Characteristics of Existing Measures of Social and Emotional Development in Early Childhood

Applications for Federal Reporting and Data Collection

Authors: Kristen Darling-Churchill, Nina Chien, Tamara Halle, Laura Lippman*, Martha Zaslow, Paula Daneri, Samantha Goldhagen, Shannon Moodie, and Grace Howarth.

*Project Director

Table of Contents

Executive Summary

Social and emotional development in early childhood lays the foundation for children's development through middle childhood, adolescence, and beyond. Young children who are able to develop positive social relationships, feel confident in themselves, express and manage their emotions, and are more likely to be prepared to learn and succeed in school (C. C. Raver, 2002). Given these findings, policymakers are placing a growing emphasis on promoting positive social and emotional development for young children as part of state and federal initiatives.

The Federal Interagency Forum on Child and Family Statistics' 2013 publication *America's Children: Key National Indicators of Well-being (Federal Interagency Forum on Child and Family Statistics, 2013)* identifies a need for indicators of early childhood development. This report addresses that need by focusing on social and emotional indicators of child development. We review existing measures of early childhood social and emotional development, focusing on children birth through age 5, with special attention given to measures with the potential to assess and report on the school readiness of 4- to 5-year-old children. This review includes an evaluation of the quality and utility of the measures and their potential for use as: (1) a regularly-appearing indicator in the *America's Children* reports; (2) an indicator appearing a single time in a special feature of the *America's Children* reports; or (3) for use in future federal survey data collection. This report also serves to aid federal stakeholders in their efforts to address key concerns related to **time and respondent burden** when administering surveys; **cost** of developing, testing, and administering short measures; and potential **policy implications** of measurement efforts.

Social and emotional development in early childhood can be subdivided into subdomains to guide the review and selection of measures. Based upon our review of the literature, the following four subdomains were identified: (1) social competence; (2) emotional competence; (3) behavior problems; and (4) self-regulation. In addition, executive function was acknowledged as a key area of development underlying social and emotional skills.

The primary resource for this report is the *Inventory of Measures of Social and Emotional Development in Early Childhood* which describes key characteristics of 72 existing measures in a detailed, tabular format. This inventory was created through extensive research as well as input from multiple federal and academic stakeholders, and it documents which measures have been used in repeated cross-sectional surveys or one-time data collections in federal surveys or large-scale studies.

A set of 10 criteria were developed specifically to evaluate the measures in the inventory. These include:

- Strength of reliability statistics;
- Strength of validity statistics;
- Size and diversity of the norming/validation sample;
- Availability of the measure in languages other than English;
- Whether a trained administrator is required;
- Availability of a parent and/or teacher form;

- Whether the measure covers two or more of the four identified subdomains of social and Emotional development or executive function;
- Length of time to administer;
- Whether the measure has a cost for use; and
- Whether the measure covers a wide early childhood age range.

Based upon application of these criteria to the measures in the Inventory, we identify the strongest candidates for inclusion as one-time or repeated indicators in *America's Children* or in future federal data collection.

Forum agency staff and academic partners provided significant input on our application of the criteria, and guidance on the need to highlight specific characteristics of the measures. Based upon the review of 72 extant measures, 28 measures are recommended for use in some capacity by federal statistical agencies. Among these, 16 are rated as "strong" (meeting more than half of the 10 criteria) and 12 are deemed "promising" (meeting five of the 10 criteria).

In response to the priorities of federal partners, this review has clearly identified those measures that are short, valid and reliable, and useful for assessing school readiness of 4- and 5-year-old children. We find that, of the 28 recommended measures, *none* meet all three of these conditions; however, 11 are short (taking less than 10 minutes to administer) and seven have strong ratings for both reliability and validity. Five measures are designed for use specifically with 0- to 3-year-old children, and seven are appropriate for use with children ages four to five, to assess school readiness. The other 16 measures offer items or scales designed for use with children in both 0-3 and 4- to 5-year-old age groups.

Additionally, the measures often cover more than one area of development. Among the 28 recommended measures, 19 address social competence; 15 measure emotional competence; 17 address behavior problems; 16 cover self-regulation; and 10 assess executive function.

Specifically for interests of the Forum and the *America's Children* report, we find that:

- There are <u>no</u> measures of social and emotional development currently collected on a recurrent basis in federal data collections that could be immediately used as a *regular indicator* of social and emotional development for any of the four subdomains or executive function.
- One measure stands out for potential use as the basis of a *special feature* in *America's Children*: the **Preschool Learning Behaviors Scale (PLBS)**. This measure has subscales which address both self-regulation and executive function.
- **Five other measures** have a number of strengths and could also be considered for a *special feature*: the **Approaches to Learning Scale from ECLS-K**; the **Social Skills Rating Scales (SSRS)**; the **Social Skills Improvement System-Rating Scales (SSIS-RS)** (an updated version of the SSRS); **Rothbart's Children's Behavior Questionnaire (CBQ);** and the **CBQ-VSF (Very Short Form)**. Notably, the **SSIS-RS** covers three subdomains of social and emotional development (social competence, emotional competence, and behavior problems) and the **CBQ** and **CBQ-VSF** cover all 5 identified areas of development.

- Four measures from the Inventory can be recommended as possible measures to collect for the first-time in future federal surveys. These include the **Behavior Assessment System for Children, Second Edition (BASC-2), the Devereux Early Childhood Assessment Clinical Form (DECA-C), the Infant Toddler Social Emotional Assessment (ITSEA), and the Matson Evaluation of Social Skills with Youngsters (MESSY).** These measures each cover at least 3 of the 5 identified areas of development (the four subdomains or executive function).

Recommendations

With the tools and processes put in place by this project, many aspects of measuring and reporting on social and emotional development can be addressed in a rigorous fashion that enjoys the benefit of the collective wisdom of the field. This information can be used by the Forum moving forward to address measurement gaps, for example:

- For the purpose of federal monitoring of early childhood social and emotional development over time, a regular indicator for *America's Children* can be developed by fielding a short measure in a repeated cross-sectional survey.
- Existing measures used in smaller scale studies that measure social and emotional development across domains can be considered for federal longitudinal studies to round out the measures that are currently available, but which are focused on one or two subdomains.

Based upon the findings presented in this report, the following specific activities are priorities for future investments in data collection and reporting:

1. Update the Indicators Needed description for the Education section of the *America's Children* report to reflect the findings of this review. Specifically, this description might include an acknowledgement of this special study commissioned by the Forum, a summary of the subdomains that were examined, and identification of existing federal surveys with specific measures that are candidates for future indicator development, as mentioned above, and a link to the ChildStats.gov site for further information.

2. Explore the potential for reporting existing data on social and emotional development of young children currently collected by Forum agencies in an *America's Children report* special feature.

3. Promote continued dialog with academic and other non-federal research partners to adapt existing measures, validate brief measures and/or create new measures that would meet the needs for federal survey data collections.

In particular, measures are needed that will address all dimensions of social and emotional development but which are also brief, valid and reliable, and which are designed for use with children of different ages and developmental stages. The development of such measures will support Forum agency efforts to expand their monitoring and reporting on early childhood development, and provide high quality data on young children's social and emotional development. In turn, the Forum's indicators can serve as benchmarks for programs and policy makers at all levels of geography working to support children's transition to and successful progress through school.

I.   Introduction

<u>Purpose and Approach</u>

The Federal Interagency Forum on Child and Family Statistics (hereafter referred to as the Forum) has identified a need for measures of young children's social and emotional development as a federal data need since the first issue of its report, *America's Children: Key National Indicators of Well-being(Federal Interagency Forum on Child and Family Statistics, 1997)*. In 2012, the Forum contracted Child Trends to review existing measures of young children's social and emotional development that potentially could be used as indicators, and to identify gaps. The goal of this report is to provide guidance on the strengths and weaknesses of existing measures of young children's (ages 0-5) social and emotional development to inform the Forum's deliberations on measures for potential use in its monitoring report, *America's Children,* and to provide recommendations for future federal survey data collections.

This report focuses on issues which are particularly important to federal data collection. These include: **time and respondent burden** when administering measures; **cost** of developing, testing, and administering necessarily short measures; and potential **policy implications** of measurement efforts. To facilitate the efforts of federal stakeholders to address these issues, this report features specific information about the characteristics of the reviewed measures, and discussion of ways to meet the needs of federal statistical agencies.

Our primary resource for the present report is the *Inventory of Measures of Social and Emotional Development in Early Childhood* (2014)[1] prepared by Child Trends, summarizing key characteristics of 72 existing measures in a detailed, tabular format, and noting which measures have been used in repeated cross-sectional surveys or one-time data collections in federal surveys or large-scale studies. This inventory was created through extensive research as well as input from multiple federal and academic stakeholders.

In addition to the inventory and this report, other resources are available to support this process. These include memos written by experts in each sub-domain of social and emotional development, a summary of expert presentations at a Forum-sponsored May 2013 event held at NIH, and the presentations themselves, as well as documentation of discussions during the year following the event with the Forum's Research and Innovation Committee, federal partners, and academic experts.[2] While this report looks in greater breadth at the current landscape of measures of young children's social and emotional development, the memos provide a review of conceptual as well as measurement considerations for specific sub-domains of social and emotional development, and recommend measures used to assess each aspect of development.

---

[1] The Inventory of measures is available separately due to its length. Readers may access it via the Forum's website: childstats.gov.

[2] *"Memos on Measurement of Social and Emotional Development in Early Childhood April 2013"*and presentations from the May 2013 Culminating Activity are available via the Forum's website: childstats.gov.

Taken together, these resources provide a summary of measures that are currently available, and explore potential gaps that may need to be filled in terms of measures development in the social-emotional domain, focusing on children birth through age five, with special attention given to measures appropriate for assessing and reporting on school readiness of 4- to 5-year-old children. Each measure is reviewed for its potential as: (1) a regularly-appearing indicator in the *America's Children* reports; (2) an indicator appearing a single time in a special feature of the *America's Children* reports; or (3) for use in federal survey data collection.

Overview of Report

As background for the identification of candidate measures of social and emotional development, this report begins by briefly summarizing why young children's social and emotional development is important for concurrent and subsequent outcomes. We then present a framework for summarizing how the broad domain of social and emotional development can be subdivided into subdomains. We note the leading theoretical frameworks for discussing social and emotional development in the literature and, across these, the components of social and emotional development discussed with some consistency. The consistency of subdomains across theoretical frameworks helps us to move forward with some confidence in grouping the measures according to these subdomains for the purpose of identifying the focus, comprehensiveness, and usefulness of each measure. For the purposes of this discussion, we distinguish between subdomains and constructs, identifying the former as broad components of children's social and emotional development and the latter as more specific and measurable skills or characteristics.

This report provides an overview of the features of the 72 existing measures of children's social and emotional development summarized in the Inventory and reviews the measures in light of a set of 10 criteria as candidates for particular consideration. We organize report content around four subdomains identified below, as well as the related area of executive function, and note that coverage of multiple dimensions is regarded as a potential advantage due to the multi-faceted and inter-related nature of social and emotional development in early childhood, and in considering efficiency in data collection. For each measure meeting at least half of the 10 criteria, we summarize its particular strengths and limitations. Following application of this set of criteria to the measures included in the Inventory, we identify the strongest candidates for inclusion as one-time or repeated indicators in the Forum report, *America's Children: Key National Indicators of Well-being,* or in federal data collections.

We conclude by looking *across* the four subdomains and executive function, identifying those measures that meet half or more of our criteria, including special attention to the strength of reliability and validity information, whether the measures are short and thus easier for uptake by federal statistical agencies, and which age groups the measures address. In response to Federal and academic input on the importance of having measures available that tap into school readiness, measures appropriate for use with 4- to 5-year-old children are clearly distinguished; measures designed for use with children ages 0-3 are also included in this review to aid initiatives supporting healthy child development in *very* early childhood. In closing, we discuss the steps that would be needed to move

from candidate measures to indicators, using the lens of the characteristics of strong indicators articulated by the Forum.

<u>Limitations</u>

It is important to underscore the limitations of our review. First, while we used multiple approaches to develop a comprehensive listing of existing measures to review in the Inventory, we cannot claim that it is exhaustive. In addition to our own review of the literature, we received guidance from academic and federal partners in identifying measures. Although we reviewed the most well-known measures of social and emotional development in early childhood, we acknowledge that there may be lesser known, but high quality, measures that we did not have the opportunity to review. Second, the criteria have been applied to entire scales rather than to subdomains or constructs. Third, we have excluded screeners from this review (discussed below). Finally, the Forum criteria for indicators would need to be applied for any measure being considered for use as an indicator in *America's Children* by the Forum (for information on these criteria, refer to the end of this report or the "About this Report" section of the most recent edition of *America's Children*, found at <u>http://www.childstats.gov/</u>). As discussed in the conclusion, further work by the Forum partners to move from measures to indicators is necessary and ongoing.

II.   The Importance of Early Childhood Social and Emotional Development

The Center on the Social Emotional Foundations for Early Learning (CSEFEL) defines early social and emotional development as the developing capacity of the child from birth through five years of age to form close and secure adult and peer relationships; experience, regulate, and express emotions in socially and culturally appropriate ways; and explore the environment and learn—all in the context of family, community, and culture (Center on the Social Emotional Foundations for Early Learning, 2008). The emergence of social and emotional skills begins at birth and early experiences influence how

> *"When young children are able to develop prosocial relationships, feel confident in themselves, and express and manage their emotions, they are more likely to be prepared to learn and succeed in school."*

children begin to understand their world and themselves. For instance, when infants' needs are consistently met by adults, they are better able to regulate their emotions, pay more attention to their surroundings, and develop secure relationships. Research has shown that as children continue to develop these social and emotional skills, they gain the confidence and competence needed to build relationships, problem-solve, and cope with emotions (Parlakian, 2003; Shonkoff & Phillips, 2000). Furthermore, when young children are able to develop prosocial relationships, feel confident in themselves, and express and manage their emotions, they are more likely to be prepared to learn and succeed in school (C.C. Raver, 2002).

The value of healthy social and emotional development in young children is well established (Jackson, Higgins, Davidson, & Copper, 2009; Snow & Van Hemel, 2008; Thompson & Lagattuta, 2006) and we know that early social and emotional competencies have associations with later outcomes that are distinct from the patterns of association between early language and cognitive skills. Children's

social and emotional development during the early childhood years lays the foundation for their development through middle childhood, adolescence, and beyond (Denham, 2006; Denham & Brown, 2010; Shonkoff & Phillips, 2000). For example, having higher levels of behavioral self-control (an aspect of self-regulation, as discussed in greater detail later as a subdomain of social and emotional development) in early childhood is linked to multiple indicators of adult well-being. Children with greater self-control are more likely to grow into adults with better health (e.g., better physical health, less substance abuse), more wealth (e.g., higher income and SES, fewer financial struggles), and fewer criminal convictions (Moffitt et al., 2011). Social competence and emotional competence (also discussed later as subdomains of social and emotional development) include multiple aspects: social skills; positive interactions and relationships with teachers, family members, and peers; positive representations of self; understanding of emotions in self and others; appropriate responses to internal emotions and emotions expressed by others; and emotion regulatory abilities. These competencies often uniquely predict academic success, even when other factors, such as earlier academic success, are taken into account (Carlton, 1999; Jacobsen & Hoffman, 1997; Pianta, Steinberg, & Rollins, 1995; Shields et al., 2001). Conversely, certain behavior problems such as hyperactivity/impulsivity, anxiety, and depression have also been found to negatively predict later achievement (Romano, Babchishin, Pagani, & Kohen, 2010).

Policymakers are placing a growing emphasis on promoting positive social and emotional development for young children. The Office of Head Start has acknowledged the importance of social and emotional development since Head Start's inception in 1965, by establishing indicators and strategies for the development of positive self-concept, self-control, cooperation, social relationships, and knowledge of families and communities (see information on the social and emotional development domain from the *Head Start Leaders Guide to Positive Child Outcomes*). In addition, the Office of Special Education Programs, U.S. Department of Education, established a system of accountability and monitoring related to prescribed areas of child outcomes. For young children, birth through five, served under Part C and B/619 of the Individuals with Disabilities Act (IDEA), those child outcomes include positive social and emotional skills (including social relationships). States are required to report the percentages of infants and toddlers with Individual Family Service Plans (IFSPs) and preschool children with Individualized Education Programs (IEPs) who demonstrate improvements in this area (Fox & Smith, 2007).

Recently, the Administration for Children and Families released a memo outlining the core components of their "well-being framework" for children and youth receiving child welfare services (U. S. Department of Health and Human Services, 2012). Social and emotional development is prioritized within this framework because, as the memo suggests: (a) the challenges that children in the child welfare system face in these areas are great; (b) there are resources and policies that can be leveraged to improve child functioning in these areas; (c) effective practices and programs for promoting social and emotional well-being are available; and (d) outcomes for children can significantly improve with an emphasis on social and emotional well-being (U. S. Department of Health and Human Services, 2012).

Similarly, since its inception, the Race to the Top Early Learning Challenge Program, a grant initiative designed to improve the quality of early learning and development programs for children ages birth

through five, has emphasized states' development of standards related to early social and emotional development, in addition to other essential domains of school readiness. Of particular interest are children from low-income families, children with disabilities or developmental delays, and English language learners. A key component of the Early Learning Challenge Program is the requirement to implement a kindergarten entry assessment aimed at gaining an understanding of children's developmental status at school entry across the various domains. At the heart of the Race to the Top Early Learning Challenge Program is the understanding that early learning programs are a vital means to improving not only children's physical and cognitive development, but also their social and emotional development (U.S. Department of Education & U.S. Department of Health and Human Services, 2011).

III.   Early Childhood Social and Emotional Development: Theoretical Frameworks and Subdomains

The Child Trends team reviewed the recent literature on social and emotional development in early childhood to better understand the theoretical and conceptual lenses through which this aspect of child development is currently viewed. While conducting this review to identify prevalent *theoretical and conceptual frameworks*, we also sought information about how the domain of children's social and emotional development is subdivided into broad areas, labeled here as *subdomains* (to distinguish them from specific measurable skills and attributes, which we consider *constructs*), within each framework. It was a high priority for the Child Trends team to determine a consistent articulation of subdomains across selected resources in the field of social and emotional development research. The rationale for this approach is as follows:

1. Consistent articulation of subdomains provides a foundation for discussing aspects of children's social and emotional development that researchers have, to date, indicated as important for indicators or future data collection to address.
2. Consistent articulation of subdomains across existing social and emotional developmental research provides a structure for our review of measures, allowing us to identify which specific subdomains are addressed by each measure.

In the following sections, we first identify prevalent conceptual/theoretical frameworks pertaining to young children's social and emotional development identified in the literature. We then note the subdomains that are discussed consistently in selected literature in each conceptual/theoretical area. To briefly anticipate a key conclusion of this review, we find that there is consistency across frameworks in the identification of subdomains. Just as the introduction to this report underscored limitations that need to be taken into account in considering the overall recommendations, we wish to be clear here in stating that we have reviewed only *selected* publications in our review of theoretical frameworks. Appendix A provides references for the 14 selected frameworks and a summary of the subdomains identified across theoretical frameworks in a tabular format. Readers will see that we have attempted to review prominent consensus documents and resources. However, we emphasize that it was not possible with available time and resources to complete an exhaustive review of all theoretical/conceptual frameworks.

Our review of the social and emotional literature found two different theoretical and conceptual frameworks to be widely referenced and discussed in relation to social and emotional development. These included:

- The bioecological model
- Attachment theories and internal-working models of development

The *bioecological model* is a contextual-developmental model that views development as an interactive process that is embedded within the broader personal and environmental context in which a child lives (Bronfenbrenner & Morris, 2006). The contexts that affect development range from those systems closest to the child (e.g., the family, peer network, or classroom) to more distal systems (e.g., cultural, political, or historical) (Bronfenbrenner & Morris, 1998; Jones & Bouffard, 2012). This model is described as interactive in that there are bidirectional influences across contextual systems and between the child and his or her environment. For example, with regard to the family and school contexts, there is communication concerning the child both from home to school and from school to home. In relation to person-environment interactions, a child's level of temperamental reactivity may elicit certain care giving behaviors by parents that may, in turn, mitigate or exacerbate the temperamental reactivity. Across contextual-developmental frameworks, children's social and emotional development reflects social interactions and the quality of relationships within each of the embedded contexts as well as the linkages across the contexts (Chen, 2011; Stevenson-Hinde, 2011).

Attachment theories and internal-working models of development are also used to frame social and emotional development in early childhood. Attachment theorists posit that the quality of early attachments with caregivers is linked to children's subsequent social relationships and behaviors. Children's ability to form secure attachments with their parents and other caregivers is thought to be foundational to positive social adjustment (Ainsworth, 1979; Dykas & Cassidy, 2011). According to Bowlby's internal working model theory, children build the foundations of their social and emotional competence through mental representations of the self and others based on the quality of their early attachments (Bowlby, 1988). These internal working models then help the child predict and understand their social environments, engage in survival-promoting behaviors, and establish a sense of psychological security (Bowlby, 1969; Bretherton, 1990). Thus, the two theories work together: early attachment relationships become foundational to subsequent social relationships and behaviors through the development of mental representations of those early attachments known as internal working models.

