Title: Journal of attention disorders.

Author: Floyd, Randy & Kirby, Edward

Imprint: Toronto ; North Tonawanda, NY : Multi-Health Systems Inc., c1996

Article: Psychometric properties of measures of behavioral inhibition with preschool-age children: Implications for assessment of children at risk for ADHD

Vol: 5 No: 2 Pages: 79

Date: Nov 2001

Maxcost: 35.00 ifm

Lender: *SCN, VCB, FNN, XFF, CLU


BILLTO: Same SHIP VIA: Library Rate MAXCOST: 35.00 ifm COPYRT COMPLIANCE: CCG FAX: (252) 328-4008 ARIEL: 150.216.67.44 E-MAIL: jis@mail.ece.edu AFFILIATION: ASERL participant, SOLINE, KUDSU LENDING CHARGES: SHIPPED: SHIP INSURANCE: LENDING RESTRICTIONS: LENDING NOTES: RETURN TO: RETURN VIA:

ILL: 2550195 Borrower: ERE

Req Date: 01/23/2004 OCLC #: 34159342

Patron: Mangione Walcott, Christy

Author: Floyd, Randy & Kirby, Edward

Title: Journal of attention disorders.

Article: Psychometric properties of measures of behavioral inhibition with preschool-age children

Vol: 5 No: 2

Date: Nov 2001 Pages: 79

Verified: <TN:50981> OCLC ISSN: 1087-0547 [Form

Maxcost: 35.00 ifm Due Date:

Lending Notes:

Bor Notes:
Psychometric properties of measures of behavioral inhibition with preschool-age children: Implications for assessment of children at risk for ADHD

R. G. Floyd and E. A. Kirby

Measures of behavioral inhibition offer promise in research with young children with ADHD. This study investigated the factorial validity, ecological validity, and temporal reliability of five performance-based measures of behavioral inhibition in a sample of 70 3-, 4-, and 5-year-old children.

An exploratory factor analysis yielded a single factor that accounted for 65% of the variance in the analysis. This factor was found to significantly correlate with teacher ratings of aggression and inattention and was found to differentiate both between sexes and between older and younger children. Several individual measures of behavioral inhibition demonstrated statistically significant correlations with teacher ratings of aggression and inattention. Adequate 1- to 2-week test-retest reliability was demonstrated for only two measures. Some limitations of the study are discussed and suggestions for future research are presented.

Attention-deficit/hyperactivity disorder (ADHD) is a psychiatric diagnosis representing a consistent pattern of inattention, hyperactivity, and impulsivity that is developmentally inappropriate or maladaptive (American Psychiatric Association, 1994). During the past 2 decades, leaders from the National Institute of Mental Health and other prominent researchers of child psychopathology have called for the increased study of assessment methodologies and treatment strategies for preschool-age children who are overactive, aggressive, and difficult to manage (Barkley, 1994; Campbell, 1985, 1990; Campbell, Pierce, March, Ewing, & Szumowski, 1994; DuPaul & Stoner, 1994; Jensen et al., 1993; Mariani & Barkley, 1997). These calls have been based on an increasing body of evidence indicating that many disruptive 3-, 4-, and 5-year-old children display poor adaptive functioning during the preschool years (Lahey et al., 1998) and experience a number of negative developmental outcomes (see Barkley, 1998 and Shelton et al., 1998 for reviews).

Research has indicated that the core deficit in ADHD is behavioral impulsivity or disinhibition (Barkley, 1997a, 1997b; Seidman, Biederman, Farone, Weber, & Ouellette, 1997). Barkley (1994, 1996a, 1997a, 1997b) has hypothesized that this deficit stems from abnormalities in the development of behavioral inhibition. According to Barkley, behavioral inhibition comprises three components: (a) the inhibition of prepotent responses, (b) the interruption of ongoing responses, and (c) interference control. The first component is the capacity to inhibit a response that has been consistently reinforced in the past or that generally is "more automatic or dominant" (Barkley, 1997a, p. 102). The second component is the ability to interrupt ongoing, but ineffective, responding to facilitate adaptive functioning. The third component is the ability to suppress the influence of external stimuli or cognitions that may interfere with current thoughts, plans, and behaviors. According to Barkley (1997a), deficits in behavioral inhibition at a young age affect the development and use of four executive functions: nonverbal working memory; verbal working memory; self-regulation of affect, motivation, and arousal; and, analytic and synthetic capacity. In turn, the direct effects of weak executive functions lead to the dysfunction in goal-directed, adaptive behaviors seen in individuals with ADHD.

Barkley (1997a) suggests that behavioral inhibition stems primarily from genetic and neurobiological origins. Behavioral inhibition appears to emerge during infancy and typically continues to develop at a rapid rate throughout
the preschool years (Lahey et al., 1994). Because behavioral inhibition is the foundation on which executive functions and self-control develop, developmentally sensitive and objective measures of behavioral inhibition offer promise in identifying young children at-risk for developing ADHD (Byrne, DeWolfe, & Bawden, 1998; Cole, Usher, & Cargo, 1993; Kindlon, Mezzacappa, & Earls, 1995). Although behavioral inhibition can be inferred from observations of behaviors in natural or analogue settings, performance-based measures may provide the most direct method to operationalize this construct in young children. It is possible that focusing these measures on what has been conceptualized as the core feature in ADHD will increase their diagnostic validity and clinical utility (Rapport, Chung, Shore, Denny, & Isaacs, 2000).

