Preliminary construct and concurrent validity of the Preschool Self-regulation Assessment (PSRA) for field-based research

Radiah Smith-Donald*, C. Cybele Raver*, Tiffany Hayes, Breeze Richardson

University of Chicago, IL, United States

Received 24 February 2006; received in revised form 23 December 2006; accepted 23 January 2007

Abstract

This study introduces a portable direct assessment of young children’s self-regulation—the Preschool Self-Regulation Assessment (PSRA). The PSRA was designed to assess self-regulation in emotional, attentional, and behavioral domains by using a brief, structured battery of tasks in conjunction with a global report of children’s behavior. Factor analyses from a pilot sample (N = 63) of Head Start children revealed two self-regulation factors reflecting children’s performance on tasks of impulse control and tasks of compliance/executive control. Assessor report of children’s behavior during the assessment was reduced into two additional factors reflecting children’s global attention/impulse control and positive emotion. Moderate correlations between self-regulation factors and children’s social competence, behavior problems, and early academic skills, provide preliminary evidence of the new measure’s validity, and additional empirical evidence for bivariate relations between academic and self-regulatory competence.

Keywords: Direct assessment; Preschool; School readiness; Self-regulation

Recent research suggests that preschoolers’ school readiness is supported by a range of skills that fall under the rubric of “self-regulation” (Blair, 2002; Diener & Kim, 2004; Duncan et al., 2007; Eisenberg, Fabes, Guthrie, & Reiser, 2000; Raver, 2002). For example, a child’s ability to manage her emotions in exciting, frustrating, or distressing situations appears to affect the way she is perceived by teachers and peers (Ladd, Birch, & Buhs, 1999; Schultz, Izard, Ackerman, & Youngstrom, 2001). Children’s ability to process new information and to develop learning strategies has been linked with their regulation of attention (Blair, 2002; Fantuzzo, Perry, & McDermott, 2004; Howse, Lange, Farran, & Boyles, 2003; McClelland, Morrison, & Holmes, 2000; Zelazo, Muller, Frye, & Marcovitch, 2003). In addition, preschoolers’ behavior regulation, including their compliance and impulse control, may affect student success as children are expected to follow classroom rules and teachers’ directions, share toys, and wait their turn (McClelland et al., 2000; Mischel, Shoda, & Rodriguez, 1989; Olson & Hoza, 1993). Emotion, attention, and behavior regulation appear to represent three important domains of development; each domain potentially overlapping with and influencing the other two (Blair & Peters, 2003; Carlson, 2003; Hughes, White, Sharpen, & Dunn, 2000; Olson, Bates, Sandy, & Schilling, 2002; Rothbart, Ahadi, Hersey, & Fisher, 2001).
However, this diverse array of self-regulatory skills does not easily fit within a single methodological framework, presenting significant challenges to research in early intervention and early education and care (Aber, Jones, & Cohen, 2000; Carter, Briggs-Gowan, Jones, & Little, 2003; McCabe, Rebello-Brito, Hernandez, & Brooks-Gunn, 2004; Raver & Zigler, 1997). In the following study, we address this challenge by testing the reliability and validity of a new multi-component assessment of children’s self-regulatory skills. First, however, we briefly review the three relevant areas of developmental research – emotion, attention, and behavior regulation – in order to provide an integrative theoretical framework for our new measures. We then discuss the costs and benefits associated with different methodologies and outline a number of potential contributions that a direct assessment of children’s self-regulatory skills offers to research in early education and care.

1. Emotion, attention, and behavior: Three domains of self-regulation

Self-regulation involves modulating systems of emotion, attention, and behavior in response to a given situation or stimulus (Calkins & Fox, 2002; Carlson, 2003; Eisenberg, Smith, Sadosky, & Spinrad, 2004). This includes managing emotions, shifting or focusing attention, and both inhibiting and activating behaviors. Research on infant temperament suggests that young children differ from one another in their biobehavioral reactivity, with variations in threshold, intensity, and duration of autonomic responses even when faced with the same stimuli (Calkins & Fox, 2002; Eisenberg & Fabes, 1992; Keenan, Grace, & Gunthorpe, 2003; Lopez, Vazquez, & Olson, 2004; Posner & Rothbart, 2000; Snidman, Kagan, Riordan, & Shannon, 1995). Research on self-regulation considers individual differences in the emotion, attention, and behavior regulatory strategies that children use to exert control over involuntary, reactivity-based responses (Calkins & Fox, 2002; Derryberry, 2002).

Children’s ability to manage and modulate emotions has been researched via their expression of positive and negative affect, their response to stressful situations, and their ability to monitor others’ emotional states (Calkins, Gill, Johnson, & Smith, 1999; Cole, Martin, & Dennis, 2004; Cole, Zahn-Waxler, Fox, Usher, & Welsh, 1996; Garner & Spears, 2000). In particular, preschoolers’ emotion regulation is correlated with their concurrent social competence, and their regulation, expression, and knowledge of emotion are predictive of later social and academic competence (Denham et al., 2003; Izard et al., 2001; Schultz et al., 2001). Emotional competence may contribute indirectly to learning, where children who are more emotionally positive, who remain engaged, and who avoid negative feelings in the face of academic challenges, tend to have higher grades and higher standardized test scores (Gumora & Arsenio, 2002). Ladd et al. (1999) report that kindergartners’ prosocial and cooperative behavior is associated with later academic achievement, further highlighting the potential importance of children’s positive emotions.