Across these conceptual framings emerges the importance of establishing competencies in the social and emotional domain by means of social interactions embedded within contexts. There is a sense that social interactions are reciprocal or bidirectional, and growth in an individual's social and emotional skills (or movement to new skills) requires building on previous experiences and accomplishments. We turn now to a discussion of the subdomains of the social and emotional skills or competencies that are

discussed within the social and emotional literature, as well as one related area of development: executive function.

<u>Subdomains of Social and Emotional Development and Executive Function</u>

An additional goal of the literature review was to identify the subdomains regularly articulated as components of social and emotional development and, within these, the constructs (i.e., the specific, definable aspects of social and emotional development that an assessment can measure). As mentioned above, Appendix A summarizes in table format the 14 resources we reviewed (in the columns) and the common subdomains identified

> *"There is good agreement on which subdomains comprise this area of development: social competence, emotional competence, behavior problems, and self-regulation. Additionally, executive function is a key contributor to the successful execution of social and emotional competencies."*

in this literature (in rows, with the first row in the table noting the theoretical framework(s) cited in the source document). The cells of the table present the more specific constructs that are identified for each subdomain. It is especially important to note the consistency with which subdomains are identified across the literature, as seen in the table. Thus, while we have identified several theoretical frameworks for understanding children's early social and emotional development, there is good agreement on which subdomains comprise this area of development: social competence, emotional competence, behavior problems, and self-regulation. In this review, we also identify executive function – comprised of distinct constructs such as inhibitory control, working memory, and cognitive flexibility/attention shifting – as a key contributor to the successful execution of social and emotional competencies. While executive function currently lacks consensus definition in the field, it is included in this review due to its interrelationship with the development and execution of social skills (Bailey & Jones, 2013)See below for additional discussion on this aspect of development.

Social Competence

The subdomain of social competence is defined in the literature as the degree to which children are effective in their social interactions with others (Halle, et al., 2014). Specific constructs in this domain include children's pro-social skills and abilities, including: the ability to recognize social cues; interact positively with peers and adults through cooperation, listening, taking turns, and initiating and maintaining conversations; engage in social problem-solving; understand the rights of others; treat others equitably; distinguish between incidental and intentional actions; and balance one's own needs with the needs of others. Measures designed to assess these specific constructs include such adult-report items as "Greets people with 'Hi' or similar expression" or "Child doesn't talk or interact during group activities."

Emotional Competence

Emotional competence is defined in the literature as the ability to understand the emotions of self and others, react to others' emotions, and regulate one's own emotional expressiveness (Denham et al., 2002). This sub-domain includes emotional processes, emotional functioning, and emotional

development. Empathy and the ability to understand others' feelings are examples of specific constructs that fall within this subdomain. Measures designed to assess children's emotional competence might ask the child to say how he or she (or others) would feel in situations such as losing a favorite toy or being hit by another child.

Behavior Problems

Behaviors that were characterized as problematic in the selected literature are those developmentally inappropriate behaviors that may impede children's ability to adapt and function in their families, early care and education settings, or with a peer group (Campbell, 1998). These may include internalizing emotions or behaviors (i.e., worry, anxiety, sadness, and social withdrawal) or externalizing behaviors (i.e., hostility and aggression). Measures designed to assess whether children exhibit behavior problems might include adult-report items such as "Child is withdrawn from or avoids children/adults" or "Child acts aggressively when frustrated."

Self-Regulation

Self-regulation (also referred to as self-control or effortful control) was generally defined in the selected literature as the ability to focus attention, manage emotions, and control behaviors (Halle et al., 2014). This area of development is also highly inter-related to other aspects of social and emotional functioning, and also executive function, as discussed below. According to the resources reviewed, specific constructs within the self-regulation subdomain include the ability to shift and focus attention as needed, activate and inhibit behavior as required, and modulate behavioral and emotional reactivity in social interactions. Measures designed to assess children's self-regulation might include adult-report items such as "Child keeps working at something until he/she is finished" or "Child interrupts others when speaking."

Executive Function: A Related Area of Development

One other category of skill development was examined because of the growing recognition of its importance in supporting both social and emotional and cognitive development in early childhood. Executive function encompasses primarily cognitive processes including working memory, attention, and inhibitory control for the purposes of planning and executing novel problem solving and goal-directed activity (Diamond, 2006). However, strong evidence has emerged underscoring that the development of executive function skills is a crucial contributor to the development of both cognitive and social capacities (Center on the Developing Child, 2011).

Executive function is recognized as both a mediator and moderator of developmental outcomes and the field is currently exploring the inter-relationships among emotion regulation, self-regulation, and executive function processes (Bailey & Jones, 2013; Riggs, Jahromi, Razza, Dillworth-Bart, & Mueller, 2006; Williford, Whittaker, Vitiollo, & Downor, 2013; Willoughby, Wirth, & Blair, 2011). For example, the ability to control the expression of emotions – or acting upon those emotions – develops in tandem with executive control during early childhood (Carlson & Wang, 2007). Further adding to the complexity, at least one recent model posits *emotion* regulation as one aspect of overall *self*-regulation, the latter of

which includes the ability to inhibit behavioral responses driven by strong emotions (Bailey & Jones, 2013). Whereas executive function is seen as a support for behavioral self-regulation, it differs from self-regulation in that it focuses primarily on the processes required for the conscious control of thought, emotion, and action – including working memory and cognitive flexibility – but does not address the beneficial uses of emotions. Self-regulation, in contrast, addresses both suppressing disruptive emotions and encouraging positive emotions.

Despite the work that still needs to be done to clarify both the inter-relationships and the distinctions among these areas of development, preliminary evidence has shown that interventions targeting executive function skills have significant positive effects (Riggs et al., 2006; Williford et al., 2013). Therefore, measurement of this critical area of development is warranted. Measures designed to assess the development of executive function include reports of whether a child calms down quickly following an exciting event or can wait before entering into new activities when asked to. Note that these behaviors are not necessarily social in nature, but are related to social behaviors captured in the subdomains of self-regulation and emotional competence described above. Similarly, there is correspondence in "effortful control" (sometimes used as a synonym for self-regulation) and the process of inhibitory control (which is a component of executive function). As we will see in the review of existing measures that follows, this correspondence may lead researchers to use the same measure to capture two or more theoretically distinct developmental competencies.

IV.   Review of Measures for the Inventory

Having confirmed that the above-mentioned subdomains of young children's social and emotional development are regularly referenced in the literature, we identified existing measures that show the strongest potential – *for each of the four subdomains and for executive function* – as candidates for inclusion in *America's Children* reports (either as regularly appearing indicators or in special feature reports) or in federal survey data collections. Information about prominent measures of social and emotional development was compiled in the Inventory and developed earlier over the course of this project. The measures that were ultimately included in the Inventory were selected based upon reviewing previous work by the Forum's Research and Innovation Committee, Child Trends' expertise on extant measures (including both small- and large-scale data sets), and suggestions and additional resources shared by academic advisors and federal stakeholders.

A set of inclusion and exclusion criteria were applied to all suggested measures. Measures that were *included* in the inventory addressed one or more aspects of social and emotional development; were designed to be used with children ages birth to five (including measures that cover part of or overlapped with that age span); utilized a mode of administration that would be feasible for a national data collection; or were currently being used in national surveys, or showed promise for such use. Measures that were *excluded* failed to meet any of the above criteria. In addition, physiological/biological measures such as sleep regulation were excluded, as were measures designed exclusively to screen for behavior problems or developmental delays. Screeners were excluded as they are designed to be very precise only at the problematic end of the developmental range, to identify children as needing additional assessment, or not, and thus do not fit the stated purpose of the project to identify measures

that can be used to describe young children's normal social and emotional development. This decision was also influenced by recommendations from academic advisors and the literature that suggests that measures which capture the whole spectrum of social and emotional development (from negative to positive and across sub-domains) may be more predictive of positive educational outcomes than measures that identify only children at-risk (Denham, 2006).

Descriptive information about the characteristics of the measures was collected from a variety of sources, including developers' websites and manuals, validation studies, published compendia of social and emotional measures, literature reviews, journal articles, personal communications with federal survey staff, and information from technical/psychometric reports produced for particular national surveys. Information collected for each measure and reported in the inventory includes:

- Source and year of development;
- Whether the measure has been previously included in, or considered for, large surveys or studies (federal or non-federal);
- Targeted age range for which each measure was developed;
- Number of items;
- Sample items (if not-copyrighted or provided in publicly available resources);
- Subscales (and number of items included in each, if that information was available);
- Type of response categories;
- Languages in which the measure is available and information about the sample used for validation in each language;
- Duration of time needed for administration;
- Copyright information;
- Cost for use;
- Short form availability (if any);
- Mode of administration (parent and/or teacher report; clinician/trained observer; direct assessment);
- Reliability and validity, including the validation/norming sample.

Following the review of the literature on frameworks for early childhood social and emotional development, each measure was also characterized according to whether it addressed one or more of the four subdomains or executive function as described above.

V.   Criteria by Which Measures Were Evaluated

We turn now to a discussion of which of the 72 measures of young children's social and emotional development included in the Inventory are the strongest candidates for further consideration by the Forum for inclusion in *America's Children* (either as a regularly appearing indicator or as part of a special feature report) or in future survey data collection efforts. In order to identify candidates for consideration, the measures were reviewed in light of 10 criteria developed specifically to evaluate appropriateness for the purposes noted above. Some criteria are related to the quality of the measure (e.g., reliability and validity) while other criteria are related to how easily the measure could be used (e.g., time of administration). All of the criteria are considered significant when choosing a measure for

use in programs and in research (Halle, Zaslow, Wessel, Moodie, & Darling-Churchill, 2011; National Research Council, 2008).

The 10 criteria include:

- *Strength of reliability statistics* (See Table 1, below, and Appendices B & C for information on how reliability was evaluated and definitions of each type of reliability or validity). *Types of Reliability considered*: internal consistency reliability; interrater reliability; and test-retest reliability. As discussed in Appendix B, the criteria for reliability were developed in a review of early childhood measures conducted for another federal project and applied to the current review (Halle et al., 2011).
- *Strength of validity statistics* (See Table 1, below, and Appendices B & C for information on how validity was evaluated and definitions of each type of reliability or validity). *Types of Validity considered*: construct validity; content validity; convergent/concurrent validity; and predictive validity. As discussed in Appendix B, the criteria for validity were developed in a review of early childhood measures conducted for another federal project and applied to the current review (Halle et al., 2011). This review positively rates measures that have evidence of any type of validity; however, keeping in mind the priority to use this information to support school readiness, based on input from academic advisors those measures with documented evidence of *a lack of* predictive validity were excluded from recommendations.[3]
- *Size and diversity of the norming/validation sample*. Especially when considering use of a measure in a nationally representative survey sample, it was considered a strength if the norming/validation sample for the measure was large and diverse with respect to socioeconomic status and race/ethnicity.
- *Availability of the measure in languages other than English*. Because of the (growing) diversity of the population of young children and their families in the United States, including languages spoken at home, it was considered a strength if a measure was available for use with respondents in a language other than English.
- *Whether the measure requires a trained administrator*, indicating the ease or difficulty in implementing the measure in a federal survey. For a large national survey, it was considered an advantage if a measure did not require administration by someone with special training beyond high quality survey administration.
- *Availability of a parent and/or teacher form*. The availability of a parent form is seen as essential for having nationally representative data. A teacher form would be collected only in early care and education settings for a subset of children participating in such settings.
- *Whether the measure covers two or more of the four identified subdomains of social and emotional development or covers executive function*, indicating that the measure provides coverage across a significant portion of the spectrum of social and emotional development and related developmental areas. As noted above, deeming coverage of multiple subdomains a

---

[3] The result of this decision was the exclusion of one measure (the Child Development Inventory [CDI]) that would otherwise have been rated as Promising (see Table 1 below for explanation of ratings). Details about this measure are available in the full Inventory, but are not presented in the remainder of this paper.

"strength" may not be the case for certain studies or purposes. However, for the current discussion it is a strength in terms of efficiency and coverage in choosing a comprehensive measure of social and emotional development.

- *Length of time to administer*, with shorter times being more desirable in the context of survey data collection.
- *Whether the measure has a cost for use.* Measures requiring a fee may be a barrier for use by the federal statistical system; and,
- *Whether the measure covers a wide age range in early childhood*, indicating suitability for use of a single measure to report on the development of children across the age span. A wide age range was considered four or more years (so for example, birth through 3 years, 11 months). In response to the request by federal and academic partners to focus our recommendations on school readiness assessment, in section VIII of this paper we further refine our review to showcase which measures are designed for use with younger vs. older children, and 4- to 5-year-old children in particular.

Each measure was rated on each criterion and given one of the following ratings:

- *Strong* (strong on criterion pertaining to reliability or validity; or easy to use on criterion having to do with access to the measure, length, and need for training in implementation);
- *Moderate* (moderate; has some promising characteristics with room for minor improvement); or
- *Weak* (needs major improvement/more challenging to use).

For some criteria, only two dichotomous options were possible (e.g., availability in languages other than English, yes or no); in these cases, the measure was given a rating of *Strong* when the criterion was met, and either *Moderate* or *Weak* when the criterion was not met, depending on how serious not meeting a particular criterion was viewed. Table 1 provides additional details on how ratings were determined.

Table 1. Definitions of criteria for rating suitability of measures.

Criteria	Strong/Easy to use	Moderate/ Acceptable to use	Weak/ More challenging to use
Reliability*	"Acceptable" reliability for TWO OR MORE types of reliability.** Also considered to be "acceptable" are instances where no actual statistic was provided by the source, but the source described the criterion using language similar to "acceptable" (e.g., "good evidence," "adequate").	"Acceptable" reliability for ONE type of reliability** (but does not meet criteria for "strong/high"). Also considered to be "acceptable" are instances where no actual statistic was provided by the source, but the source described the criteria using language similar to "acceptable" (e.g., "good evidence," "adequate").	Does not meet criteria for "moderate."
Validity*	"Strong/high" (or "provides evidence") for TWO OR MORE types of validity**; OR, one "strong/high" AND one "moderate." Considered to be "moderate" are instances where significant relationships were found but no actual statistic was provided by the source.	"Strong/high" OR "moderate" for ONE type of validity** (but does not meet criteria for "strong/high"). Considered to be "moderate" are instances where significant relationships were found but no actual statistic was provided by the source.	Does not meet criteria for "moderate." Demonstrates evidence of poor predictive validity. This condition also precludes a measure from being recommended even if 5 other criteria are rated *Strong*.
Size and diversity of the of the norming/ validation sample	Large sample (>300) AND diverse with respect to race/ethnicity AND diverse with respect to SES; AND a U.S. sample.	Diverse with respect to race/ethnicity OR diverse with respect to SES (but does not meet criteria for "strong/high") OR large sample (>300).	Does not meet criteria for "moderate."
Availability in languages other than English	Available in other languages	Not available in other languages	NA
Requirement for a trained administrator	Does not require trained administrator/observer	NA	Requires trained administrator/observer

Criteria	Strong/Easy to use	Moderate/ Acceptable to use	Weak/ More challenging to use
Parent/teacher form	Has a parent form. If the measure requires a trained administrator, then this criterion is automatically 'NA,' not applicable.	Has teacher form only (no parent form). If the measure requires a trained administrator, then this criterion is automatically 'NA,' not applicable.	NA
Covers a range of social and emotional subdomains and/or executive function	Covers two or more of the four subdomains and/or executive function	Covers only 1 subdomain or executive function.	NA
Length of time to administer ***	Less than 10 minutes	10-20 minutes	More than 20 minutes
Cost/ Requirement for purchase	Purchase not required	Purchase required	NA
Covers a wide age range	Spans 4 or more years	Does not meet criteria for "strong/high"	NA

* For the Inventory, the reliability and validity information for each measure was interpreted based on the definitions of the different types of reliability and validity, and a set of criteria for each type, presented in *Understanding and Choosing Assessments and Developmental Screeners for Young Children: Profiles of Selected Measures* (Halle et al., 2011). See Appendix B for further information on how each type of reliability and validity were evaluated.

**This includes reliability/validity indices that are "acceptable" or "strong/high" with isolated exceptions for subgroups/subscales within the full measure.

*** If the duration is unknown, information about the number of items is provided and the measure is not rated for this criterion.

It is important to note that the cut points defined in Table 1 weight reliability, validity, and size and diversity of the norming sample equally with logistics of use (such as availability of a measure without cost, and availability of a measure in more than one language). They are given equal weight because even if a measure has strong reliability and validity and has been developed with a large and diverse sample, practical considerations, like length of administration, might make it impossible to use the measure in a national survey without further costly and time-consuming measures development work (for example, developing and testing the reliability and validity of a shortened version). While we choose not to give greater weight to reliability and validity issues in our

> *"Even if a measure has strong reliability and validity and has been developed with a large and diverse sample, a practical consideration like length of administration might make it impossible to use the measure in a national survey without further costly and time-consuming measures development work."*

overall approach, in subsequent sections of this paper we highlight when criteria that were deemed particularly important have been met. See Sections VII and VIII below for more information. We do however urge the reader to note instances where measures' reliability, validity, and norming samples are coded as *Moderate* or *Weak*. This has implications for federal agencies interested in using the measure. These details are available in Appendix D.

VI.   Review of Measures by Social and Emotional Subdomain and Executive Function

Appendix D provides detailed summaries, organized by subdomain or the related area of executive function, of the extent to which the 72 measures included in the Inventory meet the criteria in Table 1. For each of the subdomains and executive function, we present a summary of measures that are already included in recurrent federal surveys or studies that have potential for use as regularly reported *America's Children* indicators; then measures included in non-recurring federal surveys or studies that are promising for use as one-time indicators in *America's Children* special feature reports; and then measures that have not previously been used in federal surveys or studies but might be considered for such use. Each section clearly distinguishes whether the measures meet more than half (six or more) of the 10 criteria, as noted in Table 1 – referred to as "strong" – or only half of the 10 criteria (i.e., five of the 10 criteria). The latter are referred to as "promising." Furthermore, discussions related to the concerns of federal statistical agencies related to administrative burden (i.e. length of the measures or cost) are presented for each considered measure. The detailed discussions in Appendix D are summarized in Section VIII below ("Summary of Strong and Promising Measures").

VII.   Further Analyses and Considerations

Based on early reviews of the findings presented in Appendix D, Forum agencies provided input about the characteristics of measures that would make uptake by statistical agencies most feasible given limited resources. Representatives of ACF/OPRE, Census, HRSA, MCHB, NCES, NIH/NICHD, and CDC/NCHS discussed the following concerns and interests:

- Consideration of these measures may be accelerated if the measures have extant evidence of reliability and validity, and have already been used with a diverse norming sample; at a minimum, information about needs and gaps related to the strength of measures reliability, validity, and norming sample should be easily accessible.
- Many "short" extant measures that might be candidates for federal surveys are not short enough; to aid in planning discussion around data collection opportunities, our review would be enhanced by featuring the length of measures or duration of administration.
- Federal agencies have an interest in fielding measures that may be used by states focused on developing school readiness assessments, and thus would like an "at a glance" guide to measures appropriate for use during the preschool years.

In response to these considerations, additional detail highlighting relevant information is presented in the summary tables. These details are discussed in the next section.

VIII.   Summary of Strong and Promising Measures

While our review gave equal weight to practical criteria and criteria related to quality, such as reliability and validity, this section reviews the measures in light of the considerations raised by federal data collection agencies. This is because practical considerations, such as length of administration, have greater practical weight, and can take an otherwise strong measure out of the range of consideration when time or issues of administration exert constraints on federal data collection efforts.

Summary Tables 2a and 2b below present measures designed for use with children ages 0-3 (2a) and 4-5 (2b) respectively, and also provide information on which measures have strong ratings for reliability, validity, and norming sample, as well as which measures are short (i.e. take less than 10 minutes to administer). In these tables, we draw together those measures deemed strongest for consideration for use as an indicator in *America's Children* (focusing on measures already being collected in recurrent surveys), for consideration as a special feature in *America's Children,* and for inclusion for the first time in a federal survey.

