the 400 total observations. However, twelve behavioral codes were not observed during the 37 observational periods. Of the 23 observed codes, 12 codes had correlation coefficients of .90 or higher, which is significant as the CAMBOPS-35 is a five coding system.

Temporal stability of the CAMBOPS-35 categories is recorded in Table 2. In an attempt to conserve space, the odd weeks are presented but similar results were found across all eight weeks. Insignificant results for temporal stability are likely explained by: (a) the categories may not be measuring a single and reliable construct as described below, and (b) observations across the eight-week period occurred in a variety of contexts ranging from organized group activities in the classroom to free time outside. Thus, temporal stability could be impacted by the lack of observations at a scheduled time during a consistent activity.

The proposed structure of the CAMBOPS-35 is presented in Figure 1. The internal consistency or homogeneity of the items within each of the five categories was assessed using Cronbach's Alpha (Table 3). Unfortunately, none of the categories had an "adequate" alpha coefficient of .70 or higher suggesting that the current behavioral codes are not measuring a single, unidimensional construct. Peer Interactions and Challenging Behaviors represent the best coefficient values (.42 and .38, respectively). The negative value for Independent and Self-Regulating Behaviors, implies a negative average covariance among items and, given the small number of items in the category, could reflect sampling error that has produced a negative average covariance in a given number of cases. Equally likely (and potentially the case for both the Atypical Behavior and Cooperation with Adults categories), is that the items truly do not have positive covariances, and therefore may not be measuring the same construct.

Lastly, convergent validity was assessed by correlating CAMBOPS-35 categories with the Composite and Primary Scales of BASC-2 Teacher Rating Scales for Preschoolers (Table 4). As shown in Table 4, many of the scales correlated in the expected direction (e.g., the CAMBOPS-35 Challenging Behaviors category was significantly correlated with the Behavioral and Externalizing Composite scales in addition to the Aggression, Attention Problems, and Hyperactivity Primary scales).

Discussion and Future Directions

Overall, an evaluation of the psychometric properties of the CAMBOPS-35 produced mixed empirical support. Investigation of the inter-rater reliability revealed significant levels of agreement on most behavioral codes. More importantly, reliability remained consistently high over an eight-week observational period, suggesting that the detailed operational definitions were appropriate and the level of training provided to the undergraduate observers was adequate.

An examination of the internal consistency of the observational system indicates that the current rationally-derived structure of the 35 behavioral codes into five categories (i.e., Cooperation with Adults, Peer Interactions, Independent and Self-Regulating Behaviors, Challenging Behaviors, and Atypical Behaviors) may not be appropriate. However, results from this pilot study may reflect the limited number of items in each of the categories. Adding, deleting, or editing items, in addition to restructuring categories may be necessary for future research. Convergent validity patterns with the TRS-P are encouraging as many of the CAMBOPS categories, with the exception of Cooperation with Adults, are significantly correlated with compatible scales in the expected direction.

The context of this study should be noted as the observations presented in this study represent the baseline data collected prior to a teacher-child intervention in a Midwestern Head Start Development Center. At the time of this study, this Center had been open for 13 months and had experienced a total of 23 departures, including the Director five months prior to our intervention. In addition, most teachers at the development center had less than two years experience working with preschool children. Thus, the classroom environment, including teachers and daily routines, was constantly changing. Temporal stability could have been impacted by multiple factors including inconsistent daily routines, unstable classroom structure, irregular teaching styles and expectations, and observations across numerous classroom and outdoor activities. Additionally, in an attempt to objectively assess behaviors while avoiding biasing factors, undergraduate observers were encouraged to observe different children each week which means that observations did not occur at a regularly scheduled time during a consistent activity.

In summary, interobserver reliability and convergent validity results are encouraging, while findings for internal consistency and temporal stability present challenges that will be addressed in future studies. Future projects utilizing the CAMBOPS-35 observational system will observe more children, interacting with teachers from diverse teaching backgrounds, in a variety of classroom environments (e.g., recently established as well as long-standing child development centers).