TESTING MODELS OF CHILDREN'S SELF-REGULATION WITHIN EDUCATIONAL CONTEXTS: IMPLICATIONS FOR MEASUREMENT

C. Cybele Raver[*], **Jocelyn Smith Carter**[†], **Dana Charles McCoy**[*], **Amanda Roy**[*], **Alexandra Ursache**[*], and **Allison Friedman**[*]

[*]Department of Applied Psychology, New York University, New York, USA

[†]Department of Psychology, Depaul University, Chicago, Illinois, USA

Abstract

Young children's self-regulation has increasingly been identified as an important predictor of their skills versus difficulties when navigating the social and academic worlds of early schooling. Recently, researchers have called for greater precision and more empirical rigor in defining what we mean when we measure, analyze, and interpret data on the role of children's self-regulatory skills for their early learning (Cole, Martin, & Dennis, 2004; Wiebe, Espy, & Charak, 2008). To address that call, this chapter summarizes our efforts to examine self-regulation in the context of early education with a clear emphasis on the need to consider the comprehensiveness and precision of measurement of self-regulation in order to best understand its role in early learning.

I. Introduction

Consider the challenges that young children face when navigating the academic and the social demands of their preschool or early elementary classrooms. Children must draw upon a host of cognitive, behavioral, and emotional skills in order to pay attention to, remember, and appropriately respond to a teacher's directives. Children must also be able to modulate their own emotions, and to read, and appropriately respond to others' emotions, particularly during heated moments in the lunch room, on the playground, or when feeling anxious or frustrated with peers. In short, doing well in school requires children to not only have the cognitive capacity to learn new material and to demonstrate what they know, but also to regulate their attention, their behavior, and their emotions. In short, self-regulation allows children to experience accomplishment and joy in the process of learning within the classroom. The recognition that cognitive skill represents one, but not the only, ingredient that contributes to children's success in school is probably the reason that attention to "self-regulation," "emotional intelligence," and "executive function" enjoy such prominence in much educational research lately (see Blair, 2010 for review).

What kind of empirical evidence is there to support the claim that self-regulation matters for children's success in school? When it has been broadly defined, children's competent self-regulation is prospectively associated with a large number of later positive outcomes in school settings, including higher social competence, higher academic competence, and lower levels of proneness to externalizing and internalizing problems (Rodriguez et al., 2005).

Conversely, children's cognitive and emotional dysregulation represents an early indicator of emerging behavioral problems for some children, and children's behavioral dysregulation is increasingly targeted for primary prevention and intervention efforts in schools (Greenberg, Kusche, & Speltz, 1991; Webster-Stratton & Taylor, 2001). Despite the recent increase in effort to understand self-regulation's importance in development, this area of research continues to struggle with a lack of consensus on ways that multiple measured indicators may tap the same (or different) subcomponents of children's self-regulation (Blair, Zelazo, & Greenberg, 2005).

The following chapter considers several theoretical constructs and empirical challenges in the study of self-regulation as our field moves from lab-based research to asking key developmental questions in educational settings. We first review what we mean by self-regulation, broadly. We then consider the evidence that emotional and cognitive components of children's self-regulation matter for educational success. Next, we offer ways that children's self-regulation may be influenced by the quality of educational settings by reviewing findings from our own recent intervention research. Finally, we consider ways that cognitive and emotional components of children's self-regulation can be tapped with different types of tools (e.g., by direct assessments as well as through adult reports), across different ages and settings, in order to be able to extend this important line of inquiry at the intersection of developmental science and educational policy research. We argue that these tools are crucial to our ability to answer fundamental questions regarding the plasticity, or malleability of children's self-regulation, as well as regarding the role of self-regulation for children's learning.

A. DEFINITIONS AND DEVELOPMENTAL DOMAINS: WHAT EXACTLY IS SELF-REGULATION?

What exactly do we mean by the term "self-regulation?" Self-regulation can be considered from both behavioral and neural systems levels, and from developmental traditions that have historically devoted greater attention to either (a) emotion-related processes of reactivity and regulation or (b) cognitive processes, including children's working memory and attention deployment (see Blair & Razza, 2007 for comprehensive review of these two disciplinary perspectives). There is consensus across these diverse disciplinary perspectives that children's self-regulation includes their ability to manage or modulate positive and negative emotions, to inhibit or control their behavior, and to shift and focus their attention (Blair, 2002; Cole, Martin, & Dennis, 2004; Eisenberg et al., 1996; Kochanska, Murray, & Coy, 1997; Kopp, 1989; Mischel, Ebbesen, & Zeiss, 1976). For the purposes of this paper, we anchor the discussion of children's behavioral self-regulation in two key domains of functioning, including children's executive functioning and their emotional regulation.

B. SELF-REGULATION ACROSS DOMAINS OF EMOTION AND COGNITION

Executive function is often defined as a set of cognitive processes that helps children remain on-task and goal-directed, including the use of working memory, the ability to sustain and shift attention, and the ability to "inhibit prepotent responding" (also defined as "inhibitory control"; Blair et al., 2005; Carlson, 2005; Espy, Kaufmann, McDiarmid, & Glisky, 1999; Jones, Rothbart, & Posner, 2003). In contrast, emotion regulation can be defined as a set of

intrapsychic and interpersonal processes whereby the intensity, valence, and time course of a child's experienced and expressed emotion is shifted, changed, or maintained (Cole et al., 2004). The construct of emotion regulation is often deployed to help researchers understand ways that children differ not only in their temperamental proneness to experience emotions with greater versus less intensity (termed "reactivity") but also in their capacity to manage or modulate negative emotions in order to be able to stay on-task or to get along with others (Calkins & Fox, 2002).

Given these two clearly contrasting domains of self-regulation, which of them matters most for young children's academic achievement? On the one hand, recent evidence from early cognitive research suggests that children's executive function plays a key role in their chances of later academic success (Blair & Razza, 2007; Diamond, Barnett, Thomas, & Munro, 2007). For example, children who demonstrate greater inhibitory control on executive function measures in preschool have been found to do better on both reading and math assessments in early elementary school (see Bull, Espy, & Weibe, 2008, for review). On the other hand, recent research on young children's socioemotional development highlights the additional role of what has been termed "affective competence," or the ability to modulate one's own emotions and to respond appropriately to the emotions of others (Halberstadt, Denham, & Dunsmore, 2001). These emotional skills have been understood to play an important role in children's approaches to learning and, when combined with measures of children's attention and persistence, are prospectively predictive of their academic achievement, as well (Li-Grining, Votruba-Drzal, Maldonado-Carreno, & Haas, 2010).

Rather than being placed in opposition, basic research in neuroscience suggests that there is tremendous intellectual benefit to consideration of cognitive and emotional domains of self-regulation as *integrated* at both neurological and behavioral levels (see Dennis & Chen, 2007). At the neurobiological level, limbic and prefrontal cortex systems must be able to work together, with children modulating their neuroendocrine and affective responses to a given environmental stressor in order to be able to encode, remember, and respond appropriately to key features of the cognitively demanding task that they face (Nater et al., 2007; Roelofs, Bakvis, Hermans, van Pelt, & van Honk, 2007). Recent research at the behavioral level suggests strong support for this neurobiologically based model, with individuals' ability to meet the cognitive demands of laboratory-based executive function tasks promoted or significantly impaired by their management of positive and negative emotions (Lench & Levine, 2005; Quas, Bauer, & Boyce, 2004; Schimmack, 2005). Conversely, children's ability to remember rules and strategies, which represents a key aspect of executive function, serves a critical role in helping them to manage negative emotions, such as anxiety and anger (Schmeichel & Demaree, 2010; Wolfe & Bell, 2007). In short, recent models of the integration of executive function and emotion regulation suggest that young children's ability to learn may be significantly supported by both of these key systems (see Ursache, Blair, & Raver, in press for review).