A number of studies of school-age children and adults with disinhibited behaviors have examined tests that are purported to measure behavioral inhibition and associated executive functions, such as the Wisconsin Card Sort (Milner, 1963), the Stroop Word-Color Interference Task (Stroop, 1935), the antisaccades task (Guitton, Buchtel, & Douglas, 1985; Roberts, Hager, & Heron, 1994), the Tower of Hanoi (Salilice, 1982), the stop-signal paradigm (Oosterlaan, Logan, & Sergeant, 1998; Schachar, Tannock, Marroitt, & Logan, 1995), and the go/no-go paradigm (Tronner, Hoeppner, Lobef, & Armstrong, 1988; Tronner, Hoeppner, & Zecker, 1991; see Lyon & Krasnegor, 1996 and Kimberg & Farah, 1993 for reviews). However, a recent review of ERIC, PsycINFO, and MEDLINE did not reveal any research that specifically examined the relationship between Barkley’s (1997a) behavioral inhibition and disinhibited behaviors in young children. Because the construct of behavioral inhibition was extrapolated from research that examined broader concepts such as self-control, self-regulation, and inhibitory control (e.g., Gaddis & Martin, 1989; Kochanska, Murray, & Coy, 1997; Kochanska, Murray, Jacquet, Koeing, & Vandeegeest, 1996; Kopp, 1982; Millich & Kramer, 1984; Mischel, Shoda, & Rodriguez, 1989), several measures of these related constructs that appear to assess behavioral inhibition in preschool-age children surfaced in the review. It is notable that measures of reflectivity-impulsivity were excluded from this review because previous research with young children has indicated that there are two types of impulsivity in young children—cognitive impulsivity and behavioral impulsivity (Gaddis & Martin, 1989; Olson, 1989; Olson, Bates, & Bayles, 1990). Drawing upon this and related research, Barkley (1997a) conveyed, “It is the ‘behavioral’ type of impulse control [i.e., behavioral inhibition] that seems to be more stable over development, to correspond more closely to parent or teacher ratings of hyperactive-impulsive behavior, and to correlate more highly with later cognitive and social competencies than does the cognitive dimension of reflectiveness (as in the Matching Familiar Figures Test, Draw-a-Line-Slowly Test)” (p. 62). Thus, measures of the cognitive type of impulse control (i.e., reflectivity-impulsivity) were excluded because they have been shown to assess construct distinct from behavioral inhibition.

Measures designed to assess delay-of-gratification (DG) and resistance-to-temptation (RTT) and those using the go/no-go paradigm appear to have demonstrated the strongest validity in measuring behavioral inhibition in young children. DG tasks tap the ability to wait for highly preferred, but delayed rewards, rather than choosing less preferred, yet immediate rewards (Funder & Block, 1989). During these tasks, the highly preferred rewards are provided only if participants inhibit choosing the more immediate rewards. In a series of studies, Mischel and colleagues used a DG task to examine the correlates and developmental competencies predicted by DG (see Mischel et al., 1989 for a review). In one study, children were presented with a preferred food item (i.e., a marshmallow) and told that they would receive two food items if the initial one was not eaten when the examiner returned. Typically children were left alone for 15 to 20 minutes. The duration of children’s delay at 4 years of age accurately predicted parent ratings of frustration and stress tolerance, and measures of social competence, cognitive skills, and academic achievement during adolescence (Mischel, Shoda, & Peake, 1988). Although studies have demonstrated that school-age children with ADHD display deficits in DG (e.g., Schweitzer, 1996; Schweitzer & Sulzer-Azaroff, 1988), no studies have examined DG abilities in young children with severe disinhibited behaviors.

A notable body of research has demonstrated the validity of RTT tasks in predicting displays of behavioral inhibition in preschool-age children. RTT tasks differ from DG tasks because attractive rewards are immediately available but are not contingent upon the successful resistance of less attractive rewards (Funder & Block, 1989). Campbell et al. (1982) incorporated a RTT task in research examining parent-referred 2- and 3-year-old children who displayed severe disruptive behaviors and their non-referred agemates. The RTT task required children to resist eating a cookie placed before them. Disinhibited responses on this measure consistently differentiated between the two groups of children across a 2-year period (Campbell et al., 1984). This finding is corroborated by a number of studies of young
children that demonstrate that measures using the RIT paradigm and a variety of attractive stimuli: (a) differentiate between younger and older preschoolers (Golden, Montare, & Bridger, 1977; Vaughn, Kopp, & Krakow, 1984); (b) discriminate between children who display severe symptoms of ADHD, aggression, and noncompliance and their typical agemates (Campbell et al., 1994; Marakozvit & Campbell, 1998); (c) correlate with more ecologically valid assessment methods, such as concurrent sociometric ratings of peer relations and parent ratings of disinhibited behaviors (Olson, 1989; G. Kochanska, personal communication, July 30, 1997; Silverman & Ragan, 1990; Silverman & Ragan, 1992); and, (d) predict later activity level and aggression, teacher ratings of disruptive behaviors, and performance on measures of problemsolving skills and sustained attention (Campbell et al., 1994; Marakozvit & Campbell, 1998; Olson, 1989; Olson & Hoza, 1993).