Table 2a. Summary of Strongest and Promising Measures for use with children ages 0-3, by Recommended Use and by Social and Emotional Subdomains and Executive Function[4]

Key:
- ***Bolded*** *measures meet the criteria for a rating of* strong *for their reliability, validity, and representativeness of the norming sample.*
- *Measures with an asterisk (*) are short (<10 minutes).*
- *Measures with a plus sign (+) are appropriate for use across some portion of both the 0-3 and 4-5 age ranges.*

Measures Appropriate for Consideration for Use as Regular Indicator[5]	Measures Appropriate for Use as One-time indicator / Special Feature[6]	Measures to Consider for First Use in Federal Surveys[7]
Social Competence		
Meets 6 or more criteria: None[8] *Meets 5 criteria:* None	*Meets 6 or more criteria:* • Social Skills and Positive Approaches to Learning***+** • Social Skills Improvement System-Rating Scales (SSIS-RS)**+** • **Social Skills Rating System (SSRS)+** *Meets 5 criteria:* • Adaptive Social Behavior Inventory (ASBI)* • Nursing Child Assessment Teaching Scale (NCATS)* • Toddler Behavior Assessment Questionnaire (TBAQ) • Two Bags Task**+**	*Meets 6 or more criteria:* • Behavior Assessment System for Children, Second Edition (BASC-2) • Colorado Childhood Temperament Inventory (CCTI)***+** • **Devereux Early Childhood Assessment Clinical Form (DECA-C)+** • **Infant Toddler Social Emotional Assessment (ITSEA)** • Matson Evaluation of Social Skills with Youngsters (MESSY)**+** *Meets 5 criteria:* • **Minnesota Preschool Affect Checklist-Revised/Shortened (MPAC-R/S)+** • Penn Interactive Preschool Play Scales (PIPPS)**+** • Rothbart's Early Childhood Behavior Questionnaire (ECBQ) • Rothbart's Infant Behavior Questionnaire-Revised (IBQ-R) • Social Competence and Behavior Evaluation Scale: The Short Form (SCBE-30)**+**

---

[4] See "Criteria for EC SED Measures Selection and Evaluation," presented earlier in this paper and Appendix B which defines the thresholds for reliability and validity.

[5] Already being collected on a recurrent basis as part of a cross-sectional federal surveys or other data collection.

[6] Currently being used in longitudinal federal studies that assess the same participants over time.

[7] Candidates for consideration for use in new federal survey data collection (i.e., they are not currently being used in a regularly-occurring federal data collection).

[8] The Child Behavior Checklist (CBCL) technically meets 6 or more criteria; however, it measures many *negative* dimensions. We therefore do not recommend this measure.

Measures Appropriate for Consideration for Use as Regular Indicator[5]	Measures Appropriate for Use as One-time indicator / Special Feature[6]	Measures to Consider for First Use in Federal Surveys[7]
Emotional Competence		
Meets 6 or more criteria: None *Meets 5 criteria:* None	*Meets 6 or more criteria:* • Rothbart's Children's Behavior Questionnaire (CBQ)**+** • Rothbart's Children's Behavior Questionnaire-Very Short Form (CBQ-VSF)**+** • **Social Skills Rating System (SSRS)+** *Meets 5 criteria:* • Adaptive Social Behavior Inventory (ASBI)***+** • Nursing Child Assessment Teaching Scale (NCATS)***** • Toddler Behavior Assessment Questionnaire (TBAQ)	*Meets 6 or more criteria:* • Colorado Childhood Temperament Inventory (CCTI)***+** • **Devereux Early Childhood Assessment Clinical Form (DECA-C)+** • **Infant Toddler Social Emotional Assessment (ITSEA)** *Meets 5 criteria:* • **Minnesota Preschool Affect Checklist-Revised/Shortened (MPAC-R/S)+** • Rothbart's Early Childhood Behavior Questionnaire (ECBQ) • Rothbart's Infant Behavior Questionnaire-Revised (IBQ-R)
Behavior Problems		
Meets 6 or more criteria: None *Meets 5 criteria:* None	*Meets 6 or more criteria:* • Rothbart's Children's Behavior Questionnaire (CBQ)**+** • Rothbart's Children's Behavior Questionnaire-Very Short Form (CBQ-VSF)**+** • Social Skills Improvement System-Rating Scales (SSIS-RS)**+** • **Social Skills Rating System (SSRS)+** *Meets 5 criteria:* • Adaptive Social Behavior Inventory (ASBI)***+** • Toddler Behavior Assessment Questionnaire (TBAQ) • Two Bags Task**+**	*Meets 6 or more criteria:* • **Behavior Assessment System for Children, Second Edition (BASC-2)+** • **Devereux Early Childhood Assessment Clinical Form (DECA-C)+** • **Infant Toddler Social Emotional Assessment (ITSEA)** • Matson Evaluation of Social Skills with Youngsters (MESSY)**+** *Meets 5 criteria:* • **Minnesota Preschool Affect Checklist-Revised/Shortened (MPAC-R/S)+** • Penn Interactive Preschool Play Scales (PIPPS)**+** • Rothbart's Early Childhood Behavior Questionnaire (ECBQ) • Rothbart's Infant Behavior Questionnaire-Revised (IBQ-R) • Social Competence and Behavior Evaluation Scale: The Short Form (SCBE-30)**+**

Key:
- ***Bolded*** measures meet the criteria for a rating of strong *for their reliability, validity, and representativeness of the norming sample.*
- *Measures with an asterisk (*****) are short (<10 minutes).*
- *Measures with a plus sign (**+**) are appropriate for use across some portion of both the 0-3 and 4-5 age ranges.*

Measures Appropriate for Consideration for Use as Regular Indicator[5]	Measures Appropriate for Use as One-time indicator / Special Feature[6]	Measures to Consider for First Use in Federal Surveys[7]
Self-Regulation		
Meets 6 or more criteria: None *Meets 5 criteria:* None	*Meets 6 or more criteria:* • **Preschool Learning Behaviors Scale (PLBS)+** • Rothbart's Children's Behavior Questionnaire (CBQ)**+** • Rothbart's Children's Behavior Questionnaire-Very Short Form (CBQ-VSF)**+** *Meets 5 criteria:* • Adaptive Social Behavior Inventory (ASBI)***+** • Approaches to Learning - shortened version tested by the FACES team***+** • Toddler Behavior Assessment Questionnaire (TBAQ) • Two Bags Task**+**	*Meets 6 or more criteria:* • **Behavior Assessment System for Children, Second Edition (BASC-2**)**+** • Behavior Rating Inventory of Executive Function –Preschool Version (BRIEF-P)**+** • Colorado Childhood Temperament Inventory (CCTI)***+** • **Devereux Early Childhood Assessment Clinical Form (DECA-C)+** • **Infant Toddler Social Emotional Assessment (ITSEA)** • Matson Evaluation of Social Skills with Youngsters (MESSY)**+** *Meets 5 criteria:* • **Minnesota Preschool Affect Checklist-Revised/Shortened (MPAC-R/S) +** • Rothbart's Early Childhood Behavior Questionnaire (ECBQ) • Rothbart's Infant Behavior Questionnaire-Revised (IBQ-R)
Executive Function		
Meets 6 or more criteria: None *Meets 5 criteria:* None	*Meets 6 or more criteria:* • **Preschool Learning Behaviors Scale (PLBS)+** • Rothbart's Children's Behavior Questionnaire (CBQ)**+** • Rothbart's Children's Behavior Questionnaire-Very Short Form (CBQ-VSF)**+** *Meets 5 criteria:* • Dimensional Change Card Sort (DCCS)**+**	*Meets 6 or more criteria:* • **Behavior Assessment System for Children, Second Edition (BASC-2)+** • Behavior Rating Inventory of Executive Function –Preschool Version (BRIEF-P)**+** • Colorado Childhood Temperament Inventory (CCTI)***+** *Meets 5 criteria:* • Rothbart's Early Childhood Behavior Questionnaire (ECBQ) • Rothbart's Infant Behavior Questionnaire – Revised (IBQ-R)

Key:
- ***Bolded*** measures meet the criteria for a rating of *strong for their reliability, validity, and representativeness of the norming sample.*
- *Measures with an asterisk (****) are short (<10 minutes).*
- *Measures with a plus sign (+) are appropriate for use across some portion of both the 0-3 and 4-5 age ranges.*

Table 2b. Summary of Strongest and Promising Measures for use with children ages 4-5, by Recommended Use and by Social and Emotional Subdomains and Executive Function

Key:
- **Bolded** measures meet the criteria for a rating of strong *for their reliability, validity, and representativeness of the norming sample.*
- *Measures with an asterisk (*) are short (<10 minutes).*
- *Measures with a plus sign (+) are appropriate for use across some portion of both the 0-3 and 4-5 age ranges.*

Measures Appropriate for Consideration for Use as Regular Indicator[9]	Measures Appropriate for Use as One-time indicator / Special Feature[10]	Measures to Consider for First Use in Federal Surveys[11]
Social Competence		
Meets 6 or more criteria: None[12] *Meets 5 criteria:* None	*Meets 6 or more criteria:* • Social Skills and Positive Approaches to Learning***+** • Social Skills Improvement System-Rating Scales (SSIS-RS)**+** • **Social Skills Rating System (SSRS) +** *Meets 5 criteria:* • Adaptive Social Behavior Inventory (ASBI)***+** • Early Development Instrument (EDI) • Two Bags Task**+**	*Meets 6 or more criteria:* • **Behavior Assessment System for Children, Second Edition (BASC-2)+** • Colorado Childhood Temperament Inventory (CCTI)***+** • **Devereux Early Childhood Assessment Clinical Form (DECA-C)+** • Matson Evaluation of Social Skills with Youngsters (MESSY)**+** • Parent Daily Report (PDR) *Meets 5 criteria:* • **Minnesota Preschool Affect Checklist-Revised/Shortened (MPAC-R/S)+** • Penn Interactive Preschool Play Scales (PIPPS)**+** • Social Competence and Behavior Evaluation Scale: The Short Form (SCBE-30)**+**

---

[9] Already being collected on a recurrent basis as part of a cross-sectional federal surveys or other data collection.

[10] Currently being used in longitudinal federal studies that assess the same participants over time.

[11] Candidates for consideration for use in new federal survey data collection (i.e., they are not currently being used in a regularly-occurring federal data collection).

[12] The Child Behavior Checklist (CBCL) technically meets 6 or more criteria (and covers all 5 subdomains); however, it measures many *negative* dimensions. We therefore do not recommend this measure.

Measures Appropriate for Consideration for Use as Regular Indicator[9]	Measures Appropriate for Use as One-time indicator / Special Feature[10]	Measures to Consider for First Use in Federal Surveys[11]
Emotional Competence		
Meets 6 or more criteria: None *Meets 5 criteria:* None	*Meets 6 or more criteria:* • Rothbart's Children's Behavior Questionnaire (CBQ)**+** • Rothbart's Children's Behavior Questionnaire-Very Short Form (CBQ-VSF)**+** • **Social Skills Rating System (SSRS) +** *Meets 5 criteria:* • Adaptive Social Behavior Inventory (ASBI)***+** • Early Development Instrument (EDI)	*Meets 6 or more criteria:* • Brief Scale of Temperament (BST)***** • Colorado Childhood Temperament Inventory (CCTI)***+** • **Devereux Early Childhood Assessment Clinical Form (DECA-C)+** • Parent Daily Report (PDR)***** *Meets 5 criteria:* • **Minnesota Preschool Affect Checklist-Revised/Shortened (MPAC-R/S) +**
Behavior Problems		
Meets 6 or more criteria: None *Meets 5 criteria:* None	*Meets 6 or more criteria:* • Rothbart's Children's Behavior Questionnaire (CBQ)**+** • Rothbart's Children's Behavior Questionnaire-Very Short Form (CBQ-VSF)**+** • Social Skills Improvement System-Rating Scales (SSIS-RS)**+** • **Social Skills Rating System (SSRS) +** *Meets 5 criteria:* • Adaptive Social Behavior Inventory (ASBI)***+** • Two Bags Task**+**	*Meets 6 or more criteria:* • **Behavior Assessment System for Children, Second Edition (BASC-2)+** • **Devereux Early Childhood Assessment Clinical Form (DECA-C)+** • Matson Evaluation of Social Skills with Youngsters (MESSY)**+** • Parent Daily Report (PDR)***** *Meets 5 criteria:* • **Minnesota Preschool Affect Checklist-Revised/Shortened (MPAC-R/S)+** • Penn Interactive Preschool Play Scales (PIPPS)**+** • Social Competence and Behavior Evaluation Scale: The Short Form (SCBE-30)**+**

Key:
- ***Bolded*** *measures meet the criteria for a rating of* strong *for their reliability, validity, and representativeness of the norming sample.*
- *Measures with an asterisk (*) are short (<10 minutes).*
- *Measures with a plus sign (+) are appropriate for use across some portion of both the 0-3 and 4-5 age ranges.*

Measures Appropriate for Consideration for Use as Regular Indicator[9]	Measures Appropriate for Use as One-time indicator / Special Feature[10]	Measures to Consider for First Use in Federal Surveys[11]
Self-Regulation		
Meets 6 or more criteria: None *Meets 5 criteria:* None	*Meets 6 or more criteria:* • Approaches to Learning Scale from ECLS-K* • **Preschool Learning Behaviors Scale (PLBS)*+** • Rothbart's Children's Behavior Questionnaire (CBQ)**+** • Rothbart's Children's Behavior Questionnaire-Very Short Form (CBQ-VSF)**+** *Meets 5 criteria:* • Adaptive Social Behavior Inventory (ASBI)***+** • Approaches to Learning - shortened version tested by the FACES team***+** • Two Bags Task**+**	*Meets 6 or more criteria:* • **Behavior Assessment System for Children, Second Edition (BASC-2)+** • Behavior Rating Inventory of Executive Function –Preschool Version (BRIEF-P)**+** • Brief Scale of Temperament (BST)*** • Childhood Executive Function Inventory (CHEXI)*** • Colorado Childhood Temperament Inventory (CCTI)*** • **Devereux Early Childhood Assessment Clinical Form (DECA-C)+** • Matson Evaluation of Social Skills with Youngsters (MESSY)**+** *Meets 5 criteria:* • **Minnesota Preschool Affect Checklist-Revised/Shortened (MPAC-R/S)+**
Executive Function		
Meets 6 or more criteria: None *Meets 5 criteria:* None	*Meets 6 or more criteria:* • **Preschool Learning Behaviors Scale (PLBS)*+** • Rothbart's Children's Behavior Questionnaire (CBQ)**+** • Rothbart's Children's Behavior Questionnaire-Very Short Form (CBQ-VSF)**+** *Meets 5 criteria:* • Dimensional Change Card Sort (DCCS)* **+**	*Meets 6 or more criteria:* • **Behavior Assessment System for Children, Second Edition (BASC-2)+** • Behavior Rating Inventory of Executive Function –Preschool Version (BRIEF-P)**+** • Childhood Executive Function Inventory (CHEXI)***+** • Colorado Childhood Temperament Inventory (CCTI)***+** *Meets 5 criteria:* None

Key:
- ***Bolded*** measures meet the criteria for a rating of strong *for their reliability, validity, and representativeness of the norming sample.*
- *Measures with an asterisk (*) are short (<10 minutes).*
- *Measures with a plus sign (+) are appropriate for use across some portion of both the 0-3 and 4-5 age ranges.*

Summary of Most Highly Rated Measures Presented by Reporting Option

Regular Indicator in *America's Children*

The summary provided in Tables 2a and 2b indicates that there are no measures of social and emotional development currently being collected on a recurrent basis within federal data collections that could be immediately used as a regular indicator of social and emotional development across all four subdomains and executive function.

Readers should note that one measure included in the inventory was not recommended for the Forum's use: the Child Behavior Check List (CBCL). The CBCL items assess mostly negative behaviors, and are thus not a good fit for the intent of this project. Please note that the information appearing in the Inventory and used to assess the strength of the CBCL is based on the full measure and not the abbreviated version, called the "Mental Health Indicator," that is currently being used in the National Health Interview Study (NHIS). There have been limited validation studies of the Mental Health Indicator (Achenbach, n.d.), and is considered by Achenbach to be a screener rather than an assessment measure. At this time, it is unclear whether the abbreviated version has similar psychometric properties to the full measure.

One-time Special Feature

In terms of an indicator that could be used as a special feature in America's Children, there appears to be only one measure that currently stands out for recommendation above the others. The **Preschool Learning Behaviors Scale (PLBS)** has strengths in a majority of the 10 criteria considered in our review, including strong reliability, strong validity, and a diverse norming sample. The PLBS has subscales which address both self-regulation and executive function (i.e., "attitude toward learning," "competence motivation," and "attention/persistence"). The measure uses teacher report, and there is a Spanish version as well as an English version. In addition, the PLBS only takes 8 minutes to collect. Although not appropriate for use with infants and toddlers, data on children ages 3 to 6 could be reported on, from the nationally representative sample of Head Start children who were assessed using the measure in the FACES and Head Start Impact Study data collections (however, items from this measure are combined with items from other measures to create composite subscales of "Behavior Problems" and "Social Skills and Positive Approaches to Learning" within the FACES and HSIS datasets). We do note, however, that ACF Forum representatives advised caution against producing a special feature for *America's Children* solely focusing on FACES data, because Head Start children represent only a small subgroup of the national population. If utilized as part of a special feature, PLBS data from the FACES sample would benefit from methodological discussion providing context for the appropriateness with this specific population of children. Based on this concern, Forum members may want to consider expanding the use of this measure to additional federal data collection efforts to gain information on a broader sample of young children.

There are also several other extant measures that could be considered for a one-time special feature indicator in America's Children. Specifically, the **Approaches to Learning Scale from the ECLS-K** has strengths in seven out of the 10 criteria considered in this review, including a strong validation

sample, ease of administration, parent and teacher forms, and a wide age span; limitations for this measure include the need to strengthen reliability and validity statistics. However, the Forum recently produced a special feature for the 2013 edition of *America's Children* that included data on approaches to learning drawing from this measure.

In addition, the **Social Skills Rating Scales (SSRS)** has strengths in more than half the criteria considered in this review and was fielded (albeit in a modified form) in the ECLS-K and the ECLS-K: 2011. It could provide a very current special feature covering three subdomains of social and emotional development: social competence, emotional competence, and behavior problems. Similarly, the **Social Skills Improvement System-Rating Scales (SSIS-RS)** (an updated version of the SSRS has "acceptable" internal consistency and test-retest reliability, and is considered valid for use with special populations, although it also has modest convergent validity and low/weak interrater reliability. The norming sample was adequately diverse, and the measure is an updated version of the Social Skills Rating System (SSRS, Gresham & Elliott, 1991), which was fielded in an adapted form in the ECLS-K and the ECLS-K: 2011. Additional work on the SSIS-RS would also be needed to shorten the measure.

Rothbart's Children's Behavior Questionnaire (CBQ) and CBQ-VSF (Very Short Form), while also having strengths in multiple scoring criteria, have only moderate psychometric properties (with limited information on exact *p* values for validity). Like the SSRS, selected and adapted items of the CBQ are gathered in the ECLS-K: 2011 and it covers multiple social and emotional subdomains (emotional competence, self-regulation, and behavior problems) as well as executive function.

Finally, there were two measures that did not meet the minimum requirements for inclusion in this review (five or more ratings of *Strong* across the 10 criteria) but which did receive strong ratings for <u>four</u> criteria AND are designed for use with 4- to 5-year-old children. If the criteria requiring measures to cover a wide age span (at least four years) were relaxed, allowing a measure to cover just the 4-5 age span, these measures could be promoted to a rating of "promising." These measures are the **Behavior Problems Index (BPI)** and the **Emotion Regulation Checklist**. Like the CBCL discussed above, the BPI items primarily assess negative behaviors, and thus are not a good fit for the intent of this project. However, the Emotion Regulation Checklist covers three of the four subdomains, has strong validity statistics, does not require a trained administrator, and offers both a parent and teacher form. This measure warrants mention and consideration for school readiness assessment purposes.

New Federal Data Collection

Measures that can be recommended from the Inventory as possible first-time measures to collect in future federal surveys include the **Behavior Assessment System for Children, Second Edition (BASC-2), the Devereux Early Childhood Assessment Clinical Form (DECA-C), the Infant Toddler Social Emotional Assessment (ITSEA),and the Matson Evaluation of Social Skills with Youngsters (MESSY).**

The Behavior Assessment System for Children, Second Edition (BASC-2) stands out with strengths in 8 of the 10 criteria considered in our review, including strong reliability, strong validity, and a diverse norming sample. The BASC-2 has subscales which address three subdomains – social competence, behavior problems, self-regulation – and executive function (i.e., "attitude toward learning,"

"competence motivation," and "attention/persistence"). The measure offers both parent and teacher report, and there is a Spanish version as well as an English version. Its primary drawback for federal purposes is that the BASC-2 takes 10-20 minutes to collect. However, it covers a large age-span (preschool form for 2- to 5-year-old children and child form for 6- to 11-year-old children) which would be advantageous for household surveys.

The DECA-C also was normed on a large, diverse sample of preschoolers, is appropriate for use with children ages two to five, has acceptable internal consistency and test-retest reliability and strong construct validity, and covers the subdomains of self-regulation, emotional competence, and behavior problems. As with the ITSEA, the major drawbacks for consideration of the DECA-C in federal data collection include the cost and length of time of administration.

The ITSEA is appropriate for use with children 12-36 months of age and gathers information that address social competence, self-regulation, and behavior problems (the ITSEA covers four domains labeled "externalizing," "internalizing," "dysregulation," and "compliance"). The ITSEA was normed on a large and diverse sample, and although its psychometrics vary, it has "acceptable" or strong internal consistency, interrater reliability, test-retest reliability, construct validity, and concurrent validity. Other benefits of this measure are that it can be completed by both parents and teachers, and is available in both English and Spanish. However, the drawbacks to using the ITSEA in new federal data collection include the cost and the length of time for administration.

The MESSY has strong reliability statistics and a diverse norming sample; however, the only information on validity ranges from weak/low to only moderate. It is available in languages other than English and offers both a parent and teacher form. And while the measure may be used with a broad age range (2-18 years), it would need modification to shorten it from its current 64 items.