C. SELF-REGULATION ACROSS PRESCHOOL AND EARLY ELEMENTARY SCHOOL CONTEXTS

In past research, we have made the case that children's self-regulation and social relationships with adults represent key foundations for learning in classroom contexts (Raver, 2002; Raver & Knitzer, 2002). Along with other colleagues, our research team carried out a cluster-randomized experimental study to test this claim, implementing a multicomponent classroom-based intervention in some classrooms but not in others to test the role of self-regulation for early learning. Specifically, we took advantage of recent innovations in the design of educational experiments, where schools, rather than children, are randomly assigned to receive new interventions or packages of curricula, teacher training, and coaching support. In this way, the research design takes into account the ways that schools (or in our case, Head Start sites) function as "systems," where the successful implementation of new curricular or instructional approach depends upon the involvement of the multiple teachers, students, and administrators within that school building or site (Bloom, 2005). In the case of our intervention approach, we considered the multiple sources of support that teachers might need in order to make significant behavioral changes to the ways that they managed children's challenging behaviors and negative emotions. Those multiple sources of support included teacher training and weekly coaching as two components of intervention that were implemented at the classroom and site levels (which we discuss in greater detail, below).

Based on the models of leveraging policy experiments to ask theoretically driven scientific questions, we sought to use early intervention services targeting children's self-regulation as an "instrumental variable (IV)" for the hypothesized mediating pathways described (Gennetian, Magnuson, & Morris, 2008). That is, the Chicago School Readiness Project (or CSRP, as it came to be called) did not focus on preschool-aged children's language or math skills; instead, it focused exclusively on preschoolers' self-regulation, allowing us to test whether investing in this key, yet less well-understood domain of development would yield academic as well as socioemotional benefit to young children. To test the efficacy of this model, we collaborated with community-based Head Start programs in seven of Chicago's most economically disadvantaged neighborhoods, randomly assigning nine Head Start sites to receive comprehensive intervention services hypothesized to support children's self-regulation. Another nine Head Start sites served as "control group" classrooms.

Specifically, the CSRP intervention provided preschool teachers with 30h of training and weekly follow-up support on improving the emotional climate of their classrooms by targeting their management of children's dysregulated and disruptive behaviors. In order to provide follow-up support, clinically trained mental health consultants worked as "coaches" with teachers, providing them with feedback on ways to implement the classroom management strategies covered in trainings. In addition, these clinically trained specialists provided direct child-focused consultation, working "1-on-1" with three to five children who exhibited the most challenging behavioral problems, with the view that these children might benefit from access to clinical psychological services.

Our recent analyses suggest that the CSRP significantly improved children's chances of school success in early childhood. Targeting children's self-regulatory skills and

relationships with teachers through classroom-based processes led to statistically significant increases in children's executive function, as well as in their attention and impulse control. Importantly, the intervention also led to gains in children's vocabulary, letter-naming, and in their early math skills, with effect sizes ranging from d=0.32 to d=0.70 (Raver et al., 2011). We also found experimentally induced reductions in children's sad, withdrawn, and disruptive behavior problems (see Raver et al., 2008, 2009). These findings lend strong empirical support to the hypothesis that improvements in children's self-regulation lead to clear benefits in their learning and achievement.

Based on these promising new findings, we have extended our research effort to follow CSRP-enrolled students as they make the transition from preschool through elementary school. The longitudinal follow-up study that is currently underway provides us with a remarkable opportunity to test hypotheses about the ways that more emotionally supportive and well structured versus negative and chaotic school contexts may alternately support or constrain children's executive function and emotion regulation over time. We have successfully followed the large majority of the preschoolers that were initially enrolled in our study as they have made the transition through kindergarten, first, and third grades. We are now currently collecting direct assessments of these children's executive function and emotional regulation as they enter the fifth grade, giving us a remarkably rich set of opportunities to develop and test models of the ways that home, school, and neighborhood environments may shape these key developmental domains.

In our preliminary analyses, it is clear that much has changed in these children's lives over the past 5years: Some children have weathered tough economic times and are thriving well according to teacher and parent reports of emotional and behavioral adjustment, while other children appear to be struggling with early signs of chronic emotional difficulty. We are particularly interested in testing the hypothesis that these key dimensions of early self-regulation, in turn, may shape low-income children's chances for success in later elementary school. For example, some CSRP children have made excellent academic progress while other children have begun to show early signs of academic difficulty and grade retention: Analyses are currently underway to understand the key predictors of these diverging academic trajectories for the children in our sample. It is to the role of measurement of those key predictors that we now turn.

II. Capturing the Role of Self-Regulation for Learning: The Importance of Measurement

From our prior research, it is clear that a strong case can be made for the claim that self-regulation matters for early educational success. But few large-scale studies to date have included direct assessments of children's self-regulation as an important outcome in educational intervention. In our view, the fields of applied developmental, prevention, and education sciences have made major methodological strides in the past several years by developing new and highly innovative research designs and corresponding analytic approaches to strengthen our capacity to make causal inferences about what works in improving children's educational outcomes. For example, these fields have been powerfully moved forward by the increased use of cluster-randomized experimental designs, and more

extensive application of advanced multivariate analytic approaches such as multilevel or hierarchical linear modeling (see Bloom, 2005; Raudenbush & Bryk, 2002). Although these design and analytic approaches have greatly strengthened the internal validity of many studies within the fields of prevention science and education, other means of strengthening studies' internal validity, including increases in the precision of measurement of key mediators and child outcomes, may have lagged behind (Shadish, Cook, & Campbell, 2002). This may be because social scientists have not, until recently, had ready access to standardized and "portable" measures of children's self-regulatory skills that have been extensively piloted or validated outside of university research laboratories. Thus, although young children's self-regulation has increasingly been identified as an important predictor of their ability to navigate the social and academic worlds of early schooling, there have been few tools available to directly assess these key skills (see recent work by Ponitz, McClelland, Matthews, & Morrison, 2009; Willoughby, Blair, Wirth, & Greenberg, 2010 as exceptions).

Does the type of measurement (and its corresponding level of precision) really make that much of a difference in detecting the impact of educational investments? Recent evidence from a meta-analysis of a large number of evaluation studies focusing on the effectiveness of the storied federal Head Start program in supporting children's academic outcomes suggest that it does (Shager et al., 2010). In that meta-analysis of over 40 studies across a period of 50years, results suggest that the type of measurement that investigators used (i.e., teacher ratings versus direct observation of children's skills) was a robust predictor of whether the study yielded statistically significant evidence of the benefit of Head Start for young children. Simply put, these results clearly indicate that the metric matters: Higher quality direct assessments with greater precision for capturing children's self-regulation are likely to maximize the chance of empirically detecting whether educational contexts have a significant impact on both cognitive and emotional domains of children's self-regulation. However, it has not been until recently that measures of children's self-regulation have been included in most educational evaluations (see Bierman, Nix, Greenberg, Blair, & Domitrovich, 2008; Diamond et al., 2007; Morris, Raver, Millensky, Jones, & Lloyd, 2010 as exceptions).