A preschooler’s ability to handle new, emotional, and stressful situations appears to also rely heavily on the development of attention regulation systems (Blair, Granger, & Razza, 2005; Derryberry & Rothbart, 1997; Raver, 1996). The ability to purposely shift and sustain visual attention provides children with new ways to cope with stressors and other stimuli (Ruff & Rothbart, 1996). These same attentional skills also help children divert or focus their attention in order to better handle conflicting stimuli, retain information, and plan their next actions; that is, to exhibit executive control or executive functioning (Chang & Burns, 2005; Derryberry, 2002; Posner & Rothbart, 2000; Zelazo et al., 2003). In addition, executive control has been explored as an aspect of effortful control, which is defined as the ability to suppress a dominant response in order to initiate a subdominant response (Eisenberg, Spinrad et al., 2004).

Much of the research on effortful control focuses on children’s behavioral self-regulation. For example, delay-of-gratification tasks tap children’s ability to stave off frustration, to wait, and to control their impulses (Eisenberg et al., 2003; Mischel et al., 1989; Murray & Kochanska, 2002). Children with higher levels of impulse control have been found to be at significantly lower risk for later adjustment problems than children who struggle to manage their behavior (Olson et al., 2002; Raver, 2002). Finally, children face situations where they must work hard not only to resist impulses, but also to demonstrate that they have internalized adults’ rules, and to act in compliant rather than defiant ways (Eisenberg & Spinrad, 2004; Kochanska, Coy, & Murray, 2001; Valiente et al., 2003). To assess preschoolers’ compliance, “clean-up” or “don’t touch” tasks (Brumfield & Roberts, 1998; Kochanska & Aksan, 1995; NICHD, 1998) are often used in concert with adult-report.

Do these dimensions of regulatory behavior reflect different underlying regulatory mechanisms, or are they different manifestations of a single self-regulatory factor? One drawback of much of the previous research on self-regulation and student achievement is that few studies include all three components of children’s self-regulation. Studies which
examine multiple constructs collect data in different modalities from different reporters (e.g., direct observation by a research assistant for one construct and teacher report for the other), making it difficult to interpret estimates of the covariance between constructs.

2. Assessing young children’s self-regulatory skills

Extant measures of child socioemotional and behavioral regulation include structured and semi-structured lab assessments, unstructured “natural” observation, teacher report, and parent report. Structured lab assessments provide a means of comparing children’s responses to the same situation. For example, separation, disappointment, and frustration paradigms (e.g., Cole, Zahn-Waxler, & Smith, 1994; Garner & Power, 1996; Molitor, Mayes, & Ward, 2003) allow researchers to directly observe children’s emotional responses and regulatory strategies in the face of specific stressors. In some cases assessor-report measures have been linked with developmental assessments to provide information on children’s affect, engagement, attention, and compliance (Bayley, 2005; Roid & Miller, 1997). While lab-based assessments have been used to collect comprehensive data concerning children’s emotional, attentional, and behavioral control, they often involve lengthy protocols, extensive assessor and coder training, and may require specialized materials (Blair, 2003; Kochanska, Murray, Jacques, Koenig, & VanDeveer, 1996; Pfeifer, Goldsmith, Davidson, & Rickman, 2002; Wakschlag et al., 2005).

In contrast, unstructured observation measures take advantage of actual regulatory challenges that children face daily (Bradley & Caldwell, 1979; Buckley, Klein, Durbin, Hayden, & Moerk, 2002; Denham et al., 2003; Fantuzzo, Hightower, Grim, & Montes, 2002). However, comparisons between children are complicated by situational variation as children are assessed while they interact with parents, teachers, and peers who may differ significantly from student to student. Similarly, teacher and parent report are context-specific. A major advantage of report measures is that they do not require expensive materials, participant access to university laboratories, or highly trained data collectors, and they allow information to be collected from large samples in a relatively short amount of time (Achenbach, 2000; Briggs-Gowan & Carter, 1998; Fantuzzo et al., 1995; Gresham & Elliot, 1990; LaFreniere & Dumas, 1996; Rothbart et al., 2001; Zill & Peterson, 1986). In addition, parents and teachers have extensive contact with a given child, and offer valuable insight into the ways that children regulate their emotions, attention, and behavior. However, these reporters often differ in expectations, experience, and personality which may lead to varied assessments of a target child’s success versus difficulty (Burchinal & Cryer, 2003).

Given the different benefits and drawbacks of each approach, researchers have noted that the best strategy is to use multiple methods and multiple reporters (Cole et al., 2004; Denham et al., 2003; Fantuzzo et al., 2004; Nelson, Robinson, & Hart, 2005; Raver, 2004). But this advice has been difficult to follow in survey research. Given that randomized interventions and dynamic models of children’s academic and self-regulatory trajectories require large sample sizes for adequate statistical power, the field faces new methodological challenges. We clearly need to develop affordable direct assessments of children’s self-regulation (Li-Grining, Pittman, & Chase-Lansdale, 2003; McCabe et al., 2004; Raver, 2004; Winston et al., 1999; Zill et al., 2003). In fact, this effort has begun and individual tasks have already been used successfully in field settings (Blair et al., 2005; Li-Grining, Votruba-Drzal, Bachman, & Chase-Lansdale, 2006; Raver, Blackburn, Bancroft, & Torp, 1999).