IX.   Conclusion

Consideration of Criteria for America's Children Indicators

In addition to the 10 criteria used in reviewing indicators for this report, another lens must be considered. Specifically, an indicator must meet the following criteria specified by the Forum for inclusion in the *America's Children* report:

- Easy to understand by broad audiences;
- Objectively based on reliable data with substantial research connecting them to child well-being;
- Balanced, so that no single area of children's lives dominate the report;
- Measured regularly, so that they can be updated and show trends over time; and
- Representative of large segments of the population, rather than one particular group.

Some of the *America's Children* criteria are addressed by this review. For example, measures featured in this paper are based, by and large, upon substantial research connecting them to child well-being outcomes, and information on their reliability, representativeness, and the periodicity of data collection is easily accessible. However, any measure selected for consideration by federal statistical agencies will need to be reviewed in light of the two remaining guidelines. Specifically, agencies need to

ensure that data collected using these measures will be presented in a manner that is easily understood by policy makers and the general public. Further, the content must provide a balanced picture of children's development in the context of other indicators already in the report.

What We Have Learned from Federal and Academic Partners

Throughout this project, a shared commitment to realizing early childhood development measurement goals within the federal statistical system was clearly evident. Regular communication between Child Trends and Forum agencies revealed multiple future data collection opportunities, and the inventory of measures has already served as a resource for federal survey and study development and redesign activities.

> *"Throughout this project, a shared commitment to realizing early childhood development measurement goals within the federal statistical system was clearly evident."*

However, federal partners expressed that resources – in terms of both funds and time – are currently very limited. In response to this concern, this review has taken care to clearly identify in our final recommendations those measures that are short, valid and reliable, and useful for assessing school readiness of four and five year olds. Additionally, while research supports the value of using measures that assess multiple, inter-related areas of social and emotional development, academic stakeholders suggest that federal agencies with limits to their staffing and survey activities might opt to focus on collecting high-quality data on a single subdomain or on executive function, rather than using measures that cover multiple aspects but which yield lower quality information. It should also be noted that some academics question the ability of survey frameworks to adequately capture relevant aspects of development, such as executive function (Willoughby, personal communication 2014).[13]

Summary of Findings from our Review

This report presents the results of a review of 72 measures of young children's social and emotional development, and the application of a set of 10 criteria assessing the characteristics of these measures. This process led to the identification of strong extant measures with potential for use as one-time *America's Children* indicators as well as for inclusion in future data collection. It also reveals the work that is needed to improve extant measures so that they can be considered for indicators and for future data collection. For example, it is noteworthy that there are <u>no</u> measures in the summary tables above that could be considered as a regular indicator in *America's Children*. In particular, a measure is needed that will address all aspects of social and emotional development but which is also brief, valid, and reliable. However, care must be taken to ensure that items or scales in such a measure can clearly distinguish theoretically distinct developmental competencies (Willoughby, personal communication 2014).

---

[13] Provided by academic advisor Dr. Michael Willoughby, memo author for the May 2013 Culminating Activity.

<u>Contributions of this Project</u>

This project has moved the measurement of social and emotional development in young children forward by:

- Providing a consensus framework for the field to consider, which heretofore had competing conceptualizations of subdomains of development.
- Systematically applying the subdomain classifications in reviewing existing measures.
- Developing consensus on ten criteria for reviewing the quality of the measures, creating an inventory of measures, and a paper making recommendations based upon our review.
- Organizing and facilitating a meeting that brought together key experts and federal stakeholders to discuss criteria for measurement, existing measures, and research and policy priorities.
- Facilitating exchange between federal partners and academic experts to prioritize the needs for new data collections and the measures to be considered for development.
- Coordinating the federal collaboration on a potential indicator for AC and recommending and suggesting refinements for measures where opportunities existed within new data collections.

<u>Suggested Next Steps for Federal Agencies</u>

The review provides not only a framework and conclusions regarding characteristics of the current landscape of measures, but also a system for evaluating future measures developed by researchers. This information supports Forum agency efforts to expand their monitoring and reporting efforts to include the development of young children. With the tools and processes put in place by this project, many aspects of measuring and reporting on social and emotional development can be addressed in a rigorous fashion that enjoys the benefit of the collective wisdom of the field.

Moving forward, the Forum can consider addressing measurement gaps, for example:

- For the purpose of federal monitoring of early childhood social and emotional development over time, a regular indicator for *America's Children* can be developed by fielding a short measure in a repeated cross-sectional survey.
- Existing measures used in smaller scale studies that measure social and emotional development across domains can be considered for federal longitudinal studies to round out the measures that are currently available, but which are focused on one or two subdomains.

Specifically, the findings of this project suggest the following potential next steps:

1. Update the Indicators Needed description for the Education section of the *America's Children* report to reflect the findings of this review. Specifically, this description might include an acknowledgement of this special study commissioned by the Forum, a summary of the areas of development that were examined, and identification of existing federal surveys with specific measures that are candidates for future indicator development, as mentioned above, and a link to the ChildStats.gov site for further information.

2. Continue collaborative discussions exploring the potential for reporting extant data on social and emotional development of young children currently being collected by Forum agencies. Within these discussions, address methodological considerations specific to certain survey designs or samples which currently act as limitations to reporting opportunities.

3. Engage in conversations with academic and other non-federal research entities about the potential for collaboration to adapt extant measures (for example, shortening them), validate brief measures, and/or create new measures that would be more feasible for use by federal survey data collections.

Gathering and reporting additional high quality data across subdomains of young children's social and emotional development will benefit programs and policy makers supporting children's transition to and progress through school. The measurement of social and emotional development across multiple subdomains as well as for children of different ages and developmental stages must be strengthened in order to meet the Nation's need for monitoring our most valuable resource—children.

X.   References

Achenbach, T. M. (n.d.). A preliminary evaluation and recommendations for use of the Mental Health Indicator in the NHIS. Center for Disease Control and Prevention. Retrieved from http://wonder.cdc.gov/wonder/sci_data/surveys/nhis/type_txt/smch97a.htm.

Ainsworth, M. (1979). Infant-mother attachment. *American Psychologist, 34*, 932-937. doi: 10.1037/0003-066X.34.10.932

Bailey, R., & Jones, S. (2013). *SECURe: An applied developmental model for social-emotional and self-regulation-related skills from birth to 3rd grade*. Paper presented at the Graduate School of Education, Salt Lake City, UT.

Blair, C. (2002). Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist, 57*(2), 111-127.

Bowlby, J. (1969). *Attachment and loss: Vol. 1*. New York, NY: Basic Books.

Bowlby, J. (1982). Attachment and loss: Retrospect and prospect. *American Journal of Orthopsychiatry, 52*, 664-678. doi: 10.1111/j.1939-0025.1982.tb01456.x

Bowlby, J. (1988). *A secure base: Parent-child attachment and healthy human development*. New York, NY: Basic Books.

Bretherton, I. (1990). Open communication and internal working models: Their role in attachment relationships. In R. Thompson (Ed.), *Socioemotional development*. Lincoln, NE: University of Nebraska Press.

Bronfenbrenner, U., & Morris, P. A. (1998). The ecology of developmental processes. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Vol. 1. Theoretical models of human development* (pp. 993-1028). Hoboken, NJ: Wiley.

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner & W. Damon (Eds.), *Handbook of child psychology: Vol. 1. Theoretical models of human development* (pp. 793-828). Hoboken, NJ: Wiley.

Bryck, R. L., & Fisher, P. A. (2012). Training the brain: Practical applications of neural plasticity from the intersection of cognitive neuroscience, developmental psychology, and prevention science. . *American Psychologist, 67(2)*(Feb-Mar 2012), 87-100.

Campbell, S. (1998). Developmental perspectives. In T. H. Ollendick & M. Herson (Eds.), *Handbook of Child Psychopathology* (3 ed., pp. 3-36). New York: Springer Science + Business Media, Inc.

Carlson, S. M., & Wang, T. W. (2007). Inhibitory control and emotion regulation in preschool children. *Cognitive Development, 22*(2007), 489-510.

Carlton, M. P. (1999). Motivation and school readiness in kindergarten children. *Dissertation Abstracts International, 60*, 3899A.

Center on the Developing Child. (2011). Building the brain's "air traffic control" system: How early experiences shape the development of the executive function. Working Paper No. 11. Boston, MA: Author.

Center on the Social Emotional Foundations for Early Learning. (2008). Handout 1.7: CSEFEL Definition of social emotional development. Retrieved from http://csefel.vanderbilt.edu/resources/training_infant.html.

Chen, X. (2011). Culture and children's socioemotional functioning: A contextual-developmental perspective. In X. Chen & K. H. Rubin (Eds.), *Socioemotional development in cultural context* (pp. 29-52). New York, NY: Guilford.

Chen, X., & Rubin, K. H. (Eds.). (2011). *Socioemotional development in cultural context*. New York, NY: Guilford.

Chien, N., Daneri, P., Goldhagen, S., Darling Churchill, K. E., & Howarth, G. (2014). Inventory of existing measures of social and emotional development in early childhood. Washington, D.C.: Child Trends.

Denham, S. A. (2006). Social-emotional competence as support for school readiness: What is it and how do we assess it? . *Early Education & Development, 17*, 57-89.

Denham, S. A., & Brown, C. (2010). "Plays nice with others": Social-emotional learning and academic success. *Early Education & Development, 21*, 652-680.

Denham, S. A., Caverly, S., Schmidt, M., Blair, K., DeMulder, E., Caal, S., & Mason, T. (2002). Preschool understanding of emotions: Contributions to classroom anger and aggression. *Journal of Child Psychology and Psychiatry, 43*(7), 901-916.

Dykas, M. J., & Cassidy, J. (2011). Attachment and the processing of social information across the life span: Theory and evidence. *Psychological Bulletin, 137*, 19-46. doi: 10.1037/a0021367.

Federal Interagency Forum on Child and Family Statistics. (1997). America's Children: Key National Indicators of Well-Being. Washington, DC: Government Printing Office.

Federal Interagency Forum on Child and Family Statistics. (2013). America's Children: Key National Indicators of Well-Being. Washington, DC: U.S. Government Printing Office.

Fox, L., & Smith, B. J. (2007). Issue Brief: Promoting social, emotional and behavioral outcomes of young children served under IDEA. Center on the Social and Emotional Foundations for Early Learning.

Halle, T., Vick Whittaker, J., Zepeda, M., Rothenberg, L., Wessel, J., Anderson, R., & Daneri, P. (2014). The Social-Emotional Development of Dual Language Learners: Looking Back at Existing Research and Moving Forward with Purpose. *Early Childhood Research Quarterly*. doi: 10.1016/j.ecresq.2013.12.002

Halle, T., Zaslow, M., Wessel, J., Moodie, S., & Darling-Churchill, K. (2011). Understanding and choosing assessments and developmental screeners for young children: Profiles of selected measures. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Jackson, E. A., Higgins, L. B., Davidson, L. L., & Copper, J. L. (2009). Indicators for Social - Emotional Development in Early Childhood: A guide for local stakeholders. New York: National Center for Children in Poverty.

Jacobsen, T., & Hoffman, V. (1997). Children's attachment representations: Longitudinal relations to school behavior and academic competency in middle childhood and adolescence. *Developmental Psychology, 33*, 703-710.

Jones, S. M., & Bouffard, S. M. (2012). Social and emotional learning in schools: From programs to strategies. *Society for Research in Child Development Social Policy Report, 26*, 1-32.

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., . . . Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 108*(7), 2693-2698.

National Research Council. (2008). *Early childhood assessment: Why, what, and how* (C. E. Snow & S. B. Van Hemel Eds.). Washington, DC: Board on Children, Youth, and Families, Board on Testing and Assessment, Division of Behavioral and Social Sciences and Education: The National Academies Press.

Parlakian, R. (2003). Before the ABCs: Promoting school readiness in infants and toddlers. Washington, DC: Zero to Three.

Pianta, R., Steinberg, M., & Rollins, K. (1995). The first two years of school: Teacher-child relationships and deflections in children's classroom adjustment. *Development and Psychopathology, 7*, 295-312.

Raver, C. C. (2002). Emotions matter: Making the case for the role of young children's emotional development for early school readiness. *Social Policy Report of the Society for Research in Child Development, 16*, 1-20.

Raver, C. C. (2002). A good beginning: Making the case for the role of young children's emotional development for early school readiness. Ann Arbor, MI: Society for Research in Child Development.

Riggs, N., Jahromi, L., Razza, R. P., Dillworth-Bart, J., & Mueller, U. (2006). Executive function and the promotion of social–emotional competence. *Journal of Applied Developmental Psychology, 27*(2006), 300-309.

Romano, E., Babchishin, L., Pagani, L. S., & Kohen, D. (2010). School readiness and later achievement: Replication and extension using a nationwide Canadian survey. *Developmental Psychology, 46*(5), 995-1007.

Shields, A., Dickstein, S., Seifer, R., Giusti, L., Magee, K. D., & Spritz, B. (2001). Emotional competence and early school adjustment: A study of preschoolers at risk. *Early Education and Development, 12*, 73-96.

Shonkoff, J., & Phillips, D. A. (2000). *From neurons to neighborhoods: The science of early childhood development*. Washington, DC: National Academy Press.

Snow, C. E., & Van Hemel, S. B. (2008). Early Childhood Assessment: Why, what and how. . In N. R. C. Committee on Developmental Outcomes and Assessments for Young Children (Ed.). Washington, D.C. : National Academies Press.

Stevenson-Hinde, J. (2011). Culture and socioemotional development, with a focus on fearfulness and attachment. In X. Chen & K. H. Rubin (Eds.), *Socioemotional development in cultural context* (pp. 11-28). New York, NY: Guilford.

Thompson, R. A., & Lagattuta, K. H. (2006). Feeling and Understanding: Early emotional development *K. McCartney & D. Phillips (Eds.), Blackwell handbook of early childhood development* (pp. 317-337). Malden, MA: Blackwell Publishing.

U. S. Department of Health and Human Services. (2012). Information memorandum: Promoting social and emotional well-being for children and youth receiving child welfare services. Washington, DC: Author.

U.S. Department of Education, & U.S. Department of Health and Human Services. (2011). Race to the Top – Early Learning Challenge: Executive Summary. Washington, DC: Author.

Waters, E., & Sroufe, L. A. (1983). Social competence as a developmental construct. *Developmental Review, 3*, 79-97.

Weissberg, R. P., & Greenberg, M. T. (1998). School and community competence-enhancement and prevention programs. In E. Siegel & A. Renninger (Eds.), *Handbook of child psychology: Vol. 4. Child psychology in practice* (pp. 955-998). New York, NY: Wiley.

Williford, A., Whittaker, J., Vitiollo, V., & Downor, J. (2013). Children's Engagement Within the Preschool Classroom and Their Development of Self-Regulation. *Early Education & Development, 24*(2), 162-187.

Willoughby, M. T., Wirth, R. J., & Blair, C. B. (2011). Contributions of modern measurement theory to measuring executive function in early childhood: An empirical demonstration. *Journal of Experimental Child Psychology, 108*(2011), 414-435.

XI. Appendix A: Review of Selected Literature to Identify Conceptual Frameworks and Common Subdomains of Social and Emotional Development

The subdomains described in this Appendix are those identified by the authors of the papers reviewed; from these we derived the four subdomains of social and emotional development used in the evaluation of measures and in the paper itself. We also included in our review measures that might capture aspects of the related developmental area of executive function. As explained in the text of the paper (see pages 8-9), we included measures of executive function because it is related to the development of social and emotional competencies and is even considered by some scholars to be foundational to those competencies (Bailey & Jones, 2013; Williford et al., 2013). As noted by our paper, there is some overlap in the operationalization of executive function and aspects of social and emotional development, most notably, the concepts of self-control or self-regulation. This overlap is acknowledged on page 9 of the paper.

Source	Title: **Promoting Social and Emotional Well-being for Children and Youth Receiving Child Welfare Services** Author: **Administration for Children and Families (ACF)** Year: **2012**	Title**: Culture and Children's Socioemotional Functioning: A Contextual-Developmental Perspective** Author**: Xinyen Chen** (in Socioemotional Development in Cultural Context, edited by Xinyen Chen and Kenneth Rubin) Year: **2011**	Title: **Social-Emotional Competence as Support for School Readiness** Author: **Susanne Denham** Year: **2006**
Theoretical/ Conceptual Framing	This source cites the Well-Being Framework developed by Lou, Anthony, Stone, Vu, and Austin (2008) and adapts it for the agency's purposes. The framework identifies four domains of well-being: 1) cognitive functioning; 2) physical health and development; 3) behavioral/emotional functioning; and 4) social functioning. The framework takes into account contextual factors that influence children, including environmental supports and personal characteristics.	This work proposed a contextual-developmental perspective that focuses on the role of social interaction in the peer context in bridging culture and socioemotional development. In this perspective, social interaction in dyadic, group, and larger settings is an important context that mediates cultural influence on individual development. The social evaluation and regulation processes in interactions play a role in building and facilitating links between cultural norms and values on the one hand, and the development of various behaviors and characteristics on the other hand.	This work draws on Rose-Krasnor's (1997) prism model: The topmost construct is effectiveness in interaction; the middle construct is success in both intrapersonal and interpersonal goals (e.g., qualities of relationships, group status, and social self-efficacy); the bottom level of the prism includes the specific behaviors, social cognitive abilities, and motivations that form part of the evaluations of self and others regarding one's social effectiveness.
Subdomains	**Emotional functioning** • Self-control • Self-esteem • Emotional management Emotional expression **Social functioning** • Social competencies • Attachment and caregiver relationships • Adaptive behavior **Behavioral functioning** • Internalizing and externalizing behaviors	**Self-control** • The ability to modulate behavioral and emotional reactivity in social interactions **Social initiative** • The tendency to initiate and maintain social participation (as often indicated by reactivity in challenging situations)	**Emotional competence** • Emotional expressiveness • Understanding of emotion • Regulation of emotion and behavior **Social competence** • Social problem-solving • Social skills • Relationship skills

Source	Title: **Compromised Emotional Competence: Seeds of Violence Sown Early?** Authors: **Susanne Denham, Kimberly Blair, Michelle Schmidt, and Elizabeth DeMulder** Year: **2002**	Title: **"Plays Nice with Others": Social-Emotional Learning and Academic Success** Authors: **Susanne Denham & Chavaughn Brown** Year: **2010**
Theoretical/ Conceptual Framing	The authors of this work posit that the foundation of emotional competence is security of attachments and the internal working model (mental representations of the self and others that help individuals predict and understand their environment, engage in survival promoting behaviors such as proximity maintenance, and establish a psychological sense of "felt security") (Bowlby, 1969); (Bretherton, 1990).	In this work social and emotional development is viewed through an organizational, bioecological lens, in which different developmental tasks are central to each age level (Waters & Sroufe, 1983; Weissberg & Greenberg, 1998). Transitions from one developmental period to another are marked by reorganization around new tasks but are also based on the accomplishments of the earlier period. This source also utilizes an adaptation of **Rose-Krasnor's (1997)** theorizing **(prism model)**. In this adaptation, the definition of the construct (social and emotional learning; SEL) at the model's topmost level is effectiveness in interaction, the result of organized behaviors that meet short- and long-term developmental needs. The next level refers to the context and assessment of SEL, the bottom level includes the very specific SEL abilities, behaviors, and motivations that form part of the evaluations of self and others regarding one's SEL effectiveness—all of which are primarily individual.
Subdomains	**Emotional competence** • Understand others' emotions • Ability to react to others' emotions • Regulate emotional expressiveness	**Self-management** • Ability to handle one's emotions in productive ways • Awareness of feelings • Handling stress • Persevering despite obstacles • Expressing emotions appropriately **Relationship skills** • Making positive overtures to play with others • Initiating and maintaining conversations • Cooperating • Listening • Taking turns • Seeking help • Developing friendship skills (e.g., joining another child/small group, expressing appreciation, negotiating, or giving feedback) **Social Awareness** • Ability to take others' perspectives • Ability to understand others' feelings and empathize with them • Ability to appreciate others' similarities and differences