The findings from our study as well as from a range of other recent studies suggest that the scientific value of including more precise levels of measurement is substantial: Not only do we increase the likelihood of detecting the impact of a given intervention or educational enhancement on key developmental domains, but we also gain a valuable lens through which to understand basic processes and mechanisms in development. In short, use of more precise, albeit more expensive, measures of children's self-regulation in the context of educational research presses the research into "double duty," answering policy questions while also answering basic scientific questions about the mind and its plasticity or malleability in the context of experimentally manipulated changes in children's environments.

To address this need for more precise, direct assessment of children's self-regulation, our research team gathered behavioral assessments from the previous laboratory-based work of other investigators that had a strong "track record" to produce valid and reliable, finely

grained behavioral data and that displayed adequate variance across children. One key question was whether an adapted set of measures would hold up when transferred from lab-based settings with higher-income samples, to use with a larger community sample scattered across a number of educational sites or settings. In short, how well does a direct assessment battery adapted from laboratory measures perform in yielding valid, reliable data in larger-scale, educational, and prevention science evaluation research? Put another way, how much can these measures be viewed as *psychometrically robust*? This represents the first key criterion that we must establish in order for findings that are yielded from these measures to be credible or trustworthy in educational evaluation and policy debates.

In addition, investigators have begun to challenge assumptions of "universality" of measures and models, by considering the moderating role of the broader ecological context and testing for measurement equivalence across groups of girls and boys, across younger and older children, and across groups differing by race/ethnicity (Knight & Hill, 1998; Little, 1997). In short, a second critical question is whether a given set of observable indicators (or measures) taps the latent skill it is designed to capture, differently or similarly for children residing in different sociocultural contexts (García Coll et al., 1996; Knight & Hill, 1998; Raver, Gershoff, & Aber, 2007). If observed indicators of self-regulation do not tap latent skills in similar enough fashion across groups of children who differ by race/ethnicity, researchers run the risk of introducing unacceptably large amounts of error into their models (Raver et al., 2007). In short, *measurement equivalence* represents the second key criterion that we must establish in order for us to trust the data that are yielded from our measures.

With these two key criteria of "high-quality" measurement in mind, we provide several examples of the benefit of using direct assessments of children's executive function and emotion regulation across multiple developmental periods. In recognition of the value of a multimethod approach, we then move on to consider the utility of adult reports of children's self-regulatory skills, considering these as "correlates" of the biobehaviorally based components of executive function and emotion regulation that we just discussed. Finally, we consider measurement and analysis of self-regulation from a nonlinear as well as linear perspective, and from a classroom- as well as a child-level framework, providing several alternative approaches to the measurement of self-regulation that are new.

A. EXECUTIVE FUNCTION: THE USE OF PENCIL-TAP AND WALK-A-LINE TASKS IN EDUCATIONAL SETTINGS

For the CSRP study, our research team adapted two laboratory-based executive function tasks for use in field-based contexts. These included a "Balance Beam" task that was adapted from Kochanska and colleagues where children were instructed to walk along a piece of masking tape placed on the floor as slowly as possible, with the assessor stating prior to the trial, "ready, set, go!" The second task, referred to as "Pencil Tap," was adapted from Luria's peg-tapping test (Blair, 2002; Diamond & Taylor, 1996). In this task, the child was instructed to tap a pencil against the table top two times whenever the assessor tapped her pencil once and to tap the pencil one time against the table top whenever the assessor tapped her pencil twice. In both tasks, children's attentiveness to the assessor's instructions, their skills in inhibiting their first impulse (to speed up rather than slow down on the

Balance Beam or to copy the number of pencil taps that the assessor completed), and children's ability to maintain the assessor's rule or instruction in working memory were key to children's success. These tasks were combined with additional cognitive and behavioral tasks to form a brief 30–40 min direct assessment, allowing us to obtain as complete a snapshot as possible of children's regulatory skills, both before and after the implementation of the CSRP preschool intervention. We collected these data with over 500 Black and Hispanic children at the beginning and end of the school year, with children ranging in age from 26 to 73months (mean=49.4, SD=8.0). Importantly, the sample with which these measures were validated was one that experienced significant socioeconomic disadvantage: Families were predominantly making ends meet on low incomes with a mean income of just over $13,000 a year.

To meet the first criterion of measurement quality outlined earlier, it was critical to test whether this pair of tasks of executive function that were "exported" from the laboratory setting to a larger scale educational evaluation could be collected reliably from a cadre of "assessors," and whether scores that were yielded from these two tasks would demonstrate discriminant validity, reliably distinguishing children's executive function from other, related domains of behavior. Staff training to collect laboratory-based tasks in the field was completed through a multistep process of certification, whereby assessors were first trained on the administration of the assessment, and then given multiple opportunities to demonstrate their skills in sticking to standardized protocols and in documenting details of children's performance with a high level of accuracy. This multistep training effort yielded highly reliable data, with assessors demonstrating high levels of "error-free" performance and with high levels of interrater reliability demonstrated on 20% of cases (which were videotaped and double coded, Cronbach's α averaging 0.93).

Our next question was whether the resulting data from the Balance Beam and Pencil Tap tasks could be said to assess a specific, discriminant latent factor of child functioning, namely executive function. Confirmatory factor analysis of our data suggested that (a) children's scores on Pencil Tap and Balance Beam tasks were sufficiently highly correlated with each other and (b) sufficiently less well correlated with children's performance on other tasks (e.g., reflecting children's levels of compliance), that they represented good observed indicators of our constructs of interest. In addition, children's performance on these two executive function tasks demonstrated criterion validity, in that children's scores on these two tasks were quite clearly and strongly predictive of teacher reports of children's difficulties with impulsivity and inattentiveness on a standardized measure of behavioral problems.

A second key criterion was that the direct assessment of children's executive function would demonstrate what is termed "measurement equivalence" across different groups of children who may differ on a set of demographic characteristics, such as age, gender, and ethnic group membership. As mentioned earlier, measurement equivalence is critical to empirical rigor and is established when the relationships between measured indicators and their corresponding latent constructs are sufficiently similar across different groups of individuals.

To test this question, confirmatory factor analysis was used to examine two primary models for invariance across each demographic group. First, a full model that allowed all model parameters to be freely estimated separately across the groups (i.e., separate factor loadings for boys and girls) was estimated (χ^2=47.38, df=17, p<0.05; CFI=94, RMSEA=0.06). Then, a reduced model that constrained all factor loadings and structural loadings to be equivalent across groups was estimated (i.e., same factor loadings for boys and girls). The reduced models were nested under the full models, and χ^2 comparison tests were used to test for a significantly worse fit in the constrained model. A significant difference between the two models indicated that the hypothesis of full measurement equivalence should be rejected. If full measurement equivalence was rejected, tests were then conducted to determine whether any of the factor loadings could be constrained to equivalence across groups.

In each of the three multiple group analyses for gender, age, and race/ ethnicity, the fully constrained models fit significantly worse than the unconstrained models (gender χ^2=15.61, df=8, p<0.05; age χ^2=35.93, df=8, p<0.001; race/ethnicity χ^2=19.36, df=8, p<0.05). This indicated that there were significant differences between groups on at least one of the factor loadings and that full measurement equivalence could not be established. Analyses were then conducted to determine which of the factor loadings could be constrained to equality between groups and which factor loadings needed to be estimated separately (see Table I). These analyses were conducted with the covariances among latent factors constrained to equality across groups.