Finally, adapting lab-based, self-regulation tasks for field research allows us to gain a clearer understanding of the normative development of low-income children who might not otherwise participate in university research. Self-regulatory competence may play a particularly important, mediating role for children facing higher socioeconomic risk (Evans & English, 2002; Gilliom, Shaw, Beck, Schonberg, & Lukon, 2002; Lengua, 2002; Raver & Spagnola, 2003; Rodriguez, Mischel, & Shoda, 1989). Low-income children’s self-regulatory skills are associated with lower levels of distress, higher social and academic functioning, and predict greater emotional and behavioral self-control in the future (Chang & Burns, 2005; Garner & Spears, 2000; Miller, Gouley, Seifer, Dickstein, & Shields, 2004; Schultz et al., 2001; Shaw, Keenan, & Vondra, 1994). Given its hypothesized role as a mediator and moderator, a number of theoretically driven interventions have tried to modify children’s self-regulation skills in an effort to improve outcomes for children at risk for emotional and behavioral difficulty (see Brotman, Gouley, Klein, Castellanos, & Pine, 2003; Izard, Fine, Mostow, Treantacosta, & Campbell, 2002; and Raver, 2002 for reviews). A comprehensive, structured measure of preschoolers’ self-regulation validated with an at-risk, ethnic-minority population can contribute to future research in both developmental psychology and education policy.
3. Study aims and hypotheses

To address this methodological gap, the following study provides preliminary evidence for a reliable and relatively quick means of collecting valid, structured data on children’s self-regulation. The aims of the study were as follows:

1. Develop a structured assessment of preschoolers’ self-regulatory skills that is reliable, viable for field research, and taps natural variation across children. Specifically, our aim is to develop a measure such that: (a) assessors can be trained in a short amount of time to collect reliable data on-site and in person, (b) the administration procedure is brief and uses materials which are inexpensive and portable, and (c) performance varies across children. Through careful task selection, development of a comprehensive, structured script, and rigorous assessor certification processes we expect the new assessment will meet these criteria.

2. Evaluate the proposed three-component structure of self-regulation using exploratory factor analysis. We predict that children’s performance on the new assessment will covary such that the data can be reduced into three overarching factors reflecting regulation of positive and negative emotion, attention, and behavior.

3. Assess concurrent validity of our new assessment with measures of school readiness. We predict that the three elements of self-regulation will be significantly correlated with concurrently assessed behavior problems, social competence, and early academic skills. Specifically, less negative emotion and more positive emotion, greater attention regulation, and greater behavior regulation will each be associated with fewer behavior problems, greater social competence, and more advanced academic skills.

Before moving on, there are several caveats to keep in mind. The causal linkages between self-regulation and school readiness are no doubt complex and bidirectional. For example, inattention and low impulse control may hinder a student’s ability to efficiently absorb new information in a classroom setting, while difficulties in information processing may lead a child to fidget and shift his attention away from classroom instruction. We will set these possible causal linkages aside in this paper. Finally, given the relatively small sample size of this preliminary round of data collection, we lack the statistical power to fully explore measurement and model equivalence across race, language, or age. These analyses are reserved for future studies, when data collection on the larger intervention project is complete.

4. Method

4.1. Participants

Two Head Starts in Chicago were selected as demonstration sites from the set of sites surveyed for the Chicago School Readiness Project (CSRP). One site serves communities comprised largely of Hispanic families while the second site serves communities which are majority Black.

Eighty-eight children aged 41–70 months were recruited for the pilot study. Of these, 64 (73%) children were consented by their parents to participate. One child completed the assessment over 2 days due to the family’s scheduling difficulties and thus was eliminated from data analyses. Therefore, complete data was collected for 63 children (54% girls, n = 34; 46% boys, n = 29; age M = 60.5 months, SD = 7.03) across both sites. The majority of participating children were Black or Hispanic (67% Hispanic, 25% Black, 5% Biracial, and 3% Other). Hispanic children were over-represented and White children were underrepresented as compared to Head Start enrollment across the nation: Hispanic, 31.2%; Black 31.1%; White, 26.9%; Multi-Racial/Other = 5% (ACF, 2005). Families were struggling financially with mean monthly income of $1392 which translates into an annual income of $16,705. In addition, 86% of families had only 1 adult living in the home.

4.2. General procedure

Data on children’s self-regulation, early academic skills, and behavior problems and competencies were collected using direct assessment and adult report. Assessors completed training and certification regarding administration of both early academic tasks and self-regulation tasks. Certified assessors administered the battery of early academic tasks first followed by the self-regulation tasks to children in a quiet area at the site of their Head Start programs. Assessor reports were completed by the assessor immediately after returning the child to the classroom. All 63 children were
Table 1
PSRA tasks

<table>
<thead>
<tr>
<th>Task</th>
<th>Description/Instructions for child</th>
<th>Measurement method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toy Wrap</td>
<td>Child was asked not to peek while the assessor noisily wrapped a “surprise”</td>
<td>Latency to first peek (s)</td>
</tr>
<tr>
<td>Toy Wait</td>
<td>Child was directed to wait without touching the wrapped “surprise”</td>
<td>Latency to touch surprise (s)</td>
</tr>
<tr>
<td>Snack Delay</td>
<td>Child was instructed to wait for a signal before “finding” an M&amp;M under a clear cup</td>
<td>Level of waiting (1–4)</td>
</tr>
<tr>
<td>Tongue Task</td>
<td>Child and assessor wait with an M&amp;M on their tongues to see who will eat it first</td>
<td>Latency to eat M&amp;M (s)</td>
</tr>
<tr>
<td>Balance Beam</td>
<td>After walking a long a line once, child was directed to walk the same line slowly</td>
<td>Difference between slow and regular trials (s)</td>
</tr>
<tr>
<td>Tower Task</td>
<td>Child was instructed to take turns with the assessor placing blocks to build a tower</td>
<td>Level of turn-sharing (0–1)</td>
</tr>
<tr>
<td>Pencil Tap</td>
<td>Child was to tap once when assessor tapped twice, and tap twice when assessor tapped once</td>
<td>Percent of correct responses (%)</td>
</tr>
<tr>
<td>Tower Clean up</td>
<td>Child was instructed to clean up the blocks from Tower Task</td>
<td>Latency to complete clean up (s)</td>
</tr>
<tr>
<td>Toy Sort</td>
<td>Child was asked to sort and put away small toys without playing with them</td>
<td>Latency to complete clean up (s)</td>
</tr>
<tr>
<td>Toy Return</td>
<td>Child was asked to return a fun toy after a brief period of play</td>
<td>Latency to return toy (s)</td>
</tr>
</tbody>
</table>