Source	Title: **Self-regulation and School Readiness** Authors: **Nancy Eisenberg, Carlos Valiente, Natale Eggum** Year: **2010**	Title: **The Social-Emotional Development of Dual Language Learners: Looking Back at Existing Research and Moving Forward with Purpose** Authors: **Tamara Halle, Jessica Vick Whittaker, Marlene Zepeda, Laura Rothenberg, Rachel Anderson, Paula Daneri, Julia Wessel, and Virginia Buysse** Year: **forthcoming**	Title**: Social and Emotional Learning in Schools: From Programs to Strategies (SRCD Social Policy Report)** Authors**: Stephanie Jones and Suzanne Bouffard** Year: **2012**
Theoretical/ Conceptual Framing	This work describes a heuristic model of the relations among effortful control, social relationships, (mal) adjustment, school engagement, and academic competence.	This work describes contextual-developmental models, in which social development is both bi-directional and transactional and occurs within the context of relationships with parents, non-parental caregivers and teachers, and peers (Chen & Rubin, 2011). This work also draws on attachment theory which suggests that children's ability to form secure attachments with parents and other caregivers is foundational to positive social adjustment (Ainsworth, 1979); (Bowlby, 1982); (Dykas & Cassidy, 2011).	This work is framed within developmental-contextual models, which view development as taking place in a nested and interactive set of contexts ranging from immediate (e.g., family, peer system, classroom, school) to more distal (e.g., cultural, political contexts).
Subdomains	**Effortful Control** • Ability to shift and focus attention • Ability to activate and inhibit behavior as required • Other executive functioning skills involved in integrating information, planning, and modulating emotion and behavior	**Self-regulation** • Ability to focus attention • Ability to manage emotions • Ability to control behaviors **Social competence** • The degree to which children are effective in their social interactions with others • Ability to understand how one relates to others • Ability to interact in a social situation **Problem behaviors** • Internalizing-worry, anxiety, sadness, and social withdrawal • Externalizing-hostile, aggressive physical behavior, impulsivity, hyperactivity	**Cognitive regulation** • Attention control • Inhibiting inappropriate responses • Working memory • Cognitive flexibility **Emotional processes** • Emotional knowledge • Emotional expression • Emotional and behavioral regulation • Empathy • Perspective-taking **Social/interpersonal skills** • Understanding social cues • Interpreting others' behaviors • Navigating social situations • Interacting positively with peers and adults

Source	Title: **Early Childhood Assessment: Why, What and How** Author: **National Research Council of the National Academies** Year: **2008**	Title: **Reconsidering Children's Early Development and Learning: Toward Common Views and Vocabulary** Author: **National Education Goals Panel-Goal 1 Technical Planning Group (Sharon Lynn Kagan, Evelyn Moore, and Sue Bredekamp)** Year: **1995**	Title: **Low-Income Children's Self-Regulation in the Classroom: Scientific Inquiry for Social Change (2012)** Author: **C. Cybele Raver** Year: **2012**
Theoretical/ Conceptual Framing	N/A	N/A	The framework cited in this work suggests that children's cognitive and affective regulation at the behavioral and neurobiological levels is relatively malleable or plastic in the face of changing environmental conditions (Blair, 2002); (Bryck & Fisher, 2012). Raver suggests that we can leverage the large amount of time that many low-income children spent in out-of-home settings in ways to support their self-regulation, which may in turn improve their chances for academic success.
Subdomains	**Self-regulation** • Modulating thought, affect, and behavior by means of deliberate as well as automated responses **Social competence** • Effectiveness in social interactions with others **Maladjustment** • Clusters of symptoms that may impede the child's ability to adapt and function in the family and peer group	**Emotional Development** • Self-concept traits, habits, abilities, motives, goals, and values that define how we perceive ourselves • Self-efficacy-the belief that one can successfully accomplish what one sets out to do • Ability to express feelings appropriately • Sensitivity to the feelings of others (empathy) **Social Development** • Cooperation • Understanding the rights of others • Ability to treat others equitably • Ability to distinguish between incidental and intentional actions • Willingness to give and receive support • Ability to balance one's own needs with those of others	**Self-regulation** • The primarily volitional regulation of attention, emotion, and executive functions for the purposes of goal-directed actions

Source	Title: **Promoting Children's Socioemotional Development in Contexts of Early Educational Intervention and Care: A Review of the Impact of Federally-Funded Research Initiatives on Young Children's School Readiness (2008)** Author: **C. Cybele Raver** Year: **2008**	Title: **Kindergarten Social-Emotional Competence: Developmental Predictors and Pyschosocial Implications** Authors: **Michelle Schmidt, Elizabeth DeMulder, and Susanne Denham** Year: **2002**	Title: **Culture and Socioemotional Development, with a Focus on Fearfulness and Attachment** Author: **Joan Stevenson-Hinde** (in Socioemotional Development in Cultural Context, edited by Xinyin Chen and Kenneth Rubin) Year: **2011**
Theoretical/ Conceptual Framing	The author cites social developmental and neurobehavioral literatures that suggest that children enter schools with distinct profiles of emotional reactivity, regulations, and executive functioning that facilitate or hinder their engagement with other learners, teachers, and the process of learning (Blair, 2002; C.C. Raver, 2002). This work also draws on attachment theory which suggests that some children establish and maintain relationships with teachers that are characterized by high degree of mutual positive engagement while others engage in relationships characterized by high level of conflict.	This work draws on Bowlby's (1988) description of internal working model which explains how individuals develop unconscious representations of their caregivers and themselves, and use these developing, self-perpetuating representations as models for choosing their own behaviors, understanding others' behaviors, and seeking out partners. This work also cites attachment theory that posits that early attachments with caregivers and subsequent social relationships and behavior are linked.	This work draws from Bronfenbrenner's (1998) framework which emphasizes dialectical relations among various levels of analysis, from internal physiological and psychological processes, to individual behavior, interactions, relationships, groups, and society. Each level influences and is influenced by the sociocultural context, with its shared values and conventions, as well as by the socioeconomic context (availability of resources such as education, work, and social mobility). This is a dynamic framework with bidirectional influences (like family systems theories that stress continuous cycles of interactions within a system that involves both stability and change).
Subdomains	**Self-regulation** • How children handle their emotions, attention, and behavior **Social competence** • Knowledge of emotions • Negotiation of interpersonal problems **Behavior problems** • Internalizing • Externalizing	**Social competence** (no additional information)	

XII. Appendix B: Criteria/Thresholds for each type of reliability and validity

For the Inventory, the reliability and validity information for each measure was interpreted based on the definitions of the different types of reliability and validity, and a set of criteria for each type, presented in *Understanding and Choosing Assessments and Developmental Screeners for Young Children: Profiles of Selected Measures* (Halle et al., 2011). The table below describes each type of reliability or validity and the criteria used to determine whether measures were weak, moderate, or strong. Sources consulted to develop these criteria follow the chart. Definitions for each type of reliability or validity are presented in Appendix C.

Type of Reliability or Validity	Description and Source of Evidence Used to Establish Criteria	Criterion and Terminology Used
Validity		
Construct Validity	Measured by examining associations between subscales within the assessment or screener instrument. Also measured by examining associations between subscale scores and child characteristics, such as age. No established standard in the field	0.50 or higher=strong/high 0.30 – 0.49=moderate 0.29 or below=weak/low
Content Validity	Measured by whether tool was reviewed by experts to determine if content reflects what the assessment or developmental screener is supposed to be measuring.	Content was or was not reviewed by experts
Convergent/Concurrent Validity	Measured by correlating the scores of the assessment/ developmental screener with scores on other assessments/ developmental screeners of similar content to determine the strength of relationships between the two. Source: Administration for Children and Families (2003)	0.50 or higher=strong/high 0.30 – 0.49=moderate 0.29 or below=weak/low
Predictive Validity	Measured by correlating the scores of the assessment/developmental screener with scores on later assessments to determine how well the assessment/developmental screener predicts later achievement or development. Source: Administration for Children and Families (2003)	0.40 or higher=provides evidence that measure may predict later achievement or development 0.39 or below=does not provide evidence that measure predicts later achievement or development

Type of Reliability or Validity	Description and Source of Evidence Used to Establish Criteria	Criterion and Terminology Used
Reliability		
Internal Consistency Reliability	Measured by correlating items within a construct to determine the interrelatedness of the items. No established standard in the field	0.70 or higher=acceptable 0.69 or below=low/weak
Interrater Reliability	Measured by the level of agreement between two raters when assessing the same children. No established standard in the field	0.80 or higher=acceptable 0.79 or below=low/weak
Test-Retest Reliability	Measured by correlating the scores on two administrations of the same assessment/ developmental screener given to the same child within a short period of time to determine consistency. No established standard in the field	0.70 or higher=acceptable (across a period of three months or less) 0.69 or below=low/weak

Sources Consulted in Determining Cut points for Reliability/Validity

Bentler, P.M., & Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.

Brown, M.W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K.A. Bollen & J.S. Long (Eds.), *Testing structural equation models*. Newbury Park, CA: Sage.

Carmines, E.G., & Zeller, R. A. (1979). *Reliability and validity assessment*. Sage University Series on Quantitative Applications in the Social Sciences, 07-001. Beverly Hills, CA, and London: Sage.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.) Hillsdale, NJ: Lawrence Erlbaum.

Council on Children with Disabilities, Section on Developmental Behavioral Pediatrics, Bright Futures Steering Committee, Medical Home Initiatives for Child with Special Needs Project Advisory Committee. (2006). Indentifying infants and young children with developmental disorders in the medical home: An algorithm for developmental surveillance and screening. *Pediatrics 118*(1), 405-420.

Lewis-Beck, M.S. (1995). *Data analysis: An introduction*. Thousand Oaks, CA: Sage.

Robinson, J.P., & Shaver, P.R., Wrightsman, S. (1991). *Measures of personality and social psychological attitudes.* San Diego, CA: Academic Press.

Snow, C.E., & Van Hemel, S.B. (Eds.). (2008). *Early childhood assessment: What, why, and how*. Washington, DC: National Research Council of the National Academies.

Administration for Children and Families (2003). *Resources for measuring services and outcomes in Head Start Programs Serving Infants and Toddlers.* E. Kisker, K. Boller, C. Nagatashi, C. Sciarrino, V. Jethwani, T. Zavitsky, M. Ford, J. Love, & Mathematica Policy Research, Inc. Washington, DC: U.S. Department of Health and Human Services. Retrieved February 15, 2009 from http://www.acf.hhs.gov/programs/opre/ehs/perf_measures/reports/resources_measuring/res_meas_cdi.html.

XIII. Appendix C: Definitions of the different types of reliability and validity

Excerpted from *Understanding and Choosing Assessments and Developmental Screeners for Young Children: Profiles of Selected Measures* (Halle et al., 2011).

Concurrent validity – This term describes the relationship between two separate measures of similar constructs which, when administered at the same time, provide results that are consistent with one another. Note: Sometimes manuals refer to this as convergent criterion validity, which could be interpreted to mean that the two tools concur or agree in the measurement of a particular construct.

Construct validity – The extent to which a tool measures a clearly defined theoretical concept. The instrument should be based on a theory, and scores from the instrument should reflect what would be expected based on that theory.

Content validity – The extent to which a tool reflects the range of possible skills or behaviors that make up the domain or construct being assessed. This is often determined through expert review.

Convergent validity – A subtype of criterion-related validity. This term indicates the degree to which a tool correlates with other tools assessing the same construct.

Internal consistency reliability – How closely items or indicators within a construct are interrelated.

Interrater reliability – How similar the results of an assessment are when different individuals administer the same assessment with the same child.

Predictive validity – The degree to which items relate to an established measurement tool that assesses the same, or a related, outcome at a later point in time.

Test-retest reliability – An indicator of whether the tool will yield the same score across two administrations of the tool within a short period of time. This tells us whether the tool provides a consistent assessment of a skill, regardless of other factors, such as the child's mood or health, the time of day, or the time of year that the child was assessed. A child should score similarly (within a defined range) if tested within a short period of time, usually defined as within three months.

XIV. Appendix D – Detailed Review of Measures by Social and Emotional Subdomains and Executive Function

The sections that follow provide a detailed summary of the extent to which the 72 measures reviewed in the Inventory meet the criteria in Table 1. Selected measures, organized by the four identified subdomains and executive function, are presented below in the following order:

1. Social competence
2. Emotional competence
3. Behavior problems
4. Self-regulation
5. Executive function

The first table in each section provides a visual guide to which measures met which criteria. These tables include all measures discussed in the subdomain and executive function section. Please note that individual measures which assess more than one subdomain and/or executive function – which we consider to be a strength – are described in each applicable section. Following the first table for each subdomain and executive function discussion, we present short tables describing the strengths and limitation of each measure, organized in subsections by their three potential uses:

- Measures currently included in recurrent federal surveys or studies with the potential for use as regularly reported *America's Children* indicators;
- Measures that are not collected regularly by federal surveys or studies but are promising for use as one-time indicators in *America's Children* special feature reports; and,
- Measures which might be considered for first use in federal surveys or studies.

Each subsection highlights first the measures that meet more than half (six or more) of the 10 criteria, as noted in Table 1, with ratings in the *Strong* range. We refer to these measures as "strong." The second part of each subsection highlights additional measures that do not meet six criteria but nonetheless could be considered for use. Measures that meet only half of the 10 criteria (i.e., five of the 10 criteria) are referred to as "promising."

Note that due to space limitations in this document, the following discussions present only selected details about measures. Full details about each measure's characteristics may be viewed in the Inventory document.

Social Competence

Table SC1 below displays the measures that assess social competence that meet criteria for use in the *America's Children* report or federal surveys. Measures were reviewed using the 10 criteria indicated in the column headings and rated as *strong*, *moderate* or *needs improvement*. We first identify the measures that met six or more criteria in the *strong* range. We then identify additional "promising" measures; that is, those that received a rating of *strong* on five of the 10 selected criteria. **See Table 1 for detailed descriptions of criteria used for evaluation.**

Based on their ratings, a total of 19 measures (out of 39 reviewed that assess social competence) are suggested for consideration either for the *America's Children* report or for federal survey data collection. Additional information on each measure's strengths and limitations is outlined in the descriptions following the summary chart.

Table SC1. Selected measures of social competence that meet at least half of the evaluation criteria

Name of Measure	Reliability	Validity	Size and diversity of norming/ validation sample	Availability in languages other than English	Requires trained administrator	Parent/teacher form	Covers 2 or more subdomains	Length of time to administer	Cost / purchase requirement	Covers a wide age range
Adaptive Social Behavior Inventory (ASBI)	Moderate	Strong	Moderate	Moderate	Strong	Strong	Strong	Strong	No info	Moderate
Behavior Assessment System for Children, Second Edition (BASC-2)	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Moderate	Moderate	Moderate
Colorado Childhood Temperament Inventory (CCTI)	Moderate	Moderate	Weak	Moderat	Strong	Strong	Strong	Strong	Strong	Strong
Devereux Early Childhood Assessment Clinical Form (DECA-C)	Strong	Strong	Storng	Moderate	Strong	Strong	Strong	Moderate	Moderate	Strong
Early Development Instrument (EDI)	Strong	Moderate	Moderate	Strong	Strong	Moderate	Strong	Moderate	Strong	Moderate

Name of Measure	Reliability	Validity	Size and diversity of norming/ validation sample	Availability in languages other than English	Requires trained administrator	Parent/teacher form	Covers 2 or more subdomains	Length of time to administer	Cost / purchase requirement	Covers a wide age range
Infant Toddler Social Emotional Assessment (ITSEA)	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Weak	Moderate	Moderate
Matson Evaluation of Social Skills with Youngsters (MESSY)	Strong	Moderate	Strong	Strong	Strong	Strong	Strong	No info; 64 items	No info	Strong
Minnesota Preschool Affect Checklist-Revised/Shortened (MPAC-R/S)	Strong	Strong	Strong	Moderate	Weak	N/A	Strong	Weak	Strong	Moderate
Nursing Child Assessment Teaching Scale (NCATS) - Child Scale	Strong	Moderate	Strong	Moderate	Weak	N/A	Strong	Strong	No info	Strong
Parent Daily Report (PDR)	Weak	Moderate	Moderate	Strong	Strong	Strong	Strong	Strong	No info	Strong
Penn Interactive Preschool Play Scales (PIPPS)	Strong	Strong	Moderate	Moderate	Strong	Strong	Strong	Moderate	Moderate	Moderate
Rothbart's Early Childhood Behavior Questionnaire (ECBQ)[14]	Moderate	Moderate	Moderate	Strong	Strong	Strong	Strong	Weak	Strong	Moderate
Rothbart's Infant Behavior Questionnaire Revised (IBQ-R)[15]	Moderate	Moderate	Moderate	Strong	Strong	Strong	Strong	Weak	Strong	Moderate

[14] The ECBQ also offers a Short Form (107 items) and a Very Short Form (36 items). Information about the duration needed to administer the short forms is not available. The rating in this chart is for the full 201 item measure.

[15] The IBQ-R also offers a Short form (91 items) and a Very Short form (36 items). Information about the duration needed to administer the short forms is not available. The rating in this chart is for the full 184 item measure.

Name of Measure	Reliability	Validity	Size and diversity of norming/ validation sample	Availability in languages other than English	Requires trained administrator	Parent/teacher form	Covers 2 or more subdomains	Length of time to administer	Cost / purchase requirement	Covers a wide age range
Social Competence and Behavior Evaluation Scale: The Short Form (SCBE-30)	Strong	Moderate	Moderate	Strong	Strong	Moderate	Strong	Moderate	Moderate	strong
Social Skills and Positive Approaches to Learning	Moderate	Strong	Strong	Strong	Strong	Strong	Moderate	Strong	Moderate	Moderate
Social Skills Improvement System-Rating Scales (SSIS-RS)	Strong	Moderate	Strong	Strong	Strong	Strong	Strong	Moderate	Moderate	Stromg
Social Skills Rating System (SSRS)[16]	Strong	Strong	Strong	Moderate	Strong	Strong	Strong	No info; 58 items	Moderate	Strong
Toddler Behavior Assessment Questionnaire (TBAQ)	Moderate	Strong	Moderate	Strong	Strong	Strong	Strong	No info; 108 items	No info	Moderate
Two Bags Task	Moderate	Weak	Strong	Strong	Weak	N/A	Strong	Moderate	Strong	Strong

[16] The reliability and validity information collected for the SSRS includes some information from the ECLS-K and ECLS-K: 2011 studies (which use selected and adapted items from the SSRS) as well as from documentation for the full measure.

<u>Measures Appropriate for Consideration for Use as Regular Indicator</u>

None of the measures of social competence are currently appropriate for consideration for use as a regular indicator in *America's Children* since they are not included in federal surveys that occur regularly.

<u>Measures Appropriate for Consideration for Use as One-time Indicator</u>

Eighteen measures included in the full Inventory assess social competence and have been used in federally sponsored surveys or large scale studies. However, they are not found in repeated cross-sectional surveys, but rather in longitudinal studies (for example, they have been included in the Early Childhood Longitudinal Study Birth Cohort (ECLS-B)). As a result, they might be considered for use as one-time indicators for a special feature of *America's Children*. Use in a repeated cross-sectional survey is necessary for a regularly- appearing indicator in the *America's Children* reports.

Among measures of social competence that are currently used by regularly occurring longitudinal federal surveys, eight meet five or more of the 10 criteria. Of those eight, three measures score a *strong* rating on six or more of the 10 criteria: the Social Skills and Positive Approaches to Learning, the Social Skills Improvement System-Rating Scales (SSIS-RS), and the Social Skills Rating Scales (SSRS), with selected and adapted items fielded in the ECLS-K and ECLS-K:2011. These three measures are thus recommended for consideration as one-time indicators. The table below summarizes the criteria on which the measures showed strengths and limitations.

Table SC2. Social competence: *Strong* measure(s) to consider for use as a one-time indicator

Name of measure	Strengths	Limitations
Social Skills and Positive Approaches to Learning, fielded in the FACES in 2006 and Head Start Impact study in 2008	1. Strong validity 2. Size and diversity of norming/validation sample 3. Available in English and Spanish 4. No training required 5. Parent form 6. Short time of administration/few items	1. Moderate reliability 2. Covers 1 subdomain 3. Purchase required 4. Narrow age range (3-5 years)
Social Skills Improvement System-Rating Scales (SSIS-RS), fielded in ECLS-K: 2011	1. Strong reliability 2. Size and diversity of norming/validation sample 3. Available in English and Spanish 4. No training required 5. Parent and teacher forms 6. Covers 2 or more subdomains (social competence and behavior problems) 7. Covers a wide age range (3-18 years)	1. Moderate validity 2. Takes a moderate amount of time to complete (10-25 minutes) 3. Purchase required

Name of measure	Strengths	Limitations
Social Skills Rating System (SSRS), selected and adapted items fielded by ECLS-K:1998 and FACES	1. Strong reliability 2. Strong validity 3. Size and diversity of norming/validation sample 4. No training required 5. Parent and teacher forms 6. Covers 2 or more subdomains (social competence, emotional competence, and behavior problems) 7. Broad age range (3-18 years)	1. English version only 2. Many items (58) 3. Purchase required[17]

The other five measures are promising. The tables below summarize the criteria on which the measures showed strengths and limitations. For each measure, the national survey in which it was used is noted with the name of the measure.

Table SC3. Social Competence: *Promising* measure(s) to consider for use as a one-time indicator

Name of measure	Strengths	Limitations
Adaptive Social Behavior Inventory (ASBI), fielded in the NICHD SECCYD Phase I	1. Strong validity 2. No training required 3. Parent and teacher forms 4. Covers two or more subdomains (social competence, emotional competence, behavioral problems, and self-regulation) 5. Short administration time (30 items)	1. Moderate reliability 2. Validated on non-diverse sample 3. English version only 4. Limited age range ("young, high-risk children")
Early Development Instrument (EDI), fielded in the National Longitudinal Survey of Children and Youth	1. Strong reliability 2. Available in multiple languages 3. No training required 4. Covers two or more subdomains (social competence and emotional competence) 5. No purchase required	1. Moderate validity 2. Validated on non-diverse sample 3. No parent form (teacher form only) 4. Takes a moderate amount of time to complete (20 minutes) 5. Limited age range (4-6 years)

---

[17] Since this measure is currently used in an existing federal study, purchase will not be required for using this measure as the basis of an indicator.