For the gender analyses, all the factors, with the exception of Toy Sort, could be constrained to be equal across boys and girls. For the age analyses, only four factors (Tower Task, Toy Peek, Toy Sort, and Tower Clean-up) could be constrained to equality among older and younger children. Finally, for the race/ethnicity analyses, all the factors, with the exception of Tongue Task/Snack Delay and Tower Clean-up, could be constrained to equality among Hispanic and African-American children.

In summary, our analyses of children's Balance Beam and Pencil Tap scores (as well as their scores in other areas of self-regulation) yielded evidence of measurement invariance, suggesting that the direct assessments of executive function performed equally well for older and younger children, for boys and girls, and for Black and Latino children (see Table I). In short, these analyses helped us to have confidence in ways that these direct assessments of executive function adapted from laboratory measures yield valid and reliable data that we could then use in larger-scale, educational and prevention science evaluation research. Recent, independent efforts to replicate these findings have yielded clear empirical evidence confirming the validity of the Preschool Self-Regulation Assessment (PSRA) across groups of children that vary by their exposure to higher versus lower levels of economic disadvantage (Denham, Warren-Khot, Bassett, Wyatt, & Perna, 2011). Our effort to validate measures of cognitive and emotional self-regulation is part of an emerging trend among a number of researchers to develop and validate measurement tools that can be used to assess multiple components of children's self-regulation in field-based contexts (Blair & Razza, 2007; Li-Grining, Votruba-Drzal, Bachman, & Chase-Lansdale, 2006; McClelland & Morrison, 2003; Raver, Blackburn, Bancroft, & Torp, 1999).

B. CAPTURING KEY COMPONENTS OF EMOTION REGULATION IN EDUCATIONAL SETTINGS

How might measures be similarly adapted from lab contexts to field contexts in order to directly assess children's emotion regulation? During the PSRA, we also trained assessors to briefly and reliably code children's expressions of frustration, anxiety, and positive engagement during the test administration, which represented an important first step in developing and validating robust measures of emotion regulation that could be used in educational settings. With these measures, we found few associations between the intervention and children's emotion regulation. We speculate that this may have been because the intervention may have helped children in experiencing lower levels of stress, lower levels of arousal, and possibly lower levels of internally felt distress in the context of improved classroom climate (see Raver et al., 2008). Those improvements in the emotional climate of the classroom may not necessarily have translated to improvements in children's *displays* of negative versus positive emotion, however. In contrast, subsequent, nonexperimental analyses using these data revealed more robust evidence of the role of preschoolers' lower levels of emotional self-regulation as an important predictor of teachers' higher ratings of their internalizing and externalizing behavior problems (McCoy & Raver, 2011). This set of analyses underscored ways that inclusion of developmentally driven measurement may sometimes yield unexpected scientific benefit: In this case, findings contributed to an emerging literature on the normative development of low-income, ethnic minority children, even though they did not directly yield evidence to support the efficacy of the preschool intervention approach for children's emotion regulation skills.

Our recent work in following CSRP-enrolled children's progress as they transition from early to later elementary school has given us additional insight into innovative ways of adapting lab-based tasks designed to tap children's emotion regulation in educational settings. Specifically, we wished to capture the roles of classroom and school climate in shaping children's emotional regulation as well as in shaping their executive function. Similar to work with executive function in preschool, we identified multiple research paradigms that had been used successfully in university- based laboratory settings with 7- to 11-year olds with lower-risk samples. We then spent significant time piloting these measures in community settings and agencies and have recently worked with a large cadre of assessors to administer these measures on laptop computers with our larger sample of CSRP-enrolled students. Prior to beginning data collection, assessors completed a two-day training where they learned how to utilize the computers, administer the measures, and transfer the data to a shared server, post-collection. To achieve certification, each individual was rated by a trainer on their ability to accurately complete all aspects of data collection. All 10 of the CSRP assessors were successfully certified. Following certification, assessors visited each student's school to administer the measures. Students were excused from their classes for approximately 40 min to complete the battery of measures. Assessors administered the tasks on laptop computers that they brought with them. All efforts were made to administer the measures in quiet, private areas to minimize disruptions. Reports from the assessors and initial analysis of the data reveal that data collection efforts were successful, and that it is possible to collect direct assessments from students in school settings with minimal effort.

Some of the measures that we used were selected to measure children's accuracy in identifying emotions from images of adult faces (called the Florida Affective Battery, FAB; Bowers, Blonder, & Heilman, 1991) and of children's ability to deploy their visual attention when presented with emotionally salient versus neutral stimuli for very brief periods of time (called the Dot Probe; Kimonis, Frick, Munoz, & Aucoin, 2007; Kujawa et al., 2011), as behavioral indicators of children's emotion regulation. We use two versions of the FAB. In version A, which measures facial affect selection, children are presented with five pictures of one person displaying five different emotions (happy, sad, frightened, angry, and neutral) and asked to point to a particular emotion (i.e., "point to the happy face"). In version B, which measures facial affect discrimination, children are presented with two pictures of two different people and are asked to indicate whether the faces are "similar" or "different." In our version of the Dot Probe task, two pictures, each displaying a different emotional valence, are simultaneously presented for 250ms on the left- and right-hand sides of a computer screen. Picture pairs include neutral/neutral, neutral/positive, and neutral/negative images drawn from the International Affective Picture database. Next, a dot either appears in the same location as the emotion image (congruent trials), in the location opposite to the emotion image (incongruent trials), or on either side of the neutral/ neutral display (neutral control). Respondents use the keyboard to indicate the location of the dot. Data drawn from across these trials help us to determine the ease or difficulty with which children marshal their attention to something as simple as the location of a "dot" on the screen, after exposure to emotion-arousing stimuli.

Methodologically, these tools have been very helpful in past research focused on the role of the environment, framed more broadly, in shaping children's emotion regulation skills. For example, using similar laboratory tasks to capture children's attention to emotionally salient cues, research at the intersection of developmental and clinical psychology has been able to clearly establish that children chronically exposed to adult anger and aggression may develop patterns of self-regulation that are biased in favor of heightened vigilance and responsiveness to emotionally negative stimuli (Kimonis et al., 2007; Kujawa et al., 2011; Pollak, Messner, Kistler, & Cohn, 2009; Pollak, Vardi, Putzer Bechner, & Curtin, 2005). Children repeatedly exposed to high levels of conflict and violence in the household have less coherent cognitive schemas of the causes, consequences, and solutions of conflict. On the other hand, these children's hypervigilance to affective cues may also mean that violence-exposed children are likely to be more sensitive readers of emotional volatility and anger in others, at the cost of being able to focus less attention to the cognitive demands and opportunities of learning environments (Macfie, Cicchetti, & Toth, 2001).

Before proceeding with tests of our key hypotheses (outlined below), we will first follow multiple steps to determine whether these adapted emotion regulation measures meet the two criteria that were described earlier. That is, were these measures effective in yielding data that are reliable and internally valid, in both the methodological and statistical senses of those terms? Feedback from our team of assessors and review of the data indicate that the tasks were easily administered on small laptop computers, children were able to complete the tasks with no major problems, and there were minimal missing data across respondents, indicating that use of the FAB and Dot Probe is feasible in field-based settings. Preliminary statistical analyses revealed that the emotionally negative stimuli in the Dot Probe task, for

example, were effective in eliciting a behaviorally distinct (and significantly slower) response from children than did the emotionally neutral stimuli, even though exposure to those stimuli were flashed on the laptop computer screen for less than a second (Ursache, Roy, & Raver, 2012). We will complete additional analyses to determine whether data resulting from the administration of the FAB also meet basic standards of internal validity. And to meet our second criterion of "high quality," did these measures yield data that suggest that the assessments are tapping emotion regulatory skills in similar enough ways across African-American and Hispanic children that the assessments can be said to demonstrate "measurement equivalence?" Our research team is moving forward to address these important empirical questions.