Go/no-go tasks assess the ability to perform a specified behavior after one cue is presented and to inhibit this response (or produce another response) after a different cue. Typically, measures using this paradigm require that a more natural, or prepotent, response be inhibited to respond correctly. School-age children with ADHD have been shown to make more errors than their typical agemates on a go/no-go task requiring them to quickly tap a finger once when an examiner taps twice and to tap twice when an examiner taps once (Trommer et al., 1988). Two studies with preschool-age children have used a tapping task that incorporates the go/no-go paradigm to measure behavioral inhibition (Cole et al., 1993; Diamond & Taylor, 1996). In these studies, the tapping task required children to tap a stick in response to an examiner's model. Children were instructed to immediately tap once after the examiner tapped twice and to tap twice after the examiner tapped once. A sample of non-referred 3- to 7-year-old children demonstrated significant age-related improvements in performance on this task (Cole et al., 1993). In another study, Reed, Pien, and Rothbart (1984) developed two go/no-go tasks—a pinball measure and a Simon-says task—to assess abilities related to behavioral inhibition in preschool-age children. The pinball measure required children to discriminate between verbal signals and to respond to the specified signal that used the release of a pinball plunger. The Simon-says task required children to follow directions given to them by one toy animal and to ignore directions given to them by another toy animal. Recent adaptations of these measures have surfaced in the research using the Inhibitory Control Battery (ICB; Kochanska et al., 1996; Kochanska et al., 1997). This research using this battery has provided evidence of the validity of these measures in predicting concurrent maternal ratings of disinhibited behaviors (G. Kochanska, personal communication, July 30, 1997).

Although evidence exists for the validity of DG RIT, and go/no-go tasks in predicting disinhibition in young children, there has been less research examining constructs related to interference control in this age group. Measures of interference control in older children and adults focus on conditions in which individuals must resist or distraction from stimuli not associated with response requirements (Barkley, 1997a). The Stroop Word-Color Interference Test (SCITT; Stroop, 1935) is an example. Scores from the SCITT have consistently differentiated between school-age children with ADHD and control children (Grodzinsky & Diamond, 1992; Pennington, Gussier, & Welch, 1993; Seidman et al., 1997) but not between children from other diagnostic groups (Barkley, Grodzinsky, & DuPaul, 1992). Several measures resembling the SCITT have been developed for use with preschool-age children, but evidence of their validity is minimal. For example, the day-night task developed by Gerstadt, Hong, and Diamond (1994) required young children to resist interference from task stimuli in order to correctly label pictures according to task rules. The day-night task required children to complete trials of the task by saying "day" to a picture of a nighttime sky and saying "night" to a picture of a daytime sky. In a sample of 3- and 4-year-old children, the percentage of younger children on this measure became progressively worse across trials, but this pattern was not observed in older children. Thus, although young children could sustain interference control during initial trials of the task, they had difficulty maintaining it throughout trials.

Another measure that appears to assess interference control by using the Stroop paradigm is the Shapes task from the Inhibitory Control Battery (Kochanska et al., 1997). During its administration, images of common objects and animals were used as stimuli. Numerous small images speckled the inside of the larger images. Some of the small images matched the larger ones, and other small images were discrepant. Children were required to resist the interference of the larger image and to name the smaller images. At present, no research specifically examining the validity of the Shapes task has been offered.

In summary, several measures that may assess behavioral inhibition in preschool-age children have evidenced significant age effects and demonstrated significant relationships with ecologically valid criteria. Although a diffuse and disorganized body of evidence has suggested...
the validity of these measures in assessing behavioral inhibition (Tannock, 1998), there is a dearth of information regarding their test-retest reliability, and limited information regarding their factorial and criterion-related validity. Following the recommendations of Rapport et al. (2000), this study examined the psychometric properties of six measures of behavioral inhibition in order to: (a) extend the research examining the measurement of the core cognitive feature in ADHD, and (b) facilitate the development of the next generation of performance-based assessment methods and their clinical applications with preschool-age children at-risk for ADHD. This study of typically developing 3- to 5-year-old children was designed to provide initial estimates of the consistency of these measures in assessing behavioral inhibition across time, to test their construct or factorial validity, and to examine their relationships with ecologically valid measures of similar constructs.

Method

Participants

Seventy 3- to 5-year-old children (30 girls and 40 boys, M age = 4.64 years) participated in the study (see Table 1). Parents of 19 children (27%) were faculty or staff at a state university, parents of 20 children (29%) were university students, and parents of 31 children (44%) were not associated with the university. Sixty-nine children attended a daycare or preschool program (M length = 2.26 years). Eight children received speech therapy for articulation errors, and 2 children had mild hearing difficulties.