Name of measure	Strengths	Limitations
Nursing Child Assessment Teaching Scale (NCATS) - Child Scale, fielded in the ECLS-B in 2007	1. Strong reliability 2. Size and diversity of norming/validation sample (ECLS-B) 3. Covers 2 or more subdomains (social competence and emotional competence) 4. Short administration time (1-6 minutes) 5. Broad age range (Birth-3 years)	1. Moderate validity 2. English version only 3. Requires trained administrator
Toddler Behavior Assessment Questionnaire (TBAQ), fielded in the NICHD Study of Early Child Care and Youth Development in 1994	1. Strong validity 2. Available in English and French 3. No training required 4. Has parent form 5. Covers 2 or more subdomains (social competence, emotional competence, behavior problems, and self-regulation)	1. Moderate reliability 2. Validated on non-diverse sample 3. Limited age range (16-36 months) 4. Many items (108)
Two Bags Task, fielded in ECLS-B	1. Size and diversity of norming/validation sample 2. Available in languages other than English (coders may speak the parent and child's language) 3. Covers 2 or more subdomains (social competence, behavioral problems, and self-regulation) 4. No fee 5. Covers a wide age range (2 years to preschool)	1. Moderate reliability 2. Weak validity 3. Trained administrator required 4. Requires a moderate amount of time to complete (10 minutes for task and then coding later)

Measures to Consider for First Use in Federal Surveys

Eleven measures included in the full Inventory assessed social competence and were evaluated as candidates for consideration for use in new federal survey data collection (i.e., they are not currently being used in a regularly-occurring federal data collection). Of those, six measures have more strengths than limitations, scoring a *strong* rating on more than half of the 10 criteria.

Table SC4. Social Competence: *Strong* measure(s) to consider for use in future data collections

Name of measure	Strengths	Limitations
Behavior Assessment System for Children, Second Edition (BASC-2)	1. Strong reliability 2. Strong validity 3. Size and diversity of norming/validation sample 4. Available in multiple languages 5. No training required 6. Parent and teacher forms 7. Covers 2 or more subdomains (social competence, behavior problems, self-regulation, and executive function) 8. Broad age range (preschool form: 2-5 years, child form: 6-11 years)	1. Long administration time (10-20 minutes) 2. Purchase required
Colorado Child Temperament Inventory (CCTI)	1. No training required 2. Parent form 3. Covers 2 or more subdomains (social competence, emotional competence, self-regulation, and executive function) 4. Few items (30) 5. Free 6. Broad age range (1-6 years)	1. Moderate reliability 2. One type of validity reported 3. Normed on non-representative sample (twins from mostly middle to upper class white families) 4. English version only
Devereux Early Childhood Assessment Clinical Form (DECA-C)	1. Strong reliability 2. Strong validity 3. Size and diversity of norming sample 4. No training required 5. Parent and teacher forms 6. Covers 2 or more subdomains (social competence, emotional competence, behavior problems, and self-regulation) 7. Broad age range (2-5 years)	1. English version only 2. Long administration time (15 minutes) 3. Purchase required
Infant Toddler Social Emotional Assessment (ITSEA)	1. Strong reliability 2. Strong validity 3. Normed on large and diverse sample 4. Available in multiple languages 5. No training required 6. Parent and teacher forms 7. Covers 2 or more subdomains (social competence, emotional competence, behavior problems, and self-regulation)	1. Long administration time (25-30 minutes) 2. Purchase required 3. Limited age range (12-36 months)

Name of measure	Strengths	Limitations
Matson Evaluation of Social Skills with Youngsters (MESSY)	1. Strong reliability 2. Size and diversity of norming/validation sample 3. Available in multiple languages 4. No training required 5. Parent and teacher forms 6. Covers 2 or more subdomains (social competence, behavior problems, and self-regulation) 7. Broad age range (2+ years)	1. One type of validity reported 2. Many items (64)
Parent Daily Report (PDR)	1. Available in multiple languages 2. No training required 3. Parent form 4. Covers 2 or more subdomains (social competence, emotional competence, and behavior problems) 5. Short administration time (5-10 minutes) 6. Covers a broad age range (4-10 years)	1. Weak reliability 2. Moderate validity 3. Normed on non-representative sample.

The ITSEA stands out for meeting the criteria for quality of the measure related to reliability, validity and norming sample, while also having several practical strengths. However it has limitations that hinder its usefulness in a large survey, such as length of administration.

Five other measures, in addition, should also be considered as promising because they scored *strong* ratings on five of the 10 criteria.

Table SC5. Social Competence: *Promising* measure(s) to consider for use in future data collections

Name of measure	Strengths	Limitations
Minnesota Preschool Affect Checklist-Revised/Shortened (MPAC-R/S)	1. Strong reliability 2. Strong validity 3. Size and diversity of norming/validation sample 4. Covers 2 or more subdomains (social competence, emotional competence, behavior problems, and self-regulation) 5. No purchase required	1. English version only 2. Requires trained administrator 3. Long administration time (5 minute observations across four different days) 4. Limited age range (3-4 years)

Name of measure	Strengths	Limitations
Penn Interactive Preschool Play Scales (PIPPS)	1. Strong reliability 2. Strong validity 3. No training required 4. Parent and teacher form 5. Covers 2 or more subdomains (social competence, behavior problems)	1. Normed on non-representative (all African American) sample 2. English version only 3. Long administration time (15 minutes) 4. Purchase required 5. Limited age range (preschool-kindergarten)
Rothbart's Early Childhood Behavior Questionnaire (ECBQ)	1. Available in multiple languages 2. No training required 3. Parent form 4. Covers 2 or more subdomains (social competence, emotional competence, behavior problems, self-regulation, and executive function) 5. No purchase required	1. Moderate reliability 2. Moderate validity 3. Normed on a non-representative sample (race) 4. Long administration time (one hour) 5. Limited age range (18-36 months)
Rothbart's Temperament Questionnaires-Infant Behavior Questionnaire Revised (IBQ-R)	1. Available in multiple languages 2. No training required 3. Parent form 4. Covers 2 or more subdomains (social competence, emotional competence, behavior problems, self-regulation, and executive function) 5. No purchase required	1. Moderate reliability 2. Moderate validity 3. Normed on a non-representative sample (race) 4. Long administration time (one hour) 5. Limited age range (3-12 months)
Social Competence and Behavior Evaluation Scale: The Short Form (SCBE-30)	1. Strong reliability 2. Available in multiple languages 3. No training required 4. Covers 2 or more subdomains (social competence and behavior problems) 5. Covers broad age range (3-6 years)	1. One type of validity reported 2. Normed on a non-representative sample 3. Teacher form only 4. Long administration time (10-15 minutes) 5. Purchase required

Summary of Review of Measures of Social Competence

Of the 40 measures reviewed that assess social competence, 19 have been identified for consideration for either a one-time special feature in the *America's Children* report or a federal survey, meeting five or more of the noted criteria. No measures for this subdomain were identified as being used in recurrent cross-sectional federal surveys, which would make them appropriate for inclusion as a regularly appearing indicator. Of the 19 measures that could be considered for one-time special feature in the *America's Children report* and measures that might be considered for regular inclusion in a federal survey, our review of measures finds:

- The SSIS-RS, fielded by ECLS-K:1998 and FACES, meets seven out of 10 criteria and, along with the Social Skills Rating System (SSRS), which also meets seven out of 10 criteria, are the strongest measures to be considered for a one-time special feature in *America's Children*. Social Skills and Positive Approaches to Learning (fielded in FACES) meets six criteria and is also considered appropriate for the same purpose. Five other measures (the ASBI, EDI, NCATS, TBAQ, and Two Bags Task) meet half the criteria and are also potentially appropriate for a one-time special feature in *America's Children*.
- Six measures of social competence meet more than half of the criteria and can be considered for use in new data collections: BASC-2, CCTI, DECA-C, ITSEA, MESSY, and PDR. Five other measures, the MPAC-R/S, PIPPS, ECBQ, IBQ-R, and SCBE-30 meet half of the criteria and are also promising for use in federal surveys.

While the suggested measures do not meet all the desirable criteria that would make a measure worthy of consideration, a number of unmet criteria could be remediated, or at least improved. For example, one of the most common unmet criteria is a shorter administration time (defined as shorter than 10 minutes). Although this limitation affects the BASC-2, DECA-C, EDI, ITSEA, MESSY, MPAC-R/S, CBQ, ECBQ, IBQ-R, PIPPS, SCBE-30, SSIS-RS, SSTRS, and Two Bags Task, abbreviated versions could be developed. We note that the psychometric properties of the shortened versions would still need to be assessed, so underscore that this is not a small task. However, it is one that could be undertaken if a high priority were placed on inclusion of a measure with this limitation.

Additionally, the existence of only an English version is a limitation for the ASBI, CCTI, DECA-C, MPAC-R/S, NCATS, PIPPS, and SSRS that could be rectified with careful work to make the measures linguistically and culturally appropriate for speakers of other languages. This is not an easy task, but again, one that could be undertaken if a priority is placed on further development of one of these measures by a Forum agency.

Nine measures are limited by lack of testing on diverse populations: the ASBI, CCTI, EDI, PDR, PIPPS, ECBQ, IBQ-R, SCBE-30, and TBAQ. This limitation is typically addressed through federal survey development which includes pilot testing on diverse samples, along with psychometric testing of the measures using the pilot and then the full sample.

In addition, limited age range is a limitation for nine suggested measures: ASBI, EDI, ITSEA, MPAC-R/S, PIPPS, ECBQ, IBQ-R, Social Skills and Positive Approaches to Learning, and TBAQ. To rectify this limitation, the measures would need to be adapted and tested for different ages before using them on a wider age group. Extension of a measure to a new age range would take substantial measures development work.

Emotional Competence

Table EC1 below displays the measures within the emotional competence subdomain that might be considered for use in the *America's Children* report and in future federal surveys. Based on their ratings, a total of 15 measures (out of 35 reviewed) are suggested for consideration for these purposes. Additional information on each measure's strengths and limitations is outlined in the descriptions following the summary chart.

Table EC1. Selected measures of emotional competence that meet at least half of the evaluation criteria

Name of Measure	Reliability	Validity	Size and diversity of norming/ validation sample	Availability in languages other than English	Requires trained administrator	Parent/teacher form	Covers 2 or more subdomains	Length of time to administer	Cost / purchase requirement	Covers a wide age range
Adaptive Social Behavior Inventory (ASBI)	Moderate	Strong	Moderate	Moderate	Strong	Strong	Strong	Strong	No info	Moderate
Brief Scale of Temperament (BST)	Moderate	Moderate	Weak	Moderate	Strong	Strong	Strong	Strong	Strong	Strong
Colorado Childhood Temperament Inventory (CCTI)	Moderate	Moderate	Weak	Moderat	Strong	Strong	Strong	Strong	Strong	Strong
Devereux Early Childhood Assessment Clinical Form (DECA-C)	Strong	Strong	Storng	Moderate	Strong	Strong	Strong	Moderate	Moderate	Strong
Early Development Instrument (EDI)	Strong	Moderate	Moderate	Strong	Strong	Moderate	Strong	Moderate	Strong	Moderate
Infant Toddler Social Emotional Assessment (ITSEA)	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Weak	Moderate	Moderate
Minnesota Preschool Affect Checklist-Revised/Shortened (MPAC-R/S)	Strong	Strong	Strong	Moderate	Weak	N/A	Strong	Weak	Strong	Moderate
Nursing Child Assessment Teaching Scale (NCATS) - Child Scale	Strong	Moderate	Strong	Moderate	Weak	N/A	Strong	Strong	No info	Strong

Name of Measure	Reliability	Validity	Size and diversity of norming/ validation sample	Availability in languages other than English	Requires trained administrator	Parent/teacher form	Covers 2 or more subdomains	Length of time to administer	Cost / purchase requirement	Covers a wide age range
Parent Daily Report (PDR)	Weak	Moderate	Moderate	Strong	Strong	Strong	Strong	Strong	No info	Strong
Rothbart's Children's Behavior Questionnaire (CBQ)	Moderate	Strong	Strong	Strong	Strong	Strong	Strong	Weak	Strong	Strong
Rothbart's Children's Behavior Questionnaire Very Short Form (CBQ-VSF)[18]	Moderate	Strong	Moderate	Strong	Strong	Strong	Strong	Moderate	Strong	Strong
Rothbart's Early Childhood Behavior Questionnaire (ECBQ)[19]	Moderate	Moderate	Moderate	Strong	Strong	Strong	Strong	Weak	Strong	Moderate
Rothbart's Infant Behavior Questionnaire Revised (IBQ-R)[20]	Moderate	Moderate	Moderate	Strong	Strong	Strong	Strong	Weak	Strong	Moderate
Social Skills Rating System (SSRS)[21]	Strong	Strong	Strong	Moderate	Strong	Strong	Strong	No info; 58 items	Moderate	Strong
Toddler Behavior Assessment Questionnaire (TBAQ)	Moderate	Strong	Moderate	Strong	Strong	Strong	Strong	No info; 108 items	No info	Moderate

---

[18] Due to special interest in short assessments for the pre-kindergarten and kindergarten period, the Very Short Form of the CBQ was reviewed in addition to the full version of the CBQ.

[19] The ECBQ also offers a Short Form (107 items) and a Very Short Form (36 items). Information about the duration needed to administer the short forms is not available. The rating in this chart is for the full 201 item measure.

[20] The IBQ-R also offers a Short form (91 items) and a Very Short form (36 items). Information about the duration needed to administer the short forms is not available. The rating in this chart is for the full 184 item measure.

[21] The reliability and validity information collected for the SSRS includes some information from the ECLS-K and ECLS-K: 2011 studies (which use selected and adapted items from the SSRS) as well as from documentation for the full measure.

Measures Appropriate for Consideration for Use as Regular Indicator

    None of the measures of emotional competence are currently appropriate for consideration for use as a regular indicator in *America's Children* since they are not included in federal surveys that occur regularly.

Measures Appropriate for Consideration for Use as One-time Indicator

    Twelve measures included in the full Inventory assess emotional competence and have been used in federally sponsored longitudinal studies. Among these, three measures met more than half of the criteria: the Rothbart's Children's Behavior Questionnaire (CBQ), the CBQ-VSF, and the Social Skills Rating Scales (SSRS). The CBQ was fielded in the ECLS-K: 2011 and scored a *strong* rating on 8 of the 10 criteria. As an extension, we include the CBQ-VSF in our recommendation of measures to consider as a one-time indicator because it is a strong measure that could be considered as a much shorter alternative to the full version of the CBQ (36 items as opposed to 195), resulting in three strong measures to consider for use as a one-time indicator. The SSRS scored a *strong* rating on 7 of the 10 criteria.

Table EC2. Emotional competence: *Strong* measure(s) to consider for use as a one-time indicator

Name of measure	Strengths	Limitations
Rothbart's Children's Behavior Questionnaire (CBQ), fielded in the ECLS-K: 2011	1. Strong validity 2. Representative norming sample 3. Available in multiple languages 4. No training required 5. Parent form 6. Covers 2 or more subdomains (emotional competence, behavior problems, self-regulation, and executive function) 7. No purchase required 8. Broad age range (3-7 years)	1. Moderate reliability 2. Long administration time (1 hour)
Rothbart's Children's Behavior Questionnaire-Very Short Form (CBQ-VSF)	1. Strong validity 2. Available in multiple languages 3. No training required 4. Parent and teacher forms 5. Covers 2 or more subdomains (emotional competence, behavior problems, self-regulation, and executive function) 6. No purchase required 7. Broad age range (3-7 years)	1. Moderate reliability 2. Moderate time of administration (< 15 minutes) 3. Small norming samples

Name of measure	Strengths	Limitations
Social Skills Rating Scales (SSRS), selected and adapted items fielded by ECLS-K:1998 and FACES	1. Strong reliability 2. Strong validity 3. Size and diversity of norming/validation sample 4. No training required 5. Parent and teacher forms 6. Covers 2 or more subdomains (social competence, emotional competence, and behavior problems) 7. Broad age range (3-18 years)	1. English version only 2. Many items (58) 3. Purchase required[22]

Four measures are considered promising, meeting five of the 10 criteria: the Adaptive Social Behavior Inventory (ASBI), which was used in the NICHD Study of Early Child Care and Youth Development; the Early Development Instrument (EDI), which was used in the National Longitudinal Survey of Children and Youth; the Nursing Child Assessment Teaching Scale (NCATS)-Child Scale, which was used in the Early Childhood Longitudinal Study Birth Cohort (ECLS-B) at nine months (2001-2002); and the Toddler Behavior Assessment Questionnaire (TBAQ), which was used in the NICHD Study of Early Child Care and Youth Development in 1994.

Table EC3. Emotional Competence: *Promising* measure(s) to consider for use as a one-time indicator

Name of measure	Strengths	Limitations
Adaptive Social Behavior Inventory (ASBI), fielded in the NICHD Study of Early Child Care and Youth Development	1. Strong validity 2. No training required 3. Parent and teacher forms 4. Covers 2 or more domains (social competence, emotional competence, behavior problems, and self-regulation) 5. Few items (30 items)	1. Moderate reliability 2. Normed on non-representative sample 3. English version only 4. Limited age range ("young high-risk children")
Early Development Instrument (EDI), fielded in the National Longitudinal Survey of Children and Youth	1. Strong reliability 2. Available in multiple languages 3. No training required 4. Covers 2 or more subdomains (social competence and emotional competence) 5. No purchase required	1. Moderate validity 2. Normed on non-representative sample 3. Teacher form only 4. Moderate administration time (20 minutes) 5. Limited age range (4-6 years)

---

[22] Since this measure is currently used in an existing federal study, purchase will not be required for using this measure as the basis of an indicator.

Nursing Child Assessment Teaching Scale (NCATS) - Child Scale, fielded in the ECLS-B in 2007	1. Strong reliability 2. Size and diversity of norming/validation sample (ECLS-B) 3. Covers 2 or more subdomains (social competence and emotional competence) 4. Short administration time (1-6 minutes) 5. Broad age range (0-3 years)	1. Moderate validity 2. English version only 3. Requires trained administrator
Toddler Assessment Behavior Questionnaire (TBAQ), fielded in the NICHD Study of Early Child Care and Youth Development in 1994	1. Strong validity 2. Covers 2 or more subdomains (behavior problems and self-regulation) 3. Available in multiple languages 4. Has parent report 5. No training required	1. Moderate reliability 2. Validated on non-diverse sample 3. Limited age range (16-36 months) 4. Long administration time (108 items)

<u>Measures to Consider for First Use in Federal Surveys</u>

Eight measures included in the full inventory assess emotional competence and could be considered for new federal data collections (i.e., they are not currently being used in a federal data collection). Of those, the Brief Scale of Temperament (BST), Colorado Childhood Temperament Inventory (CCTI), Devereux Early Childhood Assessment Clinical Form (DECA-C), Infant Toddler Social Emotional Assessment (ITSEA), and Parent Daily Report (PDR), are rated as *strong* on more than half of the 10 criteria.