Once we are confident that the resulting data and evidence are empirically trustworthy, we plan to test the hypothesis that school settings (in addition to home and neighborhood settings) may shape or canalize children's emotional regulatory skills in ways that may bode well or negatively for their ability to navigate both the cognitive and social demands of school. Based on developmental theory, we are interested in exploring the extent to which educational settings may negatively affect children's neuroendocrine and behavioral reactivity in ways that may limit and interfere with both their ability to process emotional information and their ability to engage in optimal levels of executive function.

Evidence for the hypothesis that chaotic and stressful out-of-home settings may negatively affect children's neuroendocrine response has been found for toddlers and younger children in lower versus higher quality child care settings (Gunnar, Sebanc, Tout, Donzella, & van Dulmen, 2003; Watamura, Donzella, Kertes, & Gunnar, 2004). But to our knowledge, few studies have considered the self-regulatory sequelae of chronic exposure to emotionally and physically threatening levels of chaos and conflict in elementary schools, for older children. Adaptation to an environment characterized by high levels of threat increases hypervigilance, reduces flexibility in children's neuroendocrine reactivity, and increases the likelihood of attributional bias when responding to affectively arousing social cues (Blair & Raver, in press). Importantly, with the inclusion of direct assessments of executive function and emotion regulation in the follow-up phase of data collection in our longitudinal study, we are able to examine the ways that emotion regulation and executive function skills may be behaviorally (if not neurologically) interrelated. We will also have the remarkable opportunity to test basic scientific hypotheses about ways that cognitive and emotional regulatory processes may work, independently and in concert, to support learning.

C. RELYING ON ADULT REPORTS OF THE CORRELATES OF CHILDREN'S SELF-REGULATION

An added benefit of direct assessment is that the adult who administers the battery of measures (identified on our teams as an assessor) represents an excellent source of additional information regarding children's capacity to handle cognitively and emotionally demanding tasks. In our research, we have viewed assessors' perspectives on children's attentiveness, impulsivity, level of positive engagement, and their expressions of frustration and anxiety during this moderately stressful "test-taking" situation as exceptionally valuable. It is worth noting that we view these adults' reports (as well as reports provided by teachers

and parents) as tapping the *correlates* of children's self-regulation. That is, we feel it is important to reserve terms like "executive function" and "emotion regulation" for children's specific, detailed behavioral responses to carefully identified and standardized tasks. In so doing, we maintain the precision of the terms we use, and our resulting findings can be more easily aligned with results from lab-based studies within the field of developmental science. In contrast, adults may observe a range of actions, emotional expressions, vocalizations, and other behaviors on the part of the child that are correlated with children's executive function and emotion regulation. We now turn to the adult report of those correlates.

How can assessors who administer batteries of behavioral tasks provide us with insight to these correlates? Using a brief checklist called the PSRA-Assessor Report that is completed after the adult has finished the assessment with the child, we capitalize on these adults as key informants who can provide a quick albeit somewhat broad-brush "snapshot" of children's regulatory skill. As with children's scores on the PSRA tasks, assessors' ratings of children's attention and impulse control on the PSRA-Assessor Report have demonstrated high levels of internal consistency and predictive validity in our analyses (Cronbach's α=0.92; zero-order correlation with a measure of social competence of r=0.30), suggesting this may be a productive, supplementary approach to capture important information on ways that children modulate their behavioral, emotional, and cognitive responses to stressful conditions and stimuli (Smith-Donald, Raver, Hayes, & Richardson, 2007).

In many cases, researchers often wish to include teachers' and parents' perspectives on children's self-regulatory skills in addition to the information that can be derived from direct assessments. At present, few standardized measures have been developed and validated for use by researchers, teachers, and practitioners who wish to quickly and reliably assess children's self-regulation within the classroom context. In one recent paper, we explored the psychometric properties of a teacher-reported composite of two clinical measures of self-regulation—the Barratt Impulsiveness Scale (version 11; BIS-11) and the Behavior Rating Inventory of Executive Function (BRIEF)—in our CSRP sample when the low-income African-American and Hispanic children in our study were in early elementary school (McCoy, Raver, Lowenstein, & Tirado-Strayer, 2011). We held this measure to the same set of criteria described earlier, in order to assess the quality or trustworthiness of these measures.

Results of factor analyses of data from our study revealed a two-factor solution for the composite measure formed from the BIS-11 and the BRIEF that corresponds to latent domains of cognitive and behavioral self-regulation. Our results indicated that scores on these two measures could be combined to yield a composite scale that demonstrated high internal consistency, reliability, and concurrent validity when compared to a previously validated measure of teacher-rated inattention and hyperactivity/ impulsivity. Moreover, teacher-reported data gleaned from these two measures demonstrated clear measurement invariance across multiple groups of children who differed in their racial category membership, their level of poverty risk and gender (McCoy et al., 2011). In short, these measures yielded data that can now be included in further analyses of the role of school climate and other environmental characteristics in predicting elementary-aged children's self-regulation.

In many ways, cognitive and affective domains of self-regulation have their most serious consequences when children experience high levels of dysregulation in both domains, and as a consequence, lose behavioral self-control. Several decades of research suggest that low levels of executive function, including low levels of inhibitory control, when combined with high levels of expressed anger, frustration, or anxiety can lead a small number of children to manifest high levels of impulsivity and acting out behavior that are challenging and stressful for many teachers to manage (see Campbell, Spieker, Vandergrift, Belsky, & Burchinal, 2010; Raver, Blair, & Li-Grining, in press for review). Children who are persistently emotionally dysregulated and behaviorally disruptive have been found to bear significant academic cost, receiving less instruction from teachers and experiencing fewer opportunities for learning from peers (see Arnold et al., 2006; McClelland & Morrison, 2003). For these reasons, we consistently also include standardized measures of children's externalizing and internalizing behavior problems in our data collection efforts and analyses.

D. NEW WAYS OF MODELING SELF-REGULATION IN EDUCATIONAL CONTEXTS I: THE INTEGRATION OF AFFECTIVE AND COGNITIVE DOMAINS

As we have outlined in earlier sections of this chapter, we are particularly interested in the inclusion of both executive function and emotion regulation tasks in larger-scale educational studies because of the ways that the resulting data will allow us to test pressing questions in developmental science: That is, questions of whether and how children's cognitive and emotional regulatory competence may be interrelated. Previous research, largely conducted in the lab, suggests that children's executive function is inextricably tied to their modulation of affective arousal and attention through bidirectional "bottom-up" and "top-down" processes (Blair & Ursache, 2011; Lewis & Todd, 2007). Collection of executive function and emotion regulation data in field-based contexts, such as classrooms, will allow us to learn whether previous findings of linkages between these two domains of regulatory competence are empirically robust when generalized to community samples of children facing concentrated economic disadvantage.