assessed within a 6-week timeframe. Family background data was collected at the time of project enrollment 6 months earlier. Teacher report data was collected over a 2-week time period which began 1 week after direct assessments were completed.

4.3. Self-regulation measures

Self-regulation was assessed using two instruments: the PSRA battery and an assessor report of children’s self-regulation that could potentially be used in any test-taking situation. Bilingual children were assessed in their home language. About one-third (36.5%, n = 23) of assessments were videotaped and double-coded to gauge interrater reliability on each self-regulation task. The same 23-videotaped assessments were used to assess interrater reliability on each construct of the assessor report.

4.3.1. PSRA battery

The Preschool Self-Regulation Assessment (PSRA) was administered by six trained and certified assessors. Following 2 days of training, assessors were certified once they met select administration criteria on two separate practice assessments (Goyette et al., 2006). Assessors were majority female and all were ethnic minorities with at least 4 years of college (67% female; 50% Hispanic, 33% black, and 17% biracial (Asian-American/white); 33% bachelor’s degree, 50% master’s degree, 17% > master’s degree).

The 10 PSRA tasks were selected because they were brief, required few materials, yet yielded useful data for 3- to 5-year-old children in lab-based protocols. Assessors live-coded latencies or performance levels for each task. Task descriptions and measurement methods are outlined in Table 1. Four delay tasks were adapted from the lab-based work of Kochanska and colleagues and were designed to tap children’s effortful control: Toy Wrap, Toy Wait, Snack Delay, and Tongue Task (see Murray & Kochanska, 2002). Two additional tasks from Kochanska’s lab were included as tasks of executive control because they require children to filter competing stimuli: Balance Beam and Tower Task (Maccoby, Dowley, Hagen, & Degerman, 1965; Murray & Kochanska, 2002). Pencil Tap, the third executive control task, was adapted from the peg-tapping task (Blair, 2002; Diamond & Taylor, 1996). Finally, three “do”-tasks were included to assess children’s compliance: Tower Cleanup, Toy Sort, and Toy Return (see, e.g., Brumfield & Roberts, 1998; NICHD, 1998). In addition to performance data, assessors noted the presence or absence of certain behaviors during each task (e.g., child is defiant, child engages assessor). These codes were averaged across tasks and used as global ratings in conjunction with the assessor report (described below).

4.3.2. Adapted Leiter-R assessor report

The framework and descriptors of the 28-item assessor report were adapted from the Leiter-R social-emotional rating scale (examiner version; Roid & Miller, 1997) and the Disruptive Behavior-Diagnostic Observation Schedule coding system (DB-DOS; Wakschlag et al., 2005) and provided a global picture of children’s emotions, attention, and behavior throughout the assessor-child interaction.
Fifteen items were selected from the following Leiter-R subscales: attention (e.g., pays attention during instructions), impulse control (e.g., thinks and plans before beginning), activity level (e.g., remains in seat), sociability (e.g., alert and interactive), and energy and feelings (e.g., shows pleasure in accomplishment). Two more items were adapted from the Leiter-R but do not map on to a specific original item: “has difficulty waiting between tasks” and “modulates and regulates arousal in self”. In addition, nine items were included from the DB-DOS coding system and reference defiance and noncompliance, the frequency and intensity of negative and positive affect, and the presence or absence of verbal or physical aggression. Finally, two items were included to capture children’s anxiety during the assessment (“shows frequent/intense feelings of worry”). All assessor report items included descriptors with clear, behavioral anchors as used in the DB-DOS coding system. Items were coded using a Likert scale ranging from 0 to 3 and some items were reverse-coded to minimize automatic responding.

4.4. Early academic skills

The NRS is a brief direct assessment designed to capture 4- and 5-year-olds’ early verbal, literacy, and math skills (ACF, 2003). It includes a brief screener which determines whether the bilingual child (i.e., home language is Spanish) is tested in Spanish only or in Spanish and English. The pilot sample had 25 bilingual children (39%) all of whom were assessed in both languages. For each subscale, the higher score was used in subsequent analyses. All other children were assessed once in English. The Vocabulary subscale was adapted from the Peabody Picture Vocabulary Test (PPVT-III; Dunn & Dunn, 1997) and taps children’s early verbal skills. Early literacy skills are assessed via letter naming. The Early Math subscale includes recognition of numbers and shapes, counting, size concepts, simple arithmetic, and basic chart reading and was adapted from the Math assessment of the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K). Administration of the NRS battery takes about 20 min.