Table 1. Demographic Variables and Recruitment of Participants

<table>
<thead>
<tr>
<th>Classification</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Girls</td>
<td>30</td>
</tr>
<tr>
<td>3-year-old girls</td>
<td>8</td>
</tr>
<tr>
<td>4-year-old girls</td>
<td>6</td>
</tr>
<tr>
<td>5-year-old girls</td>
<td>18</td>
</tr>
<tr>
<td>Boys</td>
<td></td>
</tr>
<tr>
<td>3-year-old boys</td>
<td>40</td>
</tr>
<tr>
<td>4-year-old boys</td>
<td>10</td>
</tr>
<tr>
<td>5-year-old boys</td>
<td>13</td>
</tr>
<tr>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>57</td>
</tr>
<tr>
<td>Black</td>
<td>6</td>
</tr>
<tr>
<td>Other</td>
<td>7</td>
</tr>
<tr>
<td>Recruitment</td>
<td></td>
</tr>
<tr>
<td>University childcare centers</td>
<td>58</td>
</tr>
<tr>
<td>University community</td>
<td>8</td>
</tr>
<tr>
<td>Private childcare centers</td>
<td>3</td>
</tr>
</tbody>
</table>

Materials

Children completed six tasks adapted from the Inhibitory Control Battery: Shapes, Tongue, Tower, Snack Delay, Pinball, and Dog and Dragon (Kochanska & Jacques, 1996).

Behavioral inhibition tasks. Shapes is a verbal inhibition task that utilizes the Stroop paradigm (Stroop, 1935). During the priming phase, children were asked to name 13 full-page images of easily recognized objects (i.e., an apple, an ice cream cone). The examiner provided verbal praise for correct responses. During the test phase, children were asked to name only a smaller image placed unsystematically on the same page as a larger image. Trials consisted of 12 items containing identical large and small images and 12 items containing discrepant images. Responses on items with discrepancies between images were scored. The latency of each response was also recorded. Analysis of latencies indicated that discrepant trials, M = 1.80, SD = .65, produced significantly longer latencies than nondiscrepant trials, M = 1.50, SD = .60, t(88) = -7.80, p < .001.

Tower is a go/no-go task that involves stacking blocks. The examiner first modeled building a tower of six blocks, and then the children built another tower with the blocks. During this phase, the examiner provided verbal praise for children's block-stacking behavior. Children then were instructed to take turns with the examiner to build a tower with 11 blocks. After children placed a block, the examiner waited 10 seconds for children either to place another block themselves or to prompt him to place the next block. Children's sequential placements of blocks were ignored during each trial. The number of blocks placed by children across three trials was recorded.

Snack Delay is a RTT task that required children to delay 20, 30, 40, 50, and 60 seconds before eating a food item placed under a clear cup. After children chose a preferred food item (i.e., mini M&Ms, raisins, ring-shape fruit cereal, gummy candies, or mini marshmallows) and completed a practice trial, the examiner placed the chosen food item under a clear cup. Children were instructed to wait with their hands on the table until the end of the delay. Behaviors during the delay period were scored. Scoring ranged from 10 points for waiting quietly with hands on the table to 0 points for eating the food item.

Pinball is a go/no-go paradigm task that required children to depress a lever on a pinball game until an appropriate signal is given. The first six test trials required children to depress the lever until they heard the examiner say “Go.” The next six trials used green and red circles as signals.
Children were instructed to release the lever when a green circle was shown and to continue to depress the lever when a red circle was shown. Scores were based on the latency until the full release of the lever. Two points were subtracted from trial scores if children released the lever from its fully depressed position but did not shoot the pinball. A maximum of three of these errors was recorded on each trial.

Dog and Dragon is a go/no-go task that follows the format of "Simon says." During the task, children watched a videotape segment on a color monitor in which a brown dog puppet and pink dragon puppet issued verbal commands at 3-second intervals. Children were instructed to obey the commands of the dog and to inhibit responding to the commands of the dragon. Six trials of each set of commands were presented. Scoring criteria for the dog trials ranged from a score of 3 for full completion of the command to a score of 0 for no response. Responses to dragon commands were scored using the reversed point scale. Scores from the dog trials and from the dragon trials were summed to produce a total score. The choice to combine both scores was supported by initial evaluations of the different combinations of scores from the Dog and Dragon task using factor analysis. These analyses revealed a similar factor structure regardless of which score was used to represent performance on this task.

Tongue is a RTT task that required children to place a preferred food item on their tongues and to delay eating it. Preliminary screening of data from the Tongue task indicated the measure had a low ceiling, which led to a severe negative skew in its distribution. Tongue scores were omitted from subsequent analyses.

Rating scales. Teachers of a subsample of children enrolled in a single daycare program (n = 34) completed the Hyperactivity, Aggression, and Attention Problems subscales of the Behavioral Assessment System for Children (BASC) Teacher Rating Scale (Reynolds & Kamphaus, 1992). The BASC Teacher Rating Scale is a multidimensional and empirically based assessment report form. Information regarding the reliability and validity of this scale is reported in the BASC manual. The Teacher Rating Scale requires teachers to rate behaviors according to a 4-point scale reflecting frequency of occurrence.

 Procedures

Children were administered the six measures in childcare centers or in university clinic rooms. These measures were administered in standard order to all children: (a) Shapes, (b) Tongue, (c) Tower, (d) Snack Delay, (e) Pinball, and (f) Dog and Dragon. The first five tasks were administered using an easel format. To test the temporal stability of the behavioral inhibition tasks, 59% of children (n = 41) were restated 1 week to 2 weeks (range = 5 to 14 days) later on one of three randomly assigned pairs of tasks: Shapes and Tongue tasks, Tower and Snack Delay task, and Pinball and Dog and Dragon tasks.

