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, 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 (2011) indicate that 10-20% of preschool children experience mild to moderate behavior problems. Reliably, 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 (Kazdin, 2003), with greater risk for children from low-income, highly stressed families who are exposed to poverty, maltreatment, and community violence (Campbell, 1997; Garberius, 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., Honkala & Lapham, 1987; Hindman, 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 (Copas, 2000).

For many, 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 (Bobbit, 1986; 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 (Barlow, 1995). However, observational assessment is not limited to challenging behaviors and has been consistently cited as the most powerful tool for assessing a variety of behavioral, affective, and social-communication skills of young children across natural contexts (e.g., Achenbach & Edelbrock, 1983; Chamberlain & Sandberg, 1982; Collen, Stern, & Baldwin, 1997; Yaple & McGinnity, 2005). Common examples of observed behaviors include interactions with peers and adults, participation in activities, play engagement, motor behaviors, emotional states, and attention (Braswell & Bodin, 2007). According to Braswell and Bodin (2007), observation is the most accurate method of obtaining information related to other assessment tools, and crucial in the identification of appropriate interventions.

The proposed structure of the CAMBOPS-35 is presented in Figure 1. The temporal stability or homogeneity of the 35-item direct observation assessment was examined 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 were evaluated. In addition, convergent validity was assessed by examining the correlation of the observation system with another teacher rating scale.

Method
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 = 4.44, SD = 0.58). The majority of the children were European American (60%), 20% were African American, 7% Native American, and 3% Asian. The majority of the children (75%) were currently raised by single mothers (biological), and the average yearly household income for participants was $17,800 (SD = $13,000).

Measures
Campbell and Martin Behavioral Observation of Preschool 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 5-second recoding interval) to observe 35 prosocial and disruptive behaviors across five domains (Figure 1).

Behavior Assessment System for Children-Teacher Rating Scales (BASC-2), TRS-P. Reynolds and Kamphaus, 2006. The TRS-P is a 100-item teacher-report measure used to assess both adaptive and atypical behavior problems in a pro-social setting. Respondents rate descriptions of behavior on a four-point scale of frequency (ranged from “Never” to “Multiple Times a Day”) for 10 subscales. The TRS-P has been shown to have high validity and reliability (Reynolds & Kamphaus, 2006).

Results
The inter-observer reliability estimates for the CAMBOPS-35 are presented in Table 1. Pearson correlations ranged from .76 (Threats, 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 = 4.06, SD = 1.71) resulting in over 400 total observations. In addition, all 27 participants were assessed at least once during seven interobserver periods conducted throughout the study. All of the 35 CAMBOPS codes were recorded during the 400 total observations. However, twelve behavioral codes were observed during the 35 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 live coding system.

The proposed structure of the CAMBOPS-35 is presented in Figure 1. The internal consistency or homogeneity of the scales within each of the five categories was assessed using Cronbach’s Alpha (Table 2). Unfortunately, some of the categories had an “acceptable” alpha coefficient of .70 to .79 suggesting that the current beta-reliability code can not measure a single, multidimensional construct. Peer Interactions and Challenging Behaviors represent the best coefficient values (.82 and .58, respectively). The negative values for Independent and Self-Regulating Behaviors, implies a negative average covariance among items and if the small number of items in the category, could reflect sampling error that has produced a negative average covariance in a given case. 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.

Largely, 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 Challenging Behavior 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, evaluation of the psychometric properties of the CAMBOPS-35 preschool 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 different operational definitions were appropriate and the level of training provided to the undergraduate observer was adequate. An examination of the internal consistency of the observational system indicates that the current rating-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. Convergent validity results from this pilot study may reflect the limited number of items in each of the categories. Adding, deleting, or editing items, additional observational rating scales in the CAMBOPS-35 may be necessary for future research. Convergent validity with the TRS-P is encouraging as many of the CAMBOPS categories, with the exception of Cooperation with Adults, are significantly correlated with comparable scales in the expected direction.

In summary, inter-rater 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).