4.5. Teacher report of behavior

Six teachers completed the short form of the Social Competence and Behavior Evaluation (SCBE-30; LaFreniere & Dumas, 1996) and the Behavior Problems Index (BPI; Zill & Peterson, 1986) for the participant children in their classrooms. Teachers’ race/ethnicity matched the predominant race/ethnicity of the children at a given site. The SCBE-30 is designed to provide information on the socio-emotional adjustment of children 3–6 years old. It provides social competence (e.g., “works easily in groups”, “shares toys with other children”; \( \alpha = .88 \)), externalizing (e.g., “easily frustrated”, “defiant when reprimanded”; \( \alpha = .89 \)), and internalizing subscales (e.g., “worries”, “remains apart, isolated from a group of children”; \( \alpha = .75 \)). The BPI is a 30-item adult-report measure which focuses on behavior problems. It was originally designed for parent respondents, but the authors gave permission to use the measure with teachers. It can also be aggregated into externalizing (e.g., “is disobedient”, “bullies or is cruel or mean to others”; \( \alpha = .91 \)) and internalizing (e.g., “is too fearful or anxious”, “clings to adults”; \( \alpha = .85 \)) subscales. Although these two measures overlap, they were analyzed separately to provide a clear picture of the new measures’ concurrent validity with previously validated measures. Teachers were given $20.00 per packet upon completion.

4.6. Background Information

Family interviews were extensive and included the CSRP Health and Demographics questionnaire. Interviews took 35–45 min and families were compensated with $10 and a bag of children’s books. Data for families who could not be interviewed face-to-face was gathered with parent permission from administrative files at the Head Start site.

5. Results

5.1. Data analysis

5.1.1. Latent factor structure

A primary goal of this study is to determine whether lab-based assessments can be translated to field contexts and still yield similar results to those found in lab contexts. As a result, exploratory factor analyses were conducted independently
for the PSRA tasks and assessor report items. Given the sample is small and not nationally representative, principal component extraction was used. Therefore the results provide a description of the sample, but cannot necessarily be generalized to the population at large. Resulting components were rotated obliquely using Direct Oblimin to allow correlation between factors. Cronbach’s alpha was calculated for each emergent construct and provides an index of internal consistency based on the average inter-item correlation within each factor. In addition, we report interrater reliability at the task level for the PSRA and at the subscale level for the assessor report using intra-class correlation (ICC) for continuous variables and Cohen’s kappa for categorical variables.

5.1.2. Concurrent validity

Factor analyses informed the composition of self-regulation variables; items with loadings greater than .4 were aggregated to form each construct. Correlation analyses were used to examine concurrent validity of PSRA-based constructs with previously validated measures of children’s social competence and behavior problems (SCBE and BPI) as well as with their early academic skills on the NRS.

5.2. Descriptive results

Reliability and descriptive statistics are presented in Table 2. Where possible, the scoring methods used in previous research were applied. We fulfilled the study’s first aim by developing an assessment which was administered in a single, 20-min session to each child. Interrater reliability between the live coder (assessor) and a reliability coder using a videotape was high on all tasks. Interrater reliability on the assessor report subscales (identified below) was also high (ICC > .8). Finally, as Table 2 shows, there was wide variability in children’s performance for most tasks.

5.3. Factor analyses

Several items on the adapted Leiter-R assessor report were observed at very low frequencies within the pilot sample resulting in overly skewed distributions (>2.0). These items were excluded from factor analyses. Factor analysis on the remaining 19 items gave three components with eigenvalues >1 reflecting attention regulation, impulse control, and positive emotion. However, several items loaded well onto both the attention regulation factor and the impulse control factor. Given the degree of overlap between these two components, factor analysis was rerun restricting the outcome to 2 factors. The restricted analysis collapsed the two overlapping components and resulted in a loss of 10% explained variation. The final structure consisted of (a) Attention/Impulse Control and (b) Positive Emotion, and explained 53.4% of the variation (Table 3). Two items (waiting and mood regulation) were excluded from both constructs as

---

**Table 2**

Descriptive statistics for PSRA (N=63)

<table>
<thead>
<tr>
<th>Task</th>
<th>M</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Reliabilitya</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balance Beam</td>
<td>2.0 s</td>
<td>3.67</td>
<td>−7.0</td>
<td>13.5</td>
<td>0.98</td>
</tr>
<tr>
<td>Pencil Tap</td>
<td>58.8%</td>
<td>0.37</td>
<td>0%</td>
<td>100%</td>
<td>1.00</td>
</tr>
<tr>
<td>Tower turns</td>
<td>0.67</td>
<td>0.48</td>
<td>0</td>
<td>1</td>
<td>0.91b</td>
</tr>
<tr>
<td>Tower Cleanup</td>
<td>43.0 s</td>
<td>21.09</td>
<td>14s</td>
<td>N (&gt;120 s)</td>
<td>0.98</td>
</tr>
<tr>
<td>Toy Sort</td>
<td>78.8 s</td>
<td>31.66</td>
<td>18s</td>
<td>N (&gt;120 s)</td>
<td>0.96</td>
</tr>
<tr>
<td>Toy Wrap – peek</td>
<td>45.7 s</td>
<td>23.13</td>
<td>0.0s</td>
<td>N (&gt;60 s)</td>
<td>0.90</td>
</tr>
<tr>
<td>Toy Wrap – touch</td>
<td>49.2 s</td>
<td>21.44</td>
<td>0.0s</td>
<td>N (&gt;60 s)</td>
<td>0.81</td>
</tr>
<tr>
<td>Toy Return</td>
<td>9.8 s</td>
<td>21.72</td>
<td>0.0s</td>
<td>N (&gt;120 s)</td>
<td>1.00</td>
</tr>
<tr>
<td>Snack Delay</td>
<td>3.6</td>
<td>0.59</td>
<td>2</td>
<td>4</td>
<td>0.84b</td>
</tr>
<tr>
<td>Tongue Task</td>
<td>36.7 s</td>
<td>9.75</td>
<td>0.0s</td>
<td>N (&gt;40 s)</td>
<td>0.91</td>
</tr>
<tr>
<td>Sample Size</td>
<td>63</td>
<td></td>
<td></td>
<td></td>
<td>23</td>
</tr>
</tbody>
</table>