 Results

Missing data points from three children were estimated using linear regression (Tabachnick & Fidell, 1994). To meet the assumptions of statistical analyses, data transformations (see Table 2) were also used to improve the normality of the distributions of the remaining variables from the behavioral inhibition tasks (Rummel, 1970; Tabachnick & Fidell, 1994). Transformations were also used to represent each variable on the same scale. As a result of these transformations, lower scores on behavioral inhibition tasks represent greater display of behavioral inhibition (i.e., more self-control). Higher scores represent greater levels of disinhibition.

Table 2. Original Descriptive Statistics and Transformation Statistics Used for Behavioral Inhibition Measures and Descriptive Statistics from BASC Subscales

<table>
<thead>
<tr>
<th>Measure</th>
<th>M</th>
<th>SD</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Transformation Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Shapes</td>
<td>17.10</td>
<td>7.90</td>
<td>-.60</td>
<td>-.35</td>
<td>Reflection Transformation</td>
</tr>
<tr>
<td>2. Tower</td>
<td>18.00</td>
<td>4.63</td>
<td>1.63</td>
<td>1.22</td>
<td>Log, Log10</td>
</tr>
<tr>
<td>3. Snack Delay</td>
<td>43.45</td>
<td>10.65</td>
<td>-.20</td>
<td>3.78</td>
<td>Inv, Inv, Rev</td>
</tr>
<tr>
<td>4. Pinball</td>
<td>138.93</td>
<td>7.22</td>
<td>-.45</td>
<td>11.11</td>
<td>Ref, Log, 10</td>
</tr>
<tr>
<td>5. Dog and Dragon</td>
<td>5.66</td>
<td>18.80</td>
<td>-.34</td>
<td>1.11</td>
<td>Ref, Sn Rt, Inv, Rev</td>
</tr>
<tr>
<td>6. Tongue</td>
<td>98.52</td>
<td>4.70</td>
<td>.22</td>
<td>.83</td>
<td>Reflection Transformation</td>
</tr>
<tr>
<td>7. BASC Hyperactivity</td>
<td>5.94</td>
<td>3.38</td>
<td>.15</td>
<td>.82</td>
<td>Log, Log10</td>
</tr>
<tr>
<td>8. BASC Aggression</td>
<td>11.32</td>
<td>3.32</td>
<td>.22</td>
<td>.04</td>
<td>Reflection Transformation</td>
</tr>
<tr>
<td>9. BASC Attn. Problems</td>
<td>7.76</td>
<td>3.32</td>
<td>.15</td>
<td>.82</td>
<td>Reflection Transformation</td>
</tr>
</tbody>
</table>

Table 3 presents the reliability coefficients from raw data from the Shapes, Tower, Snack Delay, Pinball, and Dog and Dragon tasks. Reliability coefficients ranged from .15 to .85. Scores from the two administrations of the Shapes and Tower tasks were significantly correlated, \( r (16) = .84, p < .001 \) and \( r (14) = .85, p = .001 \), respectively. Data from the two administrations of the Snack Delay, Pinball, and Dog and Dragon tasks did not display statistically significant reliability coefficients.

### Table 3. Test-Retest Reliability Coefficients for Behavioral Inhibition Measures

<table>
<thead>
<tr>
<th>Measures</th>
<th>( r )</th>
<th>( n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Shapes</td>
<td>.84**</td>
<td>16</td>
</tr>
<tr>
<td>2. Tower</td>
<td>.85**</td>
<td>14</td>
</tr>
<tr>
<td>3. Snack Delay</td>
<td>.50</td>
<td>14</td>
</tr>
<tr>
<td>4. Pinball</td>
<td>.15</td>
<td>11</td>
</tr>
<tr>
<td>5. Dog and Dragon</td>
<td>.52</td>
<td>11</td>
</tr>
</tbody>
</table>

**Note:** Data from behavioral inhibition tasks reflect raw scores.

\(^* p < .05\), \(^** p < .01\), \(^*** p < .001\).

The correlation matrix of Pearson product-moment correlations between behavioral inhibition tasks is presented in Table 4. Eight of 10 correlation coefficients were significant at least at the .05 probability level. Two correlations were significant at the .001 probability level. The relationships between variables with the effects of age removed were calculated using partial correlations. Three of 10 pair-wise correlations remained statistically significant: Snack Delay and Pinball tasks, Dog and Dragon and Pinball tasks, and Tower and Shapes tasks, \( pr (67) = .27, p = .03, pr (67) = .39, p = .001 \), and \( pr (67) = .39, p = .001 \), respectively.

To explore related clusters of variables, a factor analysis with principal axis extraction was performed on data from the five measures for the sample of 70 children. One factor was extracted with an eigenvalue of 2.24. This factor accounted for 45% of the variance in the analysis. The number of extracted factors was influenced by the following criteria: the latent root criterion; the scree test criterion; and, the residual criterion (Hair, Anderson, Tatham, & Black, 1995; Stevens, 1992; Tabachnick & Fidell, 1996). Although 6 residuals with absolute values greater than .05 were present in the residual correlation matrix, which suggested the presence of a second factor (Tabachnick & Fidell, 1996), a two-factor solution was rejected based on the latent root criterion and scree test criterion. The extracted factor demonstrated satisfactory internal consistency. Internal consistency was determined by examining the covariance matrix of the factor score and by computing Cronbach's alpha based on a composite score of variables forming the factor (Tabachnick & Fidell, 1996). The squared multiple correlation (SMC) of the factor score was acceptable to high, SMC = .71, and \( \alpha = .69 \).