Table EC4. Emotional Competence: *Strong* measure(s) to consider for use in future data collections

Name of measure	Strengths	Limitations
Brief Scale of Temperament (BST)	1. No training required 2. Parent form 3. Covers 2 or more subdomains (emotional competence and self-regulation) 4. Short administration time (< 5 minutes) 5. No purchase required 6. Broad age range (4-7 years)	1. One type of reliability reported 2. One type of validity reported 3. Normed on small non-representative sample 4. English version only

Name of measure	Strengths	Limitations
Colorado Child Temperament Inventory (CCTI)	1. No training required 2. Parent form 3. Covers 2 or more subdomains (social competence, emotional competence, self-regulation, and executive function) 4. Few items (30) 5. No purchase required 6. Broad age range (1-6 years)	1. Moderate reliability 2. One type of validity reported 3. Normed on non-representative sample (twins from mostly middle to upper class white families) 4. English version only
Devereux Early Childhood Assessment Clinical Form (DECA-C)	1. Strong reliability 2. Strong validity 3. Size and diversity of norming sample 4. No training required 5. Parent and teacher forms 6. Covers 2 or more subdomains (emotional competence, emotional competence, behavior problems, and self-regulation) 7. Broad age range (2-5 years)	1. English version only 2. Long administration time (15 minutes) 3. Purchase required
Infant Toddler Social Emotional Assessment (ITSEA)	1. Strong reliability 2. Strong validity 3. Normed on large and diverse sample 4. Available in multiple languages 5. No training required 6. Parent form 7. Covers 2 or more subdomains (social competence, behavior problems, and self-regulation)	1. Long administration time (25-30 minutes) 2. Purchase required 3. Limited age range (12-36 months)
Parent Daily Report (PDR)	1. Available in multiple languages 2. No training required 3. Parent form 4. Covers 2 or more subdomains (social competence, emotional competence, and behavior problems) 5. Short administration time (5-10 minutes) 6. Broad age range (4-10 years)	1. Weak reliability 2. Moderate validity 3. Normed on non-diverse sample

In addition, three other measures should be considered as promising because they receive a *strong* rating on five of the 10 criteria[23]:

Table EC5. Emotional Competence: *Promising* measures to consider for use in future data collections

Name of measure	Strengths	Limitations
Minnesota Preschool Affect Checklist-Revised/Shortened (MPAC-R/S)	1. Strong reliability 2. Strong validity 3. Diverse norming sample 4. Covers 2 or more subdomains (social competence, emotional competence, behavior problems, and self-regulation) 5. No purchase required	1. English version only 2. Requires a trained administrator 3. Long administration (5 minute observations across four different days)
Rothbart's Early Childhood Behavior Questionnaire (ECBQ)	1. Available in multiple languages 2. No training required 3. Parent form 4. Covers 2 or more subdomains (social competence, emotional competence, behavior problems, self-regulation, and executive function) 5. No purchase required	1. Moderate reliability 2. Moderate validity 3. Normed on a non-representative sample (race) 4. Long administration (1 hour) 5. Limited age range (18-36 months)
Rothbart's Infant Behavior Questionnaire Revised (IBQ-R)	1. Available in multiple languages 2. No training required 3. Parent form 4. Covers 2 or more subdomains (social competence, emotional competence, behavior problems, self-regulation, and executive function)	1. Moderate reliability 2. Moderate validity 3. Normed on a non-representative sample 4. Long administration (1 hour) 5. Limited age range (3-12 months)

Summary of Review of Measures of Emotional Competence

Of the 36 measures reviewed, none of them are currently included in regularly occurring federal data collection efforts and so none could be considered for inclusion in the *America's Children* report as a regularly occurring indicator. Seven measures – the ASBI, EDI, NCATS, CBQ, CBQ-VSF, SSRS, and TBAQ – could be considered for a one-time feature in *America's Children*. Eight measures could be considered for new data collection in federal surveys. Among these, the BST, CCTI, DECA-C, ITSEA, and PDR are the strongest. Three additional measures are also considered promising for first use in federal surveys, meeting half of the 10 evaluated criteria: the MPAC-R/S, ECBQ, and IBQ-R.

---

[23] In a previous version of this paper, we considered the Denham Affect Knowledge Test (AKT; Denham, 1986) to be a promising measure for use in future federal data collections. Although a new analysis of the measure disqualified it from recommendation in this paper, we consider it a promising measure for the direct assessment of emotion knowledge. See the full Inventory for additional information.

A number of noted limitations could be rectified, or at least improved, though substantial further work would be needed to do so. For example, one of the most common unmet criteria is a shorter administration time (defined as shorter than 10 minutes). Although this limitation affects the DECA-C, EDI, ITSEA, all four Rothbart temperament questionnaires (CBQ, CBQ-VSF, ECBQ, and IBQ-R), and MPAC-R/S, these measures could be shortened to take less than 10 minutes while still preserving their other qualities. Note, however, that all of the measures that require substantial time for administration cover multiple subdomains. This suggests that a longer time for administration may be acceptable given that the measure assesses multiple domains of development, therefore eliminating the need for other measures that might be used concurrently. As above, we note that creating a shortened version of a measure is not a small task, and one that would require further psychometric evaluation for the shortened version.

Seven of the measures, the ASBI, BST, CCTI, DECA-C, MPAC-R/S, NCATS, and SSRS only have English versions. This limitation could be rectified with careful work on making the measure linguistically and culturally appropriate in other languages. Again, we note that this is not a trivial further step.

Another limitation for the ASBI, BST, CCTI, EDI, PDR, CBQ-VSF, ECBQ, IBQ-R, and TBAQ is a lack of testing on large and/or diverse populations. This is typically addressed through federal survey development which includes pilot testing on diverse samples, along with psychometric testing of the measures using the pilot and then the full sample data.

The ASBI, EDI, ITSEA, MPAC-R/S, ECBQ, IBQ-R, and TBAQ are limited by a narrow age range and would require adaptation and testing for use with different age groups. As noted in the social competence section, this is a substantial undertaking that would require careful consideration of how the construct is appropriately addressed for children of different ages. Do consider however that the suite of Rothbart temperament questionnaires (including the ECBQ and IBQ-R) may be used to cover a broader age span.

In addition, the MPAC-R/s and NCATS require a trained administrator, which would necessitate additional time, money, and resources spent before data collection.

Behavior Problems

Table BP1 below displays the measures within the behavior problems subdomain that are appropriate for further consideration for use in *America's Children* and federal surveys. Based on their ratings, a total of 17 measures (out of 32 reviewed) are suggested for consideration for these purposes. Additional information on each measure's strengths and limitations is outlined in the descriptions following the summary chart.

Table BP1. Selected measures of behavior problems that meet at least half of the evaluation criteria

Name of Measure	Reliability	Validity	Size and diversity of norming/ validation sample	Availability in languages other than English	Requires trained administrator	Parent/teacher form	Covers 2 or more subdomains	Length of time to administer	Cost / purchase requirement	Covers a wide age range
Adaptive Social Behavior Inventory (ASBI)	Moderate	Strong	Moderate	Moderate	Strong	Strong	Strong	Strong	No info	Moderate
Behavior Assessment System for Children, Second Edition (BASC-2)	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Moderate	Moderate	Moderate
Devereux Early Childhood Assessment Clinical Form (DECA-C)	Strong	Strong	Storng	Moderate	Strong	Strong	Strong	Moderate	Moderate	Strong
Infant Toddler Social Emotional Assessment (ITSEA)	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Weak	Moderate	Moderate
Matson Evaluation of Social Skills with Youngsters (MESSY)	Strong	Moderate	Strong	Strong	Strong	Strong	Strong	No info; 64 items	No info	Strong
Minnesota Preschool Affect Checklist-Revised/Shortened (MPAC-R/S)	Strong	Strong	Strong	Moderate	Weak	N/A	Strong	Weak	Strong	Moderate
Parent Daily Report (PDR)	Weak	Moderate	Moderate	Strong	Strong	Strong	Strong	Strong	No info	Strong
Penn Interactive Preschool Play Scales (PIPPS)	Strong	Strong	Moderate	Moderate	Strong	Strong	Strong	Moderate	Moderate	Moderate

Name of Measure	Reliability	Validity	Size and diversity of norming/ validation sample	Availability in languages other than English	Requires trained administrator	Parent/teacher form	Covers 2 or more subdomains	Length of time to administer	Cost / purchase requirement	Covers a wide age range
Rothbart's Children's Behavior Questionnaire (CBQ)	Moderate	Strong	Strong	Strong	Strong	Strong	Strong	Weak	Strong	Strong
Rothbart's Children's Behavior Questionnaire Very Short Form (CBQ-VSF)	Moderate	Strong	Moderate	Strong	Strong	Strong	Strong	Moderate	Strong	Strong
Rothbart's Early Childhood Behavior Questionnaire (ECBQ)[24]	Moderate	Moderate	Moderate	Strong	Strong	Strong	Strong	Weak	Strong	Moderate
Rothbart's Infant Behavior Questionnaire Revised (IBQ-R)[25]	Moderate	Moderate	Moderate	Strong	Strong	Strong	Strong	Weak	Strong	Moderate
Social Competence and Behavior Evaluation Scale: The Short Form (SCBE-30)	Strong	Moderate	Moderate	Strong	Strong	Moderate	Strong	Moderate	Moderate	strong
Social Skills Improvement System-Rating Scales (SSIS-RS)	Strong	Moderate	Strong	Strong	Strong	Strong	Strong	Moderate	Moderate	Stromg
Social Skills Rating System (SSRS)[26]	Strong	Strong	Strong	Moderate	Strong	Strong	Strong	No info; 58 items	Moderate	Strong
Toddler Behavior Assessment Questionnaire (TBAQ)	Moderate	Strong	Moderate	Strong	Strong	Strong	Strong	No info; 108 items	No info	Moderate
Two Bags Task	Moderate	Weak	Strong	Strong	Weak	N/A	Strong	Moderate	Strong	Strong

[24] The ECBQ also offers a Short Form (107 items) and a Very Short Form (36 items). Information about the duration needed to administer the short forms is not available. The rating in this chart is for the full 201 item measure.

[25] The IBQ-R also offers a Short form (91 items) and a Very Short form (36 items). Information about the duration needed to administer the short forms is not available. The rating in this chart is for the full 184 item measure.

[26] The reliability and validity information collected for the SSRS includes some information from the ECLS-K and ECLS-K: 2011 studies (which use selected and adapted items from the SSRS) as well as from documentation for the full measure.

Measures Appropriate for Consideration for Use as Regular Indicator

None of the measures of behavior problems in the full Inventory are currently used by regularly occurring cross-sectional federal surveys.

Measures Appropriate for Consideration for Use as One-time Indicator

Sixteen measures included in the full Inventory assess behavior problems and have been used in federal longitudinal studies. Of those, five score a *strong* rating in more than five of the 10 criteria.

Table BP2. Behavior Problems: *Strong* measure(s) to consider for use as a one-time indicator

Name of measure	Strengths	Limitations
Rothbart's Children's Behavior Questionnaire (CBQ), fielded in the ECLS-K: 2011	1. Strong validity 2. Representative norming sample 3. Available in multiple languages 4. No training required 5. Parent form 6. Covers 2 or more subdomains (emotional competence, behavior problems, self-regulation, and executive function) 7. No purchase required 8. Broad age range (3-7 years)	1. Moderate reliability 2. Long administration time (1 hour)
Rothbart's Children's Behavior Questionnaire-Very Short Form (CBQ-VSF)	1. Strong validity 2. Available in multiple languages 3. No training required 4. Parent and teacher forms 5. Covers 2 or more subdomains (emotional competence, behavior problems, self-regulation, and executive function) 6. No purchase required 7. Broad age range (3-7 years)	1. Moderate reliability 2. Moderate administration time (< 15 minutes) 3. Small norming samples
Social Skills Improvement System-Rating Scales (SSIS-RS), fielded in ECLS-K: 2011	1. Strong reliability 2. Size and diversity of norming/validation sample 3. Available in multiple languages 4. No training required 5. Parent and teacher forms 6. Covers 2 or more subdomains (social competence and behavior problems) 7. Broad age range (3-18 years)	1. Moderate validity 2. Moderate administration time (10-25 minutes) 3. Purchase required

Name of measure	Strengths	Limitations
Social Skills Rating Scales (SSRS), selected and adapted items fielded by ECLS-K:1998 and FACES	1. Strong reliability 2. Strong validity 3. Size and diversity of norming/validation sample 4. No training required 5. Parent and teacher forms 6. Covers 2 or more subdomains (social competence, emotional competence and behavior problems) 7. Broad age range (3-18 years)	1. English version only 2. Many items (58) 3. Purchase required[27]

Three other measures should be considered as promising because they scored a *strong* rating on half of the 10 criteria.

Table BP3. Behavior Problems: *Promising* measure(s) to consider for use as a one-time indicator

Name of measure	Strengths	Limitations
Adaptive Social Behavior Inventory (ASBI), fielded in the NICHD Study of Early Child Care and Youth Development	1. Strong validity 2. No training required 3. Parent and teacher forms 4. Covers 2 or more subdomains (social competence, emotional competence, behavior problems, and self-regulation) 5. Few items (30 items)	1. Moderate reliability 2. Non-diverse norming/validation sample 3. English version only 4. Limited age range ("young high-risk children")
Toddler Assessment Behavior Questionnaire (TBAQ), fielded in the NICHD Study of Early Child Care and Youth Development in 1994	1. Strong validity 2. Available in multiple languages 3. No training required 4. Parent form 5. Covers 2 or more subdomains (social competence, emotional competence, behavior problems and self-regulation)	1. Moderate reliability 2. Validated on non-diverse sample 3. Limited age range (16-36 months) 4. Long administration time (108 items)
Two Bags Task	1. Size and diversity of norming sample 2. Available in multiple languages 3. Covers 2 or more subdomains (social competence, behavior problems, and self-regulation) 4. No purchase required 5. Broad age range (2 years-preschool)	1. Moderate reliability 2. Weak validity 3. Trained administrator required 4. Moderate administration time administer (10 minutes for task then coded later)

---

[27] Since this measure is currently used in an existing federal study, purchase will not be required for using this measure as the basis of an indicator.

Ten measures included in the full Inventory assess behavior problems and could be considered for new federal data collections. Of those, five have more strengths than limitations, with *strong* ratings on more than half of the 10 criteria.

Table BP4. Behavior problems: *Strong* measure(s) to consider for use in future data collections

Name of measure	Strengths	Limitations
Behavior Assessment System for Children, Second Edition (BASC-2)	1. Strong reliability 2. Strong validity 3. Size and diversity of norming/validation sample 4. Available in multiple languages 5. No training required 6. Parent and teacher forms 7. Covers 2 or more subdomains (social competence, behavior problems, self-regulation, and executive function) 8. Broad age range (preschool form: 2-5 years, child form: 6-11 years)	1. Long administration time (10-20 minutes) 2. Purchase required
Devereux Early Childhood Assessment Clinical Form (DECA-C)	1. Strong reliability 2. Strong validity 3. Size and diversity of norming sample 4. No training required 5. Parent and teacher forms 6. Covers 2 or more subdomains (social competence, emotional competence, behavior problems, and self-regulation) 7. Broad age range (2-5 years)	1. English version only 2. Long administration time (15 minutes) 3. Purchase required
Infant Toddler Social Emotional Assessment (ITSEA)	1. Strong reliability 2. Strong validity 3. Normed on large and diverse sample 4. Available in multiple languages 5. No training required 6. Parent and teacher forms 7. Covers 2 or more subdomains (social competence, behavior problems, and self-regulation)	1. Long administration time (25-30 minutes) 2. Purchase required 3. Limited age range (12-36 months)

Name of measure	Strengths	Limitations
Matson Evaluation of Social Skills with Youngsters (MESSY)	1. Strong reliability 2. Size and diversity of norming/validation sample 3. Available in multiple languages 4. No training required 5. Parent and teacher forms 6. Covers 2 or more subdomains (social competence, behavior problems, and self-regulation) 7. Broad age range (2-18 years)	1. One type of validity reported 2. Many items (64)
Parent Daily Report (PDR)	1. Available in multiple languages 2. No training required 3. Parent from 4. Covers 2 or more subdomains (social competence, emotional competence, and behavior problems) 5. Short administration time (5-1-minutes) 6. Broad age range (4-10 years)	1. Weak reliability 2. Moderate validity 3. Non-diverse norming/validation sample

In addition, five other measures should be considered as promising because they score a *strong* rating on half of the 10 criteria.

Table BP5. Behavior Problems: *Promising* measure(s) to consider for use in future data collections

Name of measure	Strengths	Limitations
Minnesota Preschool Affect Checklist-Revised/Shortened (MPAC-R/S)	1. Strong reliability 2. Strong validity 3. Size and diversity of norming/validation sample 4. Covers 2 or more subdomains (social competence, emotional competence, behavior problems, and self-regulation) 5. No purchase required	1. English version only 2. Requires a trained administrator 3. Long administration time (5 minute observations across 4 different days) 4. Narrow age range (3-4 years)
Penn Interactive Preschool Play Scales (PIPPS)	1. Strong reliability 2. Strong validity 3. No training required 4. Parent and teacher form 5. Covers 2 or more subdomains (social competence and behavior problems)	1. Normed on non-representative (all African American) sample 2. English version only 3. Long administration time (15 minutes) 4. Purchase required 5. Limited age range (preschool and kindergarten)

Name of measure	Strengths	Limitations
Rothbart's Early Childhood Behavior Questionnaire (ECBQ)	1. Available in multiple languages 2. No training required 3. Parent form 4. Covers 2 or more subdomains (social competence, emotional competence, behavior problems, self-regulation, and executive function) 5. No purchase required	1. Moderate reliability 2. Moderate validity 3. Normed on a non-representative sample (race) 4. Long administration time (1 hour) 6. Limited age range (18-36 months)
Rothbart's Infant Behavior Questionnaire Revised (IBQ-R)	1. Available in multiple languages 2. No training required 3. Parent form 4. Covers 2 or more subdomains (social competence, emotional competence, behavior problems, self-regulation, and executive function) 5. No purchase required	1. Moderate reliability 2. Moderate validity 3. Non-diverse norming/validation sample 4. Long administration time (1 hour) 5. Limited age range (3-12 months)
Social Competence and Behavior Evaluation Scale: The Short Form (SCBE-30)	1. Strong reliability 2. Available in multiple languages 3. No training required 4. Covers 2 or more subdomains (social competence and behavior problems) 5. Broad age range (3-6 years)	1. One type of validity reported 2. Normed on a non-representative sample 3. Teacher form only 4. Long administration time (10-15 minutes) 5. Requires purchase

Summary of Review of Measures of Behavioral Problems

Out of the 32 measures reviewed, five measures, CBQ, CBQ-VSF, SSIS-RS, and SSRS, meet more than half the criteria for *strong* and could be considered for one-time use in a special feature of *America's Children* given their strengths. The ASBI, TBAQ, and Two Bags Task meet just half of the criteria and are considered promising for this purpose. Five measures are considered to be strong candidates for first-time use in federal surveys: the BASC-2, DECA-C, ITSEA, MESSY, and PDR. In addition, the MPAC-R/S, PIPPS, ECBQ, IBQ-R, and SCBE-30 are considered promising, but not as strong as the others, for this purpose.

A number of the noted unmet criteria could be rectified, or at least improved, albeit with substantial effort. With the exception of ASBI and PDR, all other suggested measures have a limitation of long administration time (defined as longer than 10 minutes). However, it should be noted that most cover multiple domains of children's development, which may eliminate the need for multiple concurrent measures. Though this would require substantial work, a number of the measures could be shortened to take less than 10 minutes while still preserving their other qualities.

In addition, five of the measures (the ASBI, DECA-C, MPAC-R/S, PIPPS, and SSRS) only have English versions. This limitation could be rectified with careful measures development work on creating linguistically and culturally appropriate versions in other languages. Again, we caution that while this is

feasible, the effort required for development of a strong version of a measure in another language should not be underestimated.

Another unmet criterion for the ASBI, PDR, PIPPS, CBQ-VSF, ECBQ, IBQ-R, SCBE-30, and TBAQ is development and testing of the measure with large, representative samples. This limitation is typically addressed through federal survey development which includes pilot testing on diverse samples, along with psychometric testing using the pilot and then the full sample data.

Finally, limited coverage across the early childhood age range is a limitation that needs to be addressed for seven of the indicated promising measures: ASBI, ITSEA, MPAC-R/S, PIPPS, ECBQ, IBQ-R, and TBAQ. As mentioned earlier, the various Rothbart temperament questionnaires (including the ECBQ and IBQ-R) may be used to cover a broader age span. Also as suggested in previous sections, age ranges could be broadened by adapting and testing the measures with different age groups. However, this takes substantial work, starting with conceptualization of how a construct is manifested at different ages.

Self-Regulation

Table SR1 below displays measures within the self-regulation subdomain that might be given consideration for use in the *America's Children* report and in future federal surveys. Based on their ratings, a total of 19 measures (out of 31 reviewed) are suggested for consideration for these purposes. Additional information on each measure's strengths and limitations is outlined in the descriptions following the summary chart.

Table SR1. Selected measures of self-regulation that meet at least half of the evaluation criteria

Name of Measure	Reliability	Validity	Size and diversity of norming/ validation sample	Availability in languages other than English	Requires trained administrator	Parent/teacher form	Covers 2 or more subdomains	Length of time to administer	Cost / purchase requirement	Covers a wide age range
Adaptive Social Behavior Inventory (ASBI)	Moderate	Strong	Moderate	Moderate	Strong	Strong	Strong	Strong	No info	Moderate
Approaches to Learning Scale from the ECLS-K [28]	Moderate	Moderate	Strong	Strong	Strong	Strong	Moderate	Strong	Strong	Strong
Approaches to Learning Scale – shortened version tested by the FACES team [28]	Moderate	Moderate	Strong	Strong	Strong	Moderate	Moderate	Strong	Strong	Moderate
Behavior Assessment System for Children, Second Edition (BASC-2)	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Moderate	Moderate	Strong
Behavior Rating Inventory of Executive Function-Preschool (BRIEF-P)	Strong	Moderate	Strong	Strong	Strong	Strong	Moderate	Moderate	Moderate	Strong
Brief Scale of Temperament (BST)	Moderate	Moderate	Weak	Moderate	Strong	Strong	Strong	Strong	Strong	Strong
Childhood Executive Function Inventory (CHEXI)	Moderate	Moderate	Weak	Strong	Strong	Strong	Moderate	Strong	Strong	Strong

---

[28] For the purpose of this review, the Approaches to Learning Scale is considered a measure of self-regulation; however, while the scale as a whole is a report of how often children exhibit a selected set of learning behaviors, items also capture aspects of the behavior problems subdomain as well as executive function.