Note: N = did not occur within time frame.

a Reliability calculated using intra-class correlation unless otherwise noted.

b Cohen’s kappa.
they did not load well onto either component; no items loaded simultaneously onto both factors. The final constructs were internally consistent as evaluated using Cronbach’s alpha (Attention/Impulse Control, \( \alpha = .89 \); Positive Emotion, \( \alpha = .87 \)). In addition, interrater reliability was high at the construct level, ICC = .82 for Attention/Impulse Control and .83 for Positive Emotion.

As with the assessor report, PSRA task scores were checked for excessive skew and truncated range, which are evidence of limited variability. As a result, two tasks were excluded from analyses: Toy Return and Tongue Task. In addition, Tower Task provided binary information on a single item and was not included in factor analyses. Factor analysis of standardized task scores generated two components with eigenvalues >1.0: (a) Impulse Control and (b) Compliance/Executive Control (Table 4). The two component structure accounted for 50.0% of the variance; no tasks loaded above .400 on the two factors simultaneously. Compliance task scores (Toy Sort and Tower Cleanup) were reversed so that faster cleanup times yield higher compliance scores. Internal consistency of the final constructs was moderately low (Impulse Control, \( \alpha = .54 \) and Compliance/Executive Control, \( \alpha = .58 \)).

Table 3
Rotated structure matrix – adapted Leiter-R assessor report (N=63)

<table>
<thead>
<tr>
<th>Item</th>
<th>Component</th>
<th>Attention/Impulse Control(A)</th>
<th>Positive Emotion(A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concentration</td>
<td></td>
<td>.845</td>
<td></td>
</tr>
<tr>
<td>Attention</td>
<td></td>
<td>.780</td>
<td></td>
</tr>
<tr>
<td>Patience/Interruption</td>
<td></td>
<td>.773</td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td></td>
<td>.743</td>
<td></td>
</tr>
<tr>
<td>Distractability</td>
<td></td>
<td>-.721</td>
<td></td>
</tr>
<tr>
<td>Focus</td>
<td></td>
<td>-.714</td>
<td></td>
</tr>
<tr>
<td>Planning Ability</td>
<td></td>
<td>.681</td>
<td></td>
</tr>
<tr>
<td>Seated</td>
<td></td>
<td>.667</td>
<td></td>
</tr>
<tr>
<td>Impulse Control</td>
<td></td>
<td>.655</td>
<td></td>
</tr>
<tr>
<td>Non-Compliant</td>
<td></td>
<td>-.570</td>
<td></td>
</tr>
<tr>
<td>Frequent Positive</td>
<td></td>
<td></td>
<td>.847</td>
</tr>
<tr>
<td>Accomplishment</td>
<td></td>
<td></td>
<td>.837</td>
</tr>
<tr>
<td>Positive across tasks</td>
<td></td>
<td></td>
<td>.824</td>
</tr>
<tr>
<td>Intense Positive</td>
<td></td>
<td></td>
<td>.779</td>
</tr>
<tr>
<td>Positive</td>
<td></td>
<td></td>
<td>.767</td>
</tr>
<tr>
<td>Alert</td>
<td></td>
<td></td>
<td>.762</td>
</tr>
<tr>
<td>Confident</td>
<td></td>
<td></td>
<td>.703</td>
</tr>
<tr>
<td>Waitinga</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mood Regulationa</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Extraction Method: Principal Component Analysis, Rotation Method: Oblimin with Kaiser Normalization.  

\( a \) Items included in factor analysis which did not load onto any factor.

Table 4
Rotated structure matrix for PSRA tasks (N=63)

<table>
<thead>
<tr>
<th>Task</th>
<th>Component</th>
<th>Impulse Control(T)</th>
<th>Compliance/Executive Control(T)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Snack Delay</td>
<td></td>
<td>.848</td>
<td></td>
</tr>
<tr>
<td>Toy Wrap – 1st peek</td>
<td></td>
<td>.761</td>
<td></td>
</tr>
<tr>
<td>Toy Wrap – 1st touch</td>
<td></td>
<td>.439</td>
<td></td>
</tr>
<tr>
<td>Toy Sort</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tower Cleanup</td>
<td></td>
<td>.797</td>
<td></td>
</tr>
<tr>
<td>Balance Beam</td>
<td></td>
<td>.663</td>
<td></td>
</tr>
<tr>
<td>Pencil Tap</td>
<td></td>
<td>.635</td>
<td></td>
</tr>
</tbody>
</table>

Note: Extraction Method: Principal Component Analysis, Rotation Method: Oblimin with Kaiser Normalization.
Table 5
Bivariate correlations between PSRA and adapted Leiter-R constructs

<table>
<thead>
<tr>
<th>Measure</th>
<th>Construct</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adapted Leiter-R</td>
<td>Attn/Impulse Control (A)</td>
<td>−</td>
<td>−</td>
<td>.03</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td>Positive Emotion (A)</td>
<td>−</td>
<td>.03</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>PSRA Tasks</td>
<td>Compliance/Exec Control (T)</td>
<td>.45</td>
<td>.00</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td>Impulse Control (T)</td>
<td>.62</td>
<td>.09</td>
<td>.40</td>
<td>−</td>
</tr>
</tbody>
</table>

Note: N = 63.
** p < .01.