The factor loading from the Shapes task indicated statistical significance (see Table 5). All factor loadings demonstrated practical significance (Hair et al., 1995; Stevens, 1992; Tabachnick & Fidell, 1996). Analysis of the factor matrix indicated that the factor measures the components embodied in behavioral inhibition. The go/no-go tasks, RTT task, and interference task all seem to tap underlying abilities to inhibit prepotent responses through the initiation and maintenance of interference control. All measures met the conditions for drawing upon inhibition of prepotent responses, as outlined by Barkley (1997a): preexisting or conditioned prepotency of responses; conflict between this prepotency and task demands; and, temporal proximity between the signals for inhibition and the actual responses (p. 68). The statistically significant loading of the Shapes task on it also illuminates the nature of the factor. This high loading of the interference task on this factor indicates the importance of the maintenance of task rules "on line" in conscious awareness, the resistance to distractions within the tasks, and children's self-monitoring of on-going responses. These abilities are presumed to be facilitated by interference control, and most likely, working memory functions (Barkley, 1997a, 1998). Analysts of the factor matrix also indicated that the measure that appears to tax interference control the least, the Snack Delay task, demonstrated the factor loading of weakest magnitude. Because of the strong prepotency of responses and the influence of interference control during these tasks, this factor was labeled Behavioral Inhibition.1

### Table 4. Matrix of Zero-Order Correlations and Partial Correlations between Behavioral Inhibition Measures

<table>
<thead>
<tr>
<th>Measures</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Snack Delay</td>
<td>1</td>
<td>.25*</td>
<td>.22</td>
<td>.30**</td>
<td>.22**</td>
</tr>
<tr>
<td>2. Tower</td>
<td>.101</td>
<td>.22</td>
<td>.27**</td>
<td>.19</td>
<td></td>
</tr>
<tr>
<td>3. Dog and Dragon</td>
<td>.15</td>
<td>.22</td>
<td>.17</td>
<td>.39**</td>
<td>.32**</td>
</tr>
<tr>
<td>4. Pinball</td>
<td>.17</td>
<td>.27</td>
<td>.39**</td>
<td>.17</td>
<td></td>
</tr>
<tr>
<td>5. Shapes</td>
<td></td>
<td>.24</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:** Data from behavioral inhibition tasks represent transformed variables. Coefficients in parentheses represent partial correlations between behavioral inhibition measures with the effects of age removed.

\(^* p < .05\), \(^** p < .01\), \(^*** p < .001\).
Table 5. Factor Matrix for Behavioral Inhibition Measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Factor 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shapes</td>
<td>.21</td>
</tr>
<tr>
<td>Tower</td>
<td>.00</td>
</tr>
<tr>
<td>Dog and Dragon</td>
<td>.54</td>
</tr>
<tr>
<td>Pinball</td>
<td>.51</td>
</tr>
<tr>
<td>Snack Delay</td>
<td>.48</td>
</tr>
</tbody>
</table>

Intercorrelations between individual behavioral inhibition tasks, the Behavioral Inhibition factor score, and raw scores from the Hyperactivity, Aggression, and Attention Problems subscales of the BASC Teacher Rating Scale (Reynolds & Kamphaus, 1992) are presented in Table 6. Pair-wise correlations indicated significant relationships between ratings of Aggression and the Snack Delay, the Pinball, and the Shapes tasks, \( r(32) = .39, p = .02 \), \( r(32) = .46, p = .006 \), and \( r(32) = .45, p = .008 \), respectively. Significant correlations were also found between ratings of Attention Problems and the Snack Delay and the Shapes tasks, \( r(32) = .36, p = .04 \) and \( r(32) = .39, p = .02 \), respectively. Due to sizable and significant relationships between age and the behavioral inhibition task scores, semipartial correlations were computed to remove the effects of age from these relationships. Four of five initially significant relationships remained significant. Age also appeared to serve as a suppressor variable for several variables. Seven of 12 semipartial correlations were higher in magnitude than the original correlation coefficient. Although the initial correlation was not significant, the correlation between Hyperactivity and the Shapes task became statistically significant, \( r(31) = .32, p = .04 \). The factor score, Behavioral Inhibition, correlated with ratings of Aggression and Attention Problems, \( r(32) = .52, p < .01 \) and \( r(32) = .42, p < .01 \), respectively, but did not significantly correlate with ratings of Hyperactivity. Its correlations with Aggression and Attention Problems remained significant after the effects of age were removed, \( sr = .45, p < .01 \) and \( sr = .48, p < .01 \), respectively.