We hope to also test models of the differential versus common predictors of these two different domains of self-regulation. Put simply, do some educational and social contextual factors (such as high levels of classroom disorganization, noisiness, and less well-structured use of class time) predict children's executive function, whereas other educational factors (such as the level of support versus conflict experienced in social relationships with teachers and peers) predict children's capacity to regulate their emotions, in differentially clear ways? And do executive function and emotion regulation skills, in turn, differentially or similarly predict children's academic achievement, their ability to focus and sustain their attention during class time, and their motivation to tackle academic topics and problems even when those tasks are difficult? We speculate that there are some ways that executive function and emotion regulation will serve joint, mediating roles in the link between lower quality educational settings and children's ability to do well academically and socially. But we are also particularly interested in ways that executive function and emotion regulation may maintain domain-specific linkages as well as cross-domain associations with children's academic and social outcomes (such as getting along well with teachers and peers). In short, by collecting more precise direct assessments of executive function and emotion regulation,

we hope to have the remarkable opportunity to test basic scientific hypotheses about ways that cognitive and emotional regulatory processes may work independently and in concert to support learning.

E. NEW WAYS OF SEEING SELF-REGULATION IN EDUCATIONAL CONTEXTS II: THE ROLE OF CLASSROOM COMPOSITION

From neurobiological through behavioral levels of analysis, foundational models of the development of children's self-regulation also highlight the ways in which regulation of cognition and emotion is not only bidirectional but also interpersonal. For example, self-regulation in early childhood is preceded by regulation by others (notably parents) in infancy; sensitive and responsive adults routinely provide early behavioral support to scaffold infants' modulation of emotional arousal and processing of the flood of cognitively engaging stimuli in the environment (Spinrad & Stifter, 2002). Regulation by others, in turn, leads to children's increasingly autonomous ability to handle cognitive and emotional demands later on in life (see Bernier, Carlson, & Whipple, 2010; Calkins & Fox, 2002 for review).

Most critically, numerous investigators have pointed out that children not only regulate their own states but also actively regulate the attention, cognitions, and emotions of their teachers and peers. The implications of this expanded model for the development of self-regulation in classroom contexts are enormous: How much, for example, does the child's opportunity to learn to focus attention, working memory, and inhibitory control in solving a problem or learning new material depend on the behavioral support versus disruptiveness of the other children in his or her proximity? How are teachers' abilities to provide supportive and positive classroom management and regulatory support to a target child affected by the behavioral profiles of other children in the classroom?

Whether with direct assessments or teacher reports, educationally linked data on self-regulation opens up new possibilities for models of the influence of peers in shaping children's trajectories of self-regulation. One key methodological issue is that there are a number of different ways of empirically operationalizing what we mean by "classroom composition" as we tackle the challenge of aggregating children's self-regulation profiles to better understand the dynamics of the classroom as a whole. For example, is the classroom environment more susceptible to influences from one or a few children with high levels of behavior problems, or by the average level of behavior problems across all children in the classroom? We have explored these questions in preliminary analyses by operationalizing the classroom composition as the mean level of emotional and behavioral dysregulation across all children (e.g., as indicated by the classroom average in teacher ratings versus parent ratings of their externalizing behavior problems). Alternately, we have tried to capture the extent to which chaotic conditions may be dispersed rather than concentrated within a given classroom by considering the percentage of children in the classroom above a clinical threshold for externalizing behavior problems. These two different approaches yield contrasting profiles of adjustment for individual children in the classroom and differing portraits of the classroom as a whole (see Table II for illustration).

III. Summary and Conclusions

Our recent studies suggest that the PSRA and other direct measures, such as the Dot Probe, represent a promising new approach to assessment of children's self-regulation in preschool and elementary school contexts. Measures such as these offer the strong potential of helping us to evaluate the impact of interventions on children's socioemotional development much in the same way that direct assessments of children's cognition, vocabulary, and reading skills made it possible to clearly detect the benefits and costs of social programs, interventions, and policies (Duncan et al., 2007).

Our work has also involved finding innovative ways to manage the tradeoffs between the high internal validity of lab-based work with broader generalizability or external validity of collecting data with samples of children from low and moderate income households attending schools and community-based organizations in urban settings. As we continue to explore and validate these measurement approaches, we hope to strengthen multiple fields' understanding of the cognitive and emotional regulatory mechanisms that underlie children's academic and social successes and difficulties in schools. Given the increasing economic and racial/ethnic diversity of schools in the U.S., this work addresses the vitally important concern that our measures are psychometrically robust across groups of children from widely varying socioeconomic and sociocultural contexts. In so doing, this research contributes to the growing capacity of researchers to situate their work at the intersection of developmental science and social policy analysis.

REFERENCES

Arnold DH, Brown SA, Meagher S, Baker CN, Dobbs J, Doctoroff GL. Preschool-based programs for externalizing problems. Education and Treatment of Children. 2006; 29:311–339.

Bernier A, Carlson SM, Whipple N. From external regulation to self-regulation: Early parenting precursors of young children's executive functioning. Child Development. 2010; 81:326–339. [PubMed: 20331670]

Bierman KL, Nix RL, Greenberg MT, Blair C, Domitrovich CE. Executive functions and school readiness intervention: Impact, moderation, and mediation in the head start REDI program. Development and Psychopathology. 2008; 20(3):821–843. [PubMed: 18606033]

Blair C. School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. American Psychologist. 2002; 57:111–127. [PubMed: 11899554]

Blair C. Stress and the development of self-regulation in context. Child Development Perspectives. 2010; 4:181–188. [PubMed: 21779305]

Blair C, Raver CC. Individual Development and Evolution: Experiential Canalization of Self-Regulation Development. Developmental Psychology. (in press).

Blair C, Razza RP. Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. Child Development. 2007; 78(2):647–663. [PubMed: 17381795]

Blair, C.; Ursache, A. A bidirectional model of executive functions and self-regulation. In: Vohs, KD.; Baumeister, RF., editors. Handbook of Self-Regulation: Research, Theory, and Applications. 2nd edition. New York: Guilford; 2011.

Blair C, Zelazo PD, Greenberg MT. The measurement of executive function in early childhood. Developmental Neuropsychology. 2005; 28(2):561–571. [PubMed: 16144427]

Bloom, HS. Learning more from social experiments: Evolving analytic approaches. New York: Russell Sage; 2005.

Bowers, D.; Blonder, LX.; Heilman, KM. Florida affect battery. Gainesville: University of Florida; 1991.

Bull R, Espy KA, Weibe SA. Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. Developmental Neuropsychology. 2008; 33(3):205–228. [PubMed: 18473197]

Calkins SD, Fox NA. Self-regulatory processes in early personality development: A multilevel approach to the study of childhood social withdrawal and aggression. Development and Psychopathology. 2002; 14:477–498. [PubMed: 12349870]

Campbell SB, Spieker S, Vandergrift N, Belsky J, Burchinal M. Predictors and sequelae of trajectories of physical aggression in school-age boys and girls. Development and Psychopathology. 2010; 22(1):133–150. [PubMed: 20102652]

Carlson SM. Developmentally sensitive measures of executive function in preschool children. Developmental Neuropsychology. 2005; 28(2):595–616. [PubMed: 16144429]

Cole PM, Martin SE, Dennis TA. Emotion regulation as a scientific construct: Methodological challenges and directions for child development research. Child Development. 2004; 75(2):317–333. [PubMed: 15056186]

Denham SA, Warren-Khot HK, Bassett HH, Wyatt T, Perna A. Factor structure of self-regulation in preschoolers: Testing models of a field-based assessment for predicting early school readiness. Journal of Experimental Child Psychology. 2011; 111(3):386–404. [PubMed: 22104321]

Dennis T, Chen C. Neurophysiological mechanisms in the emotional modulation of attention: The interplay between threat sensitivity and attentional control. Biological Psychology. 2007; 76:1–10. [PubMed: 17582673]

Diamond A, Barnett WS, Thomas J, Munro S. Preschool program improves cognitive control. Science. 2007; 318(5855):1387–1388. [PubMed: 18048670]

Diamond A, Taylor C. Development of an aspect of executive control: Development of the abilities to remember what I said and to "Do as I say, not as I do. Developmental Psychobiology. 1996; 29:315–334. [PubMed: 8732806]

Duncan GJ, Dowsett CJ, Claessens A, Magnuson K, Huston AC, Klebanov P, et al. School readiness and later achievement. Developmental Psychology. 2007; 43(6):1428–1446. [PubMed: 18020822]

Eisenberg N, Fabes RA, Guthrie IK, Murphy BC, Maszk P, Holmgren R, et al. The relations of regulation and emotionality to problem behavior in elementary school children. Development and Psychopathology. 1996; 8:141–162.