5.4. Construct validity

_t_-Tests comparing mean performance between boys and girls on the assessor report and PSRA were nonsignificant for each construct. In contrast, older children (5 years) were significantly better regulated than younger children (3–4 years) on these two measures (p-value range: .00–.11).

To evaluate the theory that self-regulation is composed of distinct components, bivariate correlations were run between the constructs derived from the assessment—two as reported by assessors on the adapted Leiter-R (labeled [A]) and two based on children’s performance on the PSRA tasks (labeled [T]). Correlation analyses reveal moderate to high positive associations between the four components (Table 5). In particular, the delay tasks were highly correlated with assessor report of children’s attention/impulse control (r = 0.62). Assessor report of children’s global positive emotion was not correlated with any other self-regulation construct.

5.5. Concurrent validity

Concurrent validity with other dimensions of development was also assessed using bivariate correlations. Teacher-reported behavior problems and competencies were modestly to moderately correlated with children’s assessed self-regulation (Table 6). Specifically, Attention/Impulse Control (A) and Compliance/Executive Control (T) were each negatively correlated with teacher-reported behavior problems at modest levels. In addition, Compliance/Executive Control (T) was moderately positively correlated with teacher report of children’s social competence (r = 0.53). In...
Table 7
Bivariate correlations between self-regulation assessment constructs and NRS subtests

<table>
<thead>
<tr>
<th>NRS subtests</th>
<th>PSRA Tasks</th>
<th>Adapted Leiter-R</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Compliance/Exec. Control</td>
<td>Impulse Control</td>
</tr>
<tr>
<td>Early Math</td>
<td>0.08</td>
<td>0.22*</td>
</tr>
<tr>
<td>Letter Naming</td>
<td>0.10</td>
<td>0.23*</td>
</tr>
<tr>
<td>Verbal Skills</td>
<td>0.12</td>
<td>0.38***</td>
</tr>
</tbody>
</table>

Note: N = 63.
* p < .10.
** p < .05.
*** p < .01.

contrast, assessor report of children’s Positive Emotion(A) was not associated with teacher report of behavior problems or competencies.

PSRA constructs were also significantly correlated with children’s early academic skills (Table 7). In particular, children’s performance on Impulse Control(T) tasks, global Attention/Impulse Control(A), and global Positive Emotion(A) were positively associated with children’s early math and verbal skills ($r = 0.22–0.38$). Additionally, children’s Impulse Control(T) was positively associated with their performance on the letter naming subscale.

6. Discussion

Analyses from this study suggest that the PSRA makes several contributions to research in early childhood and education. First, our piloted measures were successfully administered in the field following a brief training process. PSRA tasks were coded reliably and elicited substantial variability in children’s performance, suggesting that Kochanska and colleagues’ landmark, lab-based tasks translate well to field research, as do short tasks of executive control and compliance. Comparison of PSRA data against others’ findings indicate that our new measure yielded similar results to those observed in prior work. For example, children’s performance on the Pencil Tap (percent correct = 58.6%) was only slightly lower than that reported by Blair (2003) for children’s overall executive control (percent correct = 66% across peg-tapping and Stroop tasks). Similarly, children in our pilot sample took about 9.0 s on average to slowly “walk along the balance beam” (having slowed down by 2.0 s on average), which is in line with Kochanska and colleagues’ observation of approximately 10.5 s across studies (Kochanska, Murray, & Coy, 1997; Kochanska et al., 1996; Murray & Kochanska, 2002).

There were several items which failed to induce sufficient variability and which were omitted from analyses. Mean ratings on the assessor report indicated that the children in our low-income pilot sample were typically cooperative and rarely displayed intense negative emotions or aggression. Similarly, results from the PSRA highlight the regulatory tasks that our pilot sample handled particularly well: Almost all children immediately returned the toy and held candy on their tongue for 40 s to “win” the M&M game. A larger sample recruited from a broader range of sites may provide greater variance in future research. Alternatively, administering these tasks in less controlled circumstances outside of the lab may interfere with our ability to tap individual differences in these areas of preschoolers’ self-regulation (McCabe et al., 2004).

Second, factor analyses exploited the covariance between PSRA tasks and between items on the assessor report and suggested multiple, overarching domains which reflect emotion, attention, and behavior regulation. Internal consistency, evaluated via Cronbach’s alpha, was somewhat low for PSRA constructs, perhaps as a result of the low number of items (tasks) used to estimate each construct. Emotion-based data was derived from the assessor report and consisted of positive affect, engagement, and cooperation; negative affect was rarely reported. Information on attention regulation (“concentration”, “focus”) was gathered via the assessor report, while executive control was tapped by the PSRA Pencil Tap and Balance Beam tasks. Finally, behavior regulation, specifically impulse control, was captured by both the adapted Leiter-R and the PSRA delay tasks. Once aggregated, the two components yielded from the PSRA tasks were moderately correlated, indicating different, albeit related, constructs had been tapped. In addition, these constructs overlapped with assessor report of children’s Attention/Impulse Control(A). In contrast, Positive Emotion(A) was not significantly correlated with any other factor.
The assessment also revealed substantial overlap across certain domains. For example, assessors’ global ratings of preschoolers’ attention regulation and impulse control (i.e., behavior regulation) were highly correlated and were collapsed into a single construct. It is unclear whether the measure prevented assessors’ from accurately discriminating fine-grained aspects of self-regulation, or whether the construct items are related to a common latent factor of regulation at this age. One can easily see how these components overlap: Engaging in impulsive behavior sometimes entails inattention to directions or to the task at hand in preference for the impulse. Similarly, lack of attention is often assessed by observing children’s tendency to follow their impulses and engage in other activities.