The Behavioral Inhibition factor score was also used to examine age and sex differences (see Table 7). Although homogeneity of variance across groups was not demonstrated, \( F_{max}(df=69, k=6) = 3.14, p < .01 \), Levene’s test of equality of variances indicated that error variance was equal across groups, \( F(df1=5, df2=64) = .84, p = .53 \). A 2 by 3 factorial analysis of variance (ANOVA) using sex and age groups as independent variables was conducted. Due to unequal group sizes, estimated marginal means were calculated, and a corrected model was examined. The ANOVA indicated significant main effects for sex and age. \( F(1, 70) = 12.75, p < .001 \) and \( F(2, 70) = 6.36, p = .01 \), respectively. The interaction between independent variables was nonsignificant, \( F(2, 70) = 1.95, p = .15 \). Girls demonstrated significantly greater levels of behavioral inhibition than boys, \( F(1, 70) = .64, p = .01 \). Post-hoc analysis using the Games-Howell test (Cone & Foster, 1993) indicated 5-year-old children demonstrated greater behavioral inhibition than 3- and 4-year-old children. \( M \) difference = -1.05, \( p < .001 \) and \( M \) difference = -.57, \( p = .04 \), respectively. Differences in performance between 3- and 4-year-old children were nonsignificant, \( M \) difference = .50, \( p = .20 \).

Table 6. Matrix of Correlations between the Behavioral Inhibition Factor Score, Behavioral Inhibition Measures, and Teacher Ratings

<table>
<thead>
<tr>
<th>Measures</th>
<th>Hyperactivity</th>
<th>BASC Scale</th>
<th>Attention Problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Behavioral Inhibition Factor</td>
<td>.18</td>
<td>.52**</td>
<td>.42*</td>
</tr>
<tr>
<td>2. Shapes</td>
<td>.27</td>
<td>.45**</td>
<td>.48**</td>
</tr>
<tr>
<td>3. Snack Delay</td>
<td>.34**</td>
<td>.37**</td>
<td>.41**</td>
</tr>
<tr>
<td>4. Pinball</td>
<td>.23</td>
<td>.39*</td>
<td>.38*</td>
</tr>
<tr>
<td>5. Tower</td>
<td>.24</td>
<td>.38**</td>
<td>.38**</td>
</tr>
<tr>
<td>6. Dog and Dragon</td>
<td>.29</td>
<td>.46**</td>
<td>.28</td>
</tr>
</tbody>
</table>

Note: Data from behavioral inhibition tasks represent transformed variables. Coefficients in parentheses represent semipartial correlations between age-corrected scores from behavioral inhibition measures and teacher ratings.

* Scores from the Pinball task were not significantly correlated with age. Therefore, zero-order correlations were computed.

* \( p < .05 \), ** \( p < .01 \).
Table 7. Factorial Analysis of Variance of Scores on Behavioral Inhibition Factor by Sex and Age

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>Ratio</th>
<th>Prob. Level</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main Effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>2</td>
<td>12.75</td>
<td>6.10</td>
<td>12.20</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Sex</td>
<td>1</td>
<td>3.04</td>
<td>3.04</td>
<td>6.36</td>
<td>.01</td>
</tr>
<tr>
<td><strong>2-Way Interaction</strong></td>
<td></td>
<td>1.87</td>
<td>.93</td>
<td>1.86</td>
<td>.15</td>
</tr>
<tr>
<td>Age x Sex</td>
<td>2</td>
<td>18.28</td>
<td>9.15</td>
<td>7.64</td>
<td>&lt; .001</td>
</tr>
<tr>
<td><strong>Explained</strong></td>
<td>54</td>
<td>30.03</td>
<td>4.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Residual</strong></td>
<td>72</td>
<td>49.91</td>
<td>49.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Discussion**

The study offers an important step toward establishing psychometrically sound measures of behavioral inhibition for use in clinical practice with preschool-age children. First, it provides supporting evidence for Barkley's theory concerning the components of behavioral inhibition and their interrelation in young children (Barkley, 1997a). Tasks included in this study that appeared to measure behavioral inhibition were influenced by the same underlying factor.

Second, individual measures of behavioral inhibition and the Behavioral Inhibition factor score demonstrated strong concurrent relationships with ecologically valid indices of behavior. The scores from three measures and the Behavioral Inhibition factor score were significantly correlated with teacher ratings of aggressive and inattentive behaviors. These relationships support previous findings examining the ecological validity of measures of behavioral inhibition with preschool-age children (Campbell et al., 1982; Campbell et al., 1994; Olson, 1989; Olson & Hoza, 1993; Silverman & Ruggia, 1990, 1992). However, this study provides perhaps the first evidence of the ecological validity of a measure of interference control (i.e., the Shapes task) and a go/no-go task (i.e., the Pinball task) in a sample of preschool-age children. A surprising finding was that the Shapes task was the only measure that was significantly correlated with teacher ratings of hyperactivity, which is considered the hallmark sign of ADHD in young children (Lahey et al., 1994). It may be that in reference to a developmental perspective, the activity level has a broader range of normality or acceptability within the preschool-age range (Campbell, 1985, 1990; Campbell et al., 1982). It may not be until the school-age years that a comparable level of activity may be seen as outside the normal or acceptable range (Byrne, DeWolfe, & Bawden, 1998, p. 64).

Although studies of preschool-age children have demonstrated strong relationships between measures of behavioral inhibition and variables representing disinterested behaviors, such as observations of aggression, inattention, and overactivity in clinic settings (e.g., Campbell et al., 1982, Campbell et al., 1994; Olson & Hoza, 1993), no studies of preschool-age children were found that examined this relationship with caregivers' ratings of hyperactivity.