Espy KA, Kaufmann PM, McDiarmid MD, Glisky ML. Executive functioning in preschool children: Performance on A-not-B and other delayed response format tasks. Brain and Cognition. 1999; 41(2):178–199. [PubMed: 10590818]

García Coll C, Lamberty G, Jenkins R, McAdoo HP, Crnic K, Wasik BH, et al. An integrative model for the study of developmental competencies in minority children. Child Development. 1996; 67:1891–1914. [PubMed: 9022222]

Gennetian LA, Magnuson K, Morris PA. From statistical associations to causation: What developmentalists can learn from instrumental variables techniques coupled with experimental data. Developmental Psychology. 2008; 44(2):381–394. [PubMed: 18331130]

Greenberg, MT.; Kusche, CA.; Speltz, M. Emotional regulation, self-control and Psychopathology: The role of relationships in early childhood. In: Cicchetti, D.; Toth, S., editors. Rochester Symposium on Developmental Psychopathology. Vol. 2. New York: Cambridge University Press; 1991. p. 21-56.

Gunnar MR, Sebanc AM, Tout K, Donzella B, van Dulmen MM. Peer rejection, temperament, and cortisol activity in preschoolers. Developmental Psychobiology. 2003; 43(4):346–368. [PubMed: 15027418]

Halberstadt A, Denham SA, Dunsmore JC. Affective social competence. Social Development. 2001; 10(1):79–119.

Jones LB, Rothbart MK, Posner MI. Development of executive attention in preschool children. Developmental Science. 2003; 6(5):498–504.

Kimonis ER, Frick PJ, Munoz LC, Aucoin KJ. Can a laboratory measure of emotional processing enhance the statistical prediction of aggression and delinquency in detained adolescents with

callous-unemotional traits? Journal of Abnormal Child Psychology. 2007; 35(5):773–785. [PubMed: 17510788]

Knight, GP.; Hill, NE. Measurement equivalence in research involving minority adolescents. In: McLoyd, VC.; Steinberg, L., editors. Studying minority adolescents. Mahwah, NJ: Lawrence Erlbaum Associates; 1998. p. 183-210.

Kochanska G, Murray K, Coy KC. Inhibitory control as a contributor to conscience in childhood: From toddler to early school age. Child Development. 1997; 68(2):263–277. [PubMed: 9180001]

Kopp CB. Regulation of distress and negative emotions: A developmental view. Developmental Psychology. 1989; 25(3):343–354.

Kujawa AJ, Torpey D, Kim J, Hajcak G, Rose S, Gotlib IH, et al. Attentional biases for emotional faces in young children of mothers with chronic or recurrent depression. Journal of Abnormal Child Psychology. 2011; 39(1):125–135. [PubMed: 20644991]

Lench H, Levine L. Effects of fear on risk and control judgements and memory: Implications for health promotion messages. Cognition and Emotion. 2005; 19(7):1049–1069.

Lewis M, Todd R. The self-regulating brain: Cortical-subcortical feedback and the development of intelligent action. Cognitive Development. 2007; 22:406–430.

Li-Grining CP, Votruba-Drzal E, Bachman HJ, Chase-Lansdale PL. Are certain preschoolers at risk in the era of welfare reform? The moderating role of children's temperament. Children and Youth Services Review. 2006; 28(9):1102–1123.

Li-Grining CP, Votruba-Drzal E, Maldonado-Carreno C, Haas K. Children's early approaches to learning and academic trajectories through fifth grade. Developmental Psychology. 2010; 46(5): 1062–1077. [PubMed: 20822223]

Little TD. Mean and covariance structures (MACS) analyses of cross-cultural data: practical and theoretical issues. Multivariate Behavioral Research. 1997; 32(1):53–76.

Macfie J, Cicchetti D, Toth SL. Dissociation in maltreated versus non-maltreated preschool-aged children. Child Abuse & Neglect. 2001; 25:1253–1267. [PubMed: 11700697]

McClelland MM, Morrison FJ. The emergence of learning-related social skills in preschool children. Early Childhood Research Quarterly. 2003; 18(2):206–224.

McCoy DC, Raver CC. Caregiver emotional expressiveness, child emotion regulation, and child behavior problems among Head Start families. Social Development. 201110.1111/j. 1467-9507.2011.00608.x

McCoy DC, Raver CC, Lowenstein AE, Tirado-Strayer N. Assessing self-regulation in the classroom: Validating the BIS-11 and the BRIEF in a low-income, ethnic minority sample. Early Education and Development. 2011; 22(6):883–906.

Mischel W, Ebbesen EB, Zeiss AM. Determinants of selective memory about the self. Journal of Consulting and Clinical Psychology. 1976; 44(1):92–103. [PubMed: 1245637]

Morris P, Raver CC, Millensky M, Jones SM, Lloyd CM. Making preschool more productive: How classroom management training can help teachers. MDRC Working Paper. 2010

Nater UM, Moor C, Okere U, Stallkamp R, Martin M, Ehlert U, et al. Performance on a declarative memory task is better in high than low cortisol responders to psychosocial stress. Psychoneuroendocrinology. 2007; 32(6):758–763. [PubMed: 17606328]

Pollak SD, Messner M, Kistler DJ, Cohn JF. Development of perceptual expertise in emotion recognition. Cognition. 2009; 110(2):242–247. [PubMed: 19059585]

Pollak SD, Vardi S, Putzer Bechner AM, Curtin JJ. Physically abused children's regulation of attention in response to hostility. Child Development. 2005; 76(5):968–977. [PubMed: 16149995]

Ponitz CC, McClelland MM, Matthews JS, Morrison FJ. A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. Developmental Psychology. 2009; 45(3): 605–619. [PubMed: 19413419]

Quas JA, Bauer A, Boyce WT. Physiological reactivity, social support, and memory in early childhood. Child Development. 2004; 75(3):797–814. [PubMed: 15144487]

Raudenbush, SW.; Bryk, AS. Hierarchical linear models: Applications and data analysis methods. Thousand Oaks, CA: Sage; 2002.

Raver CC. Emotions matter: Making the case for the role of young children's emotional development for early school readiness. SRCD Social Policy Report. 2002; 16:3–18.

Raver CC, Blackburn EK, Bancroft M, Torp N. Relations between effective emotional self-regulation, attentional control, and low-income preschoolers' social competence with peers. Early Education and Development. 1999; 10(3):333–350.

Raver, CC.; Blair, C.; Li-Grining, CP. Chapter to be included as part of the National Center for Research on Early Childhood Education book series on early childhood policy and practice. Baltimore, MD: Brookes; Extending models of emotional self-regulation to classroom settings: Implications for professional development. (in press).