In addition, we initially classified compliance as a subcategory of behavior regulation, yet the “do”-tasks loaded with executive control, a construct tightly aligned with attention (Chang & Burns, 2005; Zelazo et al., 2003). Again, this can be interpreted in at least two different ways. First, children’s ability to plan and manage conflicting stimuli and their tendency to efficiently carry out adult instructions may be part and parcel of the same construct. Alternatively, it may indicate that we mislabeled our “compliance” tasks. For example, the toy sort task involves following directions, but also calls on children’s ability to organize their actions. Assessors noted, anecdotally, that some highly compliant children took a long time to complete this task because they were making an obvious effort to execute the task carefully and correctly. Therefore, slow task completion may demonstrate weaker executive control skills, rather than noncompliant behavior. Such overlaps between tasks and constructs present an important set of new questions regarding measure refinement and the nature of preschoolers’ self-regulation.

Finally, how are children’s self-regulatory skills related to indicators of early school readiness? While t-test results suggest that assessed self-regulation did not vary by gender, they indicate that older children were better regulated, as previous research would predict (Posner & Rothbart, 2000). In addition, our results suggest that children’s attention and impulse control, evaluated via assessor report and task performance, and their task-based compliance/executive control are associated with their behavior and social skills in the classroom. Children who were reported by assessors as more attentive and less impulsive were less likely to be reported by teachers as impulsive, disobedient, or withdrawn. This was also the case, although to a lesser degree, for children who inhibited their impulses and waited during delay tasks. Attentive and patient behavior on the assessment was also associated with children’s ability to negotiate conflicts and work well with other children as reported by teachers on the SCBE. An even stronger positive association was observed between teacher-perceived social competence and children’s organized and effective task execution—Compliance/Executive Control(T). And these same, well-organized children were also less likely to be reported as anxious, clingy, or withdrawn by their teachers. Overall, the PSRA, which takes children out of the classroom and places them in a highly structured interaction with a stranger, provided a portrait of children’s competence and difficulties that is congruent with that reported by teachers.

Non-significant associations between Positive Emotion(A) and other variables may reveal a limitation of our assessment. Specifically, we did not present paradigms which were intensely emotionally evocative (such as disappointment paradigms, e.g., Garner & Power, 1996) or which encouraged children to elicit emotions from the assessor. In addition, the PSRA does not allow for observation of children’s emotional responses during interactions with teachers and peers (e.g., during play, conflict, or structured class time), times at which emotion regulation is probably most important for school readiness. This suggests that our assessment should be considered as one among multiple tools in assessing preschoolers’ self-regulation in school readiness research, with strengths in some areas (e.g. tapping impulse control) but not in others (e.g. emotion regulation).

Although Positive Emotion(A) was not significantly correlated with teacher-reported behavior, it was modestly correlated with children’s early math and verbal skills. It seems that children who received higher scores on the NRS were more likely to be positive, express confidence, and engage the examiner in conversation than those with lower scores. This corresponds to previous findings concerning the relations between positive emotions and academic proficiency (Gumora & Arsenio, 2002). Children’s performance on delay tasks was also associated with higher NRS scores, possibly reflecting that children who can delay gratification for toys and candy also provide more thoughtful (i.e., correct) responses on cognitive recall tasks. Against predictions, performance on tasks of executive control and compliance were not associated with demonstrated academic skills. Analyses with a larger sample can provide more compelling evidence of the relations between these domains.

In sum, this study presents the first step in the development and validation of a new set of measures: the PSRA and a global assessor report adapted from the Leiter-R. Findings provide support for the PSRA as a short, reliable, multidimensional measure of preschoolers’ self-regulation that is relatively inexpensive to administer in field settings. In addition, our findings suggest that this approach yields data that are psychometrically valid, and this pilot study
replicates previous findings on the associations between self-regulation and student success (Blair, 2002; Gumora & Arsenio, 2002; Howse et al., 2003; Raver, 2002). Although these relations are likely bidirectional, this is a promising area for policy and prevention science. Socioemotional interventions have recently been undertaken with the goal of improving both regulatory and academic outcomes for children in early elementary school, particularly those at risk for school difficulty (August, Realmuto, Hektner, & Bloomquist, 2001; Flay & Allred, 2003; Linares et al., 2005). However, most emotion- and behavior-focused interventions face substantial methodological obstacles when forced to rely on teacher- and parent-reported affect, attention, and impulse control. This study aims to help researchers surmount those obstacles by field-testing a standardized and valid direct assessment of children’s emotional, attentional, and behavioral self-regulation.

It is important to highlight that our findings, though promising, are preliminary. There are many methodological steps ahead. These new measures must be validated with larger, more diverse samples, with careful attention paid to testing their equivalence across multiple groups of children differing by age, gender, and race/ethnicity. Analyses of stability versus change over time in children’s performance on multiple dimensions of self-regulation are also important next steps, and are currently underway. In pursuing these steps, our aim is to provide efficient, high-quality tools to help maximize the field’s understanding of young children’s self-regulation and school readiness.

Acknowledgments

This research was funded by the McCormick Tribune Foundation and the NICHD Interagency Consortium on School Readiness. Special thanks go to the families and staff of the participating Head Start programs. In addition, we gratefully acknowledge the contributions of the staff of the Chicago School Readiness Project and we appreciate the helpful comments of Molly Metzger and Christine Li-Grining.

References


Guilford Press.