Third, two measures of behavioral inhibition, the Shapes and the Tower tasks, demonstrated preliminary evidence of satisfactory to good reliability across weeks in small subsamples of children. Reliability is a necessary characteristic for validity (Anastasi, 1988); however, error variance related to the random fluctuation of children's performances across time and practice effects might have accounted for the low reliability coefficients for these measures. Furthermore, it is difficult to establish a strong relationship between variables because of the small sample sizes used for this reliability analysis. The effects of testing, instrumentation, statistical regression, and selection biases may have affected the internal validity of each of the test-retest correlations (Cook & Campbell, 1979). Studies using larger samples are needed to establish the stability of the two measures that demonstrated acceptable reliability before confidence can be placed in their validity. In addition, additional study of the reliability of the remaining measures is needed.

Although there were only five variables entered into the factor analysis and there were strong correlations among all variables, the study's sample size can be considered very poor to poor (Tabachnick & Fidell, 1996). In addition, the use of linear regression to estimate three missing data points may have inflated correlations between variables. Selection biases may have also affected the results because children were not selected by random, stratified, proportionate, or cluster sampling of the population of 3-, 4-, and 5-year-old children. In addition, educational levels of parents and socioeconomic status (SES) were not systematically examined. Over half of the children who participated in the study had parents who were either university staff or students. Thus, the sample may not be representative of the U.S. population. In addition, although all tasks were standardized, no data were gathered regarding the interrater reliability of task scoring. Factors such as experimenter bias (Rosenthal, 1976) may have influenced the scoring of behaviors and administration of tasks.
Additional research is needed to refine the construct of behavioral inhibition. For example, the apparent interrelationships between purported measures of behavioral inhibition may be due to their relationships with other cognitive processes and executive functions such as working memory (Barkley, 1996b, 1997a; Cole et al., 1993; Diamond & Taylor, 1996; Lee, Vaughn, & Kopp, 1983; Rapport et al., 2000) or more global constructs such as delay aversion (Songa-Barke, 1993), general intelligence (Golden et al., 1977; Silverman & Ippolito, 1995, 1997; Vaughn et al., 1984), and compliance to adult requests (Silverman & Ragusa, 1990; Vaughn et al., 1984). In addition, additional evidence supporting the differentiation of the construct of behavioral inhibition from cognitive impulsivity or reflectiveness is needed. To investigate further the construct of behavioral inhibition in preschool-age children, research is needed that examines various purported measures of the construct in large samples of young children. Studies examining the convergent and discriminant validity and the method effects of measures of behavioral inhibition for preschool-age children will clarify these relationships. These measures can be examined through structural equation modeling or a multitrait-multimethod matrix (Campbell & Fiske, 1959; Marsh & Grayson, 1995). Ratings of disinhibited behaviors by teachers and parents, behavioral observations of children in analogue settings, and other laboratory measures of behavioral measures can be included (Barkley, 1991). The tapping task (Cole et al., 1993; Diamond & Taylor, 1996), the Stroop-like day-night task (Gerstadt et al., 1994), sorting tasks for young children (Cole et al., 1993), and standardized and normed continuous performance tests for young children, such as the Gordon Diagnostic System (Gordon, 1987) could be examined. Promising newly standardized and normed measures of behavioral inhibition that were designed for young children include the Statue subtest from the NEPSY (Korkman, Kirk, & Kemp, 1998) and the Conners’ Continuous Performance Test: Kiddie Version (Conners, 2000). To promote the utility of the measures of behavioral inhibition in identifying children with ADHD, studies are needed that examine the discriminant power, incremental validity, and treatment utility of these instruments (Barkley & Grodzinsky, 1994; Harper & Ottinger, 1992; Rapport et al., 2000).

Research has indicated that the core deficit in ADHD is behavioral disinhibition. Psychometrically strong performance-based measures of this construct may promote the early identification of children with deficits in behavioral inhibition and subsequent interventions. This study supplies some initial empirical support for Barkley’s (1997a) conceptualization of behavioral inhibition. It also extends previous research examining self-control in young children and provides one step toward applying Barkley’s model to clinical practice with disinhibited preschool-age children. However, additional research is needed to establish the psychometric properties of measures of behavioral inhibition and to demonstrate their clinical utility in identifying young children at-risk for ADHD.

Acknowledgements

We would like to thank the Indiana State University child care centers for their assistance in completing this research. We extend our thanks to Chris Quarto, Lisa Bischoff, Christy Coleman, William Osmon, and Kenneth Walker for their reviews of previous drafts of this article. This research was completed in partial fulfillment of the requirements of the doctoral degree in School Psychology at Indiana State University. Portions of this article were presented at the 1999 annual convention of the American Psychological Association in Boston, MA.

Notes

Perhaps this factor is more aptly named Interference Control. Until additional research supports the differentiation between the three components of Barkley’s behavioral inhibition, the more conservative label of Behavioral Inhibition, which subsumes the component of interference control, is offered.

References


**Randy G. Floyd** was with the Department of Psychology at Middle Tennessee State University at the time of article submission. **Edward A. Kirby** is with the Department of Educational and School Psychology at Indiana State University.

Correspondence concerning this article should be addressed to Randy G. Floyd, who is now at The University of Memphis; Department of Psychology; Memphis, TN 38152-6400. Electronic mail may be sent to rgfloyd@memphis.edu.