Raver CC, Gershoff ET, Aber JL. Testing equivalence of mediating models of income, parenting, and school readiness for White, Black, and Hispanic children in a national sample. Child Development. 2007; 78:96–115. [PubMed: 17328695]

Raver CC, Jones SM, Li-Grining CP, Metzger M, Champion KM, Sardin L. Improving preschool classroom processes: Preliminary findings from a randomized trial implemented in Head Start settings. Early Childhood Research Quarterly. 2008; 23(1):10–26.

Raver CC, Jones SM, Li-Grining CP, Zhai F, Bub K, Pressler E. CSRP's impact on low-income preschoolers' pre-academic skills: Self-regulation as a mediating mechanism. Child Development. 2011; 82(1):362–378. [PubMed: 21291447]

Raver CC, Jones SM, Li-Grining C, Zhai F, Metzger MW, Solomon B. Targeting children's behavior problems in preschool classrooms: A cluster-randomized controlled trial. Journal of Consulting and Clinical Psychology. 2009; 77(2):302–316. [PubMed: 19309189]

Raver, CC.; Knitzer, J. Ready to enter: What research tells policymakers about strategies to promote social and emotional school readiness among three-and four-year-old children. New York: National Center for Children in Poverty, Mailman School of Public Health, Columbia University; 2002.

Rodriguez ML, Ayduk O, Aber JL, Mischel W, Sethi A, Shoda Y. A contextual approach to the development of self-regulatory competencies: The role of maternal unresponsivity and toddlers' negative affect in stressful situations. Social Development. 2005; 14(1):136–157.

Roelofs K, Bakvis P, Hermans EJ, van Pelt J, van Honk J. The effects of social stress and cortisol responses on the preconscious selective attention to social threat. Biological Psychology. 2007; 75(1):1–7. [PubMed: 17030399]

Schimmack U. Attentional interference effects of emotional pictures: Threat, negativity, or arousal? Emotion. 2005; 5:55–66. [PubMed: 15755219]

Schmeichel BJ, Demaree HA. Working memory capacity and spontaneous emotion regulation: High capacity predicts self-enhancement in response to negative feedback. Emotion. 2010; 10(5):739–744. [PubMed: 21038959]

Shadish, WR.; Cook, TD.; Campbell, DT. Experimental and quasi- experimental designs for generalized causal inference. Boston, MA: Houghton Mifflin; 2002.

Shager, HM.; Schindler, HS.; Hart, CM.; Duncan, GJ.; Magnuson, KA.; Yoshikawa, H. Using meta-analysis to explain variation in Head Start research results: The role of research design. In: Zaslow, M., editor. Early childhood teacher professional development and meta-analysis. Symposium conducted at the meeting of the Society for Research on Educational Effectiveness; Washington, DC. 2010.

Smith-Donald R, Raver CC, Hayes T, Richardson B. Preliminary construct and concurrent validity of the Preschool Self-Regulation Assessment (PSRA) for field research. Early Childhood Research Quarterly. 2007; 22(2):173–187.

Spinrad, TL.; Stifter, CA. Maternal sensitivity and infant emotional reactivity: Concurrent and longitudinal relations. In: Fabes, RA., editor. Emotions and the Family. Binghamton, NY: Haworth; 2002. p. 243-264.

Ursache A, Blair C, Raver CC. The promotion of self-regulation as a means of enhancing school readiness and early achievement in children at risk for school failure. Child Development Perspectives. (in press).

Ursache, A.; Roy, AL.; Raver, CC. Measuring emotion regulation in field-based settings: Understanding the dot probe task. Poster presented at the Society for Research in Child Development Themed Meeting: Developmental Methodology; Tampa, FL. 2012.

Watamura SE, Donzella B, Kertes DA, Gunnar MR. Developmental changes in baseline cortisol activity in early childhood: Relations with napping and effortful control. Developmental Psychobiology. 2004; 45(3):125–133. [PubMed: 15505801]

Webster-Stratton C, Taylor T. Nipping early risk factors in the bud: Preventing substance abuse, delinquency, and violence in adolescence through interventions targeted at young children (0–8 years). Prevention Science. 2001; 2(3):165–192. [PubMed: 11678292]

Wiebe SA, Espy KA, Charak D. Using confirmatory factor analysis to understand executive control in preschool children: I Latent structure. Developmental Psychology. 2008; 44:575–587. [PubMed: 18331145]

Willoughby MT, Blair CB, Wirth RJ, Greenberg M. The measurement of executive function at age 3 years: Psychometric properties and criterion validity of a new battery of tasks. Psychological Assessment. 2010; 22(2):306–317. [PubMed: 20528058]

Wolfe CD, Bell MA. The integration of cognition and emotion during infancy and early childhood: Regulatory processes associated with the development of working memory. Brain and Cognition. 2007; 65(1):3–13. [PubMed: 17630061]

Table I

Standardized Factor Loadings for PSRA Total and Equivalence Models

		Equivalence models						
		Gender		Age		Race/ethnicity		
Factor	Total	Girls	Boys	Young	Old	Black	Latino
Balance Beam	0.45	0.47	0.43	**0.29**	**0.32**	0.43	0.50
Pencil Tap	0.74	0.73	0.74	**0.58**	**0.90**	**0.71**	0.67
Tower Task	0.49	0.48	0.52	0.45	0.45	0.49	0.53
Tongue/Snack	0.66	0.70	0.61	**0.71**	**0.58**	0.71	**0.54**
Toy Peek	0.63	0.62	0.60	0.51	0.56	0.63	0.65
Toy Touch	0.46	0.48	0.44	**0.60**	**0.38**	0.43	0.47
Toy Sort	0.45	**0.59**	**0.30**	0.44	0.34	0.41	0.39
Tower Clean-up	0.52	0.51	0.42	0.39	0.54	**0.37**	**0.94**

Note: All loadings significant at *p*<0.01. Bolded items indicate factor loadings that could not be constrained to equality across groups. Other factor loadings differ across groups due to differences in standard errors. Factor loadings for the equivalence models are based on the reduced model in which a set of factor loadings and all covariances were constrained to equality across groups.

Table II

Example of Operationalizing Peer Composition in a Sample Classroom

Child ID	Mean externalizing behavior problems score			Above clinical threshold	
	Teacher report	Parent report	Highest report	Teacher report	Highest report
1	6	**18**	18	No	Yes
2	12	**16**	16	No	Yes
3	**11**	10	11	No	No
4	10	**12**	12	No	No
5	2	**7**	7	No	No
6	**3**	0	3	No	No
7	9	**12**	12	No	No
8	**11**	9	11	No	No
9	**19**	15	19	Yes	Yes
10	**24**	17	24	Yes	Yes
11	5	**12**	12	No	No
12	**10**	8	10	No	No
13	**16**	10	16	Yes	Yes
14	20	**21**	21	Yes	Yes
15	**7**	5	7	No	No
16	8	**10**	10	No	No
17	**11**	10	11	No	No
18	**6**	3	6	No	No
19	14	**18**	18	Yes	Yes
20	**16**	10	16	Yes	Yes
Average	11	11	13	30%	40%

Note: Bolded numbers indicate the highest reported externalizing behavior problems score for a each child. As this table illustrates, our reliance on teacher reports, parent reports, or selection of the highest score (regardless of whether provided by a parent or a teacher) would yield somewhat different portraits of each of these children and of this classroom, as a whole.