Name of Measure	Reliability	Validity	Size and diversity of norming/ validation sample	Availability in languages other than English	Requires trained administrator	Parent/teacher form	Covers 2 or more subdomains	Length of time to administer	Cost / purchase requirement	Covers a wide age range
Colorado Childhood Temperament Inventory (CCTI)	Moderate	Moderate	Weak	Moderate	Strong	Strong	Strong	Strong	Strong	Strong
Devereux Early Childhood Assessment Clinical Form (DECA-C)	Strong	Strong	Storng	Moderate	Strong	Strong	Strong	Moderate	Moderate	Strong
Infant Toddler Social Emotional Assessment (ITSEA)	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Weak	Moderate	Moderate
Matson Evaluation of Social Skills with Youngsters (MESSY)	Strong	Moderate	Strong	Strong	Strong	Strong	Strong	No info; 64 items	No info	Strong
Preschool Learning Behaviors Scale (PLBS)	Strong	Strong	Strong	Strong	Strong	Moderate	Moderate	Strong	Moderate	Moderate
Rothbart's Children's Behavior Questionnaire (CBQ)	Moderate	Strong	Strong	Strong	Strong	Strong	Strong	Weak	Strong	Strong
Rothbart's Children's Behavior Questionnaire Very Short Form (CBQ-VSF)	Moderate	Strong	Moderate	Strong	Strong	Strong	Strong	Moderate	Strong	Strong
Rothbart's Early Childhood Behavior Questionnaire (ECBQ)[29]	Moderate	Moderate	Moderate	Strong	Strong	Strong	Strong	Weak	Strong	Moderate
Rothbart's Infant Behavior Questionnaire Revised (IBQ-R)[30]	Moderate	Moderate	Moderate	Strong	Strong	Strong	Strong	Weak	Strong	Moderate

[29] The ECBQ also offers a Short Form (107 items) and a Very Short Form (36 items). Information about the duration needed to administer the short forms is not available. The rating in this chart is for the full 201 item measure.

[30] The IBQ-R also offers a Short form (91 items) and a Very Short form (36 items). Information about the duration needed to administer the short forms is not available. The rating in this chart is for the full 184 item measure.

Name of Measure	Reliability	Validity	Size and diversity of norming/ validation sample	Availability in languages other than English	Requires trained administrator	Parent/teacher form	Covers 2 or more subdomains	Length of time to administer	Cost / purchase requirement	Covers a wide age range
Toddler Behavior Assessment Questionnaire (TBAQ)	Moderate	Strong	Moderate	Strong	Strong	Strong	Strong	No info; 108 items	No info	Moderate
Two Bags Task	Moderate	Weak	Strong	Strong	Weak	N/A	Strong	Moderate	Strong	Strong

<u>Measures Appropriate for Consideration for Use as Regular Indicator</u>

No measures of self-regulation included in the full Inventory are currently used by regularly occurring cross-sectional federal surveys.

<u>Measures Appropriate for Consideration for Use as One-time Indicator</u>

Thirteen measures included in the full Inventory assess self-regulation and have been used in longitudinal studies. Four of these measures score a *strong* rating on six or more of the 10 criteria.

Table SR2. Self-regulation: *Strong* measure(s) to consider for use as a one-time indicator

Name of measure	Strengths	Limitations
Approaches to Learning Scale, fielded in the ECLS-K: 1998 and ECLS-K: 2011	1. Normed on large and diverse sample 2. Available in multiple languages 3. No training required 4. Parent and teacher forms 5. Short administration time (5 minutes) 6. No purchase required 7. Broad age range (K-5th grade)	1. Moderate reliability 2. Strong validity on only one type of validity 3. Self-regulation subdomain only
Preschool Learning Behaviors Scale (PLBS), fielded in FACES in 2006	1. Strong reliability 2. Strong validity 3. Normed on large and diverse sample 4. Available in multiple languages 5. No training required 6. Short administration time (8 minutes)	1. Only teacher form available 2. Purchase required 3. Limited age range (3-5 years) 4. Only covers 1 subdomain (and EF)
Rothbart's Children's Behavior Questionnaire (CBQ), fielded in the ECLS-K: 2011	1. Strong validity 2. Representative norming sample 3. Available in multiple languages 4. No training required 5. Parent form 6. Covers 2 or more subdomains (emotional competence, behavior problems, self-regulation, and executive function) 7. No purchase required 7. Broad age range (3-7 years)	1. Moderate reliability 2. Long administration time (1 hour)

Name of measure	Strengths	Limitations
Rothbart's Children's Behavior Questionnaire-Very Short Form (CBQ-VSF)	1. Strong validity 2. Available in multiple languages 3. No training required 4. Parent and teacher forms 5. Covers 2 or more subdomains (emotional competence, behavior problems, self-regulation, and executive function) 6. No purchase required 7. Broad age range (3-7 years)	1. Moderate reliability 2. Moderate time of administration (< 15 minutes) 3. Small norming samples

Four other measures should be considered as promising because they score a *strong* rating on half of the 10 criteria.

Table SR3. Self-Regulation: *Promising* measure(s) to consider for use as a one-time indicator

Name of measure	Strengths	Limitations
Adaptive Social Behavior Inventory (ASBI), fielding in the NICHD Study of Early Child Care and Youth Development	1. Strong validity 2. No training required 3. Parent and teacher forms 4. Covers 2 or more subdomains (social competence, emotional competence, behavior problems, and self-regulation) 5. Few items (30 items)	1. Moderate reliability 2. Non-diverse norming/validation sample 3. Limited age range ("young high-risk children"
Approaches to Learning-shortened version tested by the FACES team	1. Normed on large and diverse sample 2. Available in multiple languages 3. No training required 4. Short administration time (5 minutes) 5. No purchase required	1. Moderate reliability 2. Moderate validity 3. Only teacher form available 4. Self-regulation subdomain only 5. Limited age range (3-5 years)
Toddler Assessment Behavior Questionnaire (TBAQ), fielded in the NICHD Study of Early Child Care and Youth Development	1. Strong validity 2. Covers 2 or more subdomains (social competence, emotional competence, behavior problems and self-regulation) 3. Available in multiple languages 4. Has parent report 5. No training required	1. Moderate reliability 2. Validated on non-diverse sample 3. Limited age range (16-36 months) 4. Long administration time (108 items)

Name of measure	Strengths	Limitations
Two Bags Task, fielded in the ECLS-B	1. Size and diversity of norming/validation sample 2. Available in multiple languages 3. Covers 2 or more subdomains (social competence, behavior problems, and self-regulation) 4. No purchase required 5. Broad age range (2 years-preschool)	1. Moderate reliability 2. Weak validity 3. Trained administrator required 4. Long administration time (10 minutes for task and then coded later)

Measures to Consider for First Use in Federal Surveys

Eleven measures included in the full Inventory assess self-regulation and could be considered for new federal data collections (i.e., they are not currently being used in a federal data collection). Of those, eight measures score a *strong* rating on more than half of the 10 criteria.

Table SR4. Self-Regulation: *Strong* measure(s) to consider for use in future data collections

Name of measure	Strengths	Limitations
Behavior Assessment System for Children, Second Edition (BASC-2)	1. Strong reliability 2. Strong validity 3. Normed on large and diverse sample 4. Available in multiple languages 5. No training required 6. Parent form 7. Covers 2 or more subdomains (behavior problems, self-regulation, and executive function) 8. Broad age range	1. Long administration time (10-20 minutes) 2. Purchase required
Behavior Rating Inventory of Executive Function-Preschool (BRIEF-P)	1. Strong reliability 2. Normed on a large and diverse sample 3. Available in multiple languages 4. No training required 5. Parent and teacher forms 6. Broad age range (2-5)	1. Moderate validity 2. Only covers 1 subdomain (and EF) 3. Longer administration time (10-15 minutes)

Name of measure	Strengths	Limitations
Brief Scale of Temperament (BST)	1. No training required 2. Parent form 3. Covers 2 or more subdomains (self-regulation and emotional competence) 4. Short administration time (< 5 minutes) 5. No purchase required 6. Broad age range (4-7 years)	1. Moderate reliability 2. Moderate validity 3. Normed on small non-representative sample 4. English version only
Childhood Executive Function Inventory (CHEXI)	1. Available in multiple languages 2. No training required 3. Parent and teacher forms 4. Short administration time (5-10 minutes) 5. No purchase required 6. Broad age range (4-15 years)	1. Moderate reliability 2. Moderate validity 3. Non-representative norming sample 4. Only covers 1 subdomain (and EF)
Colorado Childhood Temperament Inventory (CCTI)	1. No training required 2. Parent form 3. Covers 2 or more subdomains (social competence, emotional competence, self-regulation, and executive function) 4. Few items (30 items) 5. No purchase required 5. Broad age range (1-6 years)	1. Moderate reliability 2. Moderate validity 3. Non-diverse or small norming/validation sample 5. English version only
Devereux Early Childhood Assessment Clinical Form (DECA-C)	1. Strong reliability 2. Strong validity 3. Normed on large and diverse sample 4. No training required 5. Parent and teacher forms 6. Covers 2 or more subdomains (social competence, self-regulation, emotional competence, and behavior problems) 7. Broad age range (2-5 years)	1. English version only 2. Purchase required 3. Long administration time (15 minutes)

Name of measure	Strengths	Limitations
Infant Toddler Social Emotional Assessment (ITSEA)	1. Strong reliability 2. Strong validity 3. Size and diversity of norming sample 4. Available in multiple languages 5. No training required 6. Parent and teacher forms 7. Covers 2 or more subdomains (social competence, emotional competence, behavior problems, and self-regulation)	1. Long administration time (25-30 minutes) 2. Purchase required 3. Limited age range (12-36 months)
Matson Evaluation of Social Skills with Youngsters (MESSY)	1. Strong reliability 2. Normed on large and diverse sample 3. Available in multiple languages 4. No training required 5. Parent and teacher forms 6. Covers 2 or more subdomains (social competence, behavior problems, and self-regulation) 7. Broad age range (2-18 years)	1. Moderate validity

In addition, three other measures should be considered as promising because they score a *strong* rating on half of the 10 criteria.

Table SR5. Self-Regulation: *Promising* measure(s) to consider for use in future data collections

Name of measure	Strengths	Limitations
Minnesota Preschool Affect Checklist-Revised/Shortened (MPAC-R/S)	1. Strong reliability 2. Strong validity 3. Size and diversity of norming sample 4. Covers 2 or more subdomains (social competence, emotional competence, behavior problems, and self-regulation) 5. No purchase required	1. English version only 2. Trained administrator required 3. Long administration time (5 minutes observations across 4 different days) 4. Limited age range (3-4 years)
Rothbart's Early Childhood Behavior Questionnaire (ECBQ)	1. Available in multiple languages 2. No training required 3. Parent form 4. Covers 2 or more subdomains (social competence, emotional competence, behavior problems, self-regulation, and executive function) 5. No purchase required	1. Moderate reliability 2. Moderate validity 3. Normed on a non-representative sample (race) 4. Long administration time (1 hour) 5. Limited age range (18-36 months)

Rothbart's Infant Behavior Questionnaire Revised (IBQ-R)	1. Available in multiple languages 2. No training required 3. Parent form 4. Covers 2 or subdomains (social competence, emotional competence, behavior problems, self-regulation, and executive function) 5. No purchase required	1. Moderate reliability 2. Moderate validity 3. Non-diverse or small norming/validation sample 4. Limited age range (3-12 months)

Summary of Review of Measures of Self-Regulation

The Approaches to Learning Scale fielded in the ECLS-K:1998 and ECLS-K:2011, the CBQ fielded in the ECLS-K:2011, the CBQ-VSF, and the PLBS fielded in the FACES 2006 study, meet six or more of the criteria and could be considered for a one-time special feature in *America's Children.* Four measures already used in federal data collection efforts – the ASBI, the shortened version of the Approaches to Learning scale fielded by the FACES team, TBAQ, and Two Bags Task – met half of the selection criteria and therefore could be considered as promising. Seven measures are considered to be strong candidates for first-time use in federal surveys: the BASC-2, BRIEF-P, BST, CHEXI, CCTI, DECA-C, ITSEA, and MESSY. In addition, the MPAC-R/S, ECBQ, and IBQ-R are rated as promising for first time use.

A number of the noted unmet criteria could be rectified, or at least improved, with additional effort. Ten of the suggested measures have a limitation of long administration time: the BASC-2, BRIEF-P, DECA-C, ITSEA, MPAC-R/S, the Rothbart temperament questionnaires (IBQ-R, ECBQ, CBQ, and CBQ-VSF), and Two Bags Task. Though this would require substantial work, a number of the measures could be shortened to take less than 10 minutes while still preserving their other qualities.

In addition, five of the measures, the ASBI, BST, CCTI, DECA-C, and the MPAC-R/S, only have English versions. This limitation could be rectified with careful measures development work to create linguistically and culturally appropriate versions in other languages. Again, we caution that while this is feasible, the effort required for development of a strong version of a measure in another language should not be underestimated.

Another unmet criterion for the ASBI, BST, CHEXI, CCTI, CBQ-VSF, ECBQ, IBQ-R, and TBAQ is development and testing of the measure with representative samples. This limitation is typically addressed through federal survey development which includes pilot testing on diverse samples, along with psychometric testing using the pilot and then the full sample data.

Finally, limited age range is a limitation that needs to be addressed for eight of the indicated promising measures: ASBI, Approaches to Learning–shortened version tested by the FACES team, ITSEA, MPAC-R/S, PLBS, ECBQ, IBQ-R, and TBAQ. As suggested in previous sections, age ranges could be broadened by adapting and testing the measures with different age groups. But this takes substantial work, starting with conceptualization of how a construct is manifested at different ages.

Executive Function

Table EF1 below displays the measures of executive function that could be given further consideration for inclusion in the *America's Children* report and in future federal surveys. As noted earlier, while executive function is not technically a subdomain of social and emotional development, it is included because of its relationship to the development of cognitive skills that support the acquisition and execution of social and emotional behaviors and skills. Based on their ratings, a total of 10 measures (out of 22 reviewed) are suggested for consideration for these purposes. Additional information on each measure's strengths and limitations is outlined in the descriptions following the summary chart.

Table EF1. Selected measures of executive function that meet at least half of the evaluation criteria

Name of Measure	Reliability	Validity	Size and diversity of norming/ validation sample	Availability in languages other than English	Requires trained administrator	Parent/teacher form	Covers 2 or more subdomains	Length of time to administer	Cost / purchase requirement	Covers a wide age range
Behavior Assessment System for Children, Second Edition (BASC-2)	Strong	Strong	Strong	Strong	Strong	Strong	Strong	Moderate	Moderate	Strong
Behavior Rating Inventory of Executive Function-Preschool (BRIEF-P)	Strong	Moderate	Strong	Strong	Strong	Strong	Moderate	Moderate	Moderate	Strong
Childhood Executive Function Inventory (CHEXI)	Moderate	Moderate	Weak	Strong	Strong	Strong	Strong	Strong	Moderate	Strong
Colorado Childhood Temperament Inventory (CCTI)	Moderate	Moderate	Weak	Moderate	Strong	Strong	Strong	Strong	Strong	Strong
Dimensional Change Card Sort (DCCS)	Moderate	Moderate	Moderate	Moderate	Weak	N/A	Moderate	Strong	Strong	Strong
Preschool Learning Behaviors Scale (PLBS)	Strong	Strong	Strong	Strong	Strong	Moderate	Moderate	Strong	Moderate	Moderate

Name of Measure	Reliability	Validity	Size and diversity of norming/validation sample	Availability in languages other than English	Requires trained administrator	Parent/teacher form	Covers 2 or more subdomains	Length of time to administer	Cost / purchase requirement	Covers a wide age range
Rothbart's Children's Behavior Questionnaire (CBQ)	Moderate	Strong	Strong	Strong	Strong	Strong	Strong	Weak	Strong	Strong
Rothbart's Children's Behavior Questionnaire Very Short Form (CBQ-VSF)	Moderate	Strong	Moderate	Strong	Strong	Strong	Strong	Moderate	Strong	Strong
Rothbart's Early Childhood Behavior Questionnaire (ECBQ)[31]	Moderate	Moderate	Moderate	Strong	Strong	Strong	Strong	Weak	Strong	Moderate
Rothbart's Infant Behavior Questionnaire Revised (IBQ-R)[32]	Moderate	Moderate	Moderate	Strong	Strong	Strong	Strong	Weak	Strong	Moderate

[31] The ECBQ also offers a Short Form (107 items) and a Very Short Form (36 items). Information about the duration needed to administer the short forms is not available. The rating in this chart is for the full 201 item measure.

[32] The IBQ-R also offers a Short form (91 items) and a Very Short form (36 items). Information about the duration needed to administer the short forms is not available. The rating in this chart is for the full 184 item measure.

<u>Measures Appropriate for Consideration for Use as Regular Indicator</u>

No measure of executive function included in the full Inventory is currently used by regularly occurring cross-sectional federal surveys.

<u>Measures Appropriate for Consideration for Use as One-time Indicator</u>

Eight measures of executive function included in the full Inventory have been used in longitudinal federal surveys. Of those, three meet the criteria for use as one-time indicators and score a *strong* rating on six or more of the 10 criteria.

Table EF2. Executive Function: *Strong* measure(s) to consider for use as a one-time indicator

Name of measure	Strengths	Limitations
Preschool Learning Behaviors Scale (PLBS), fielded in FACES in 2006	1. Strong reliability 2. Strong validity 3. Normed on large and diverse sample 4. Spanish and English versions available 5. No training required 6. Short administration time (8 minutes)	1. Only teacher form available 2. Purchase required 3. Limited age range (3-5.5 years) 4. Only covers 1 subdomain (and EF)
Rothbart's Children's Behavior Questionnaire (CBQ), fielded in the ECLS-K: 2011	1. Strong validity 2. Representative norming sample 3. Available in multiple languages 4. No training required 5. Parent form 6. Covers 2 or more subdomains (emotional competence, behavior problems, self-regulation, and executive function) 7. No purchase required 8. Broad age range (3-7 years)	1. Moderate reliability 2. Long administration time (1 hour)
Rothbart's Children's Behavior Questionnaire-Very Short Form (CBQ-VSF)	1. Strong validity 2. Available in multiple languages 3. No training required 4. Parent and teacher forms 5. Covers 2 or more subdomains (emotional competence, behavior problems, self-regulation, and executive function) 6. No purchase required 7. Broad age range (3-7 years)	1. Moderate reliability 2. Small norming sample 3. Moderate administration time (less than 15 minutes)

One other measure should be considered as promising because it scores a *strong* rating on half of the 10 criteria.

Table EF3. Executive Function: *Promising* measure(s) to consider for use as a one-time indicator

Name of measure	Strengths	Limitations
Dimensional Change Card Sort (DCCS), fielded in the ECLS-K: 2011	1. Normed on a large and diverse sample 2. Available in multiple languages 3. Short time of administration (5-10 minutes) 4. No purchase required 5. Broad age range (2.5-5 for standard version, 5-7 for broader version)	1. Moderate reliability 2. Moderate validity 3. Requires trained administrator 4. Covers only one subdomain (executive function)

Measures to Consider for First Use in Federal Surveys

Six measures of executive function included in the full Inventory could be considered for new federal data collections. Of these, four measures score a *strong* rating on more than half of the 10 criteria.

Table EF4. Executive Function: Strong measure(s) to consider for use in future data collections

Name of measure	Strengths	Limitations
Behavior Assessment System for Children, Second Edition (BASC-2)	1. Strong reliability 2. Strong validity 3. Normed on large and diverse sample 4. Available in multiple languages 5. No training required 6. Parent and teacher forms 7. Covers 2 or more subdomains (social competence, executive function, behavior problems, and self-regulation) 8. Broad age range (preschool form: 2-5 years; child form: 6-11 years)	1. Long administration time (10-20 minutes) 2. Purchase required
Behavior Rating Inventory of Executive Function-Preschool (BRIEF-P)	1. Strong reliability 2. Normed on a large and diverse sample 3. Available in multiple languages 4. No training required 5. Parent and teacher forms 6. Broad age range (2-5)	1. Moderate validity 2. Only covers 1 subdomain (self-regulation) in addition to EF 3. Longer administration time (10-15 minutes)

Name of measure	Strengths	Limitations
Childhood Executive Function Inventory (CHEXI)	1. Available in multiple languages 2. No training required 3. Parent and teacher forms 4. Short administration time (10 minutes) 5. No purchase required 6. Broad age range (4-15 years)	1. Moderate reliability 2. Moderate validity 3. Non-representative norming sample 4. Only covers 1 subdomain (and EF)
Colorado Childhood Temperament Inventory (CCTI)	1. No training required 2. Parent report 3. Covers 2 or more subdomains (social competence, emotional competence, self-regulation, and executive function) 4. Short administration time (30 items) 5. No purchase required 6. Broad age range (1-6 years)	1. Moderate reliability 2. Moderate validity 3. Small validation sample 4. English version only

In addition, two other measures should be considered as promising because it scores a *strong* rating on half of the 10 criteria.

Table EF5. Executive Function: *Promising* measure(s) to consider for use in future data collections

Name of measure	Strengths	Limitations
Rothbart's Early Childhood Behavior Questionnaire (ECBQ)	1. Available in multiple languages 2. No training required 3. Parent form 4. Covers 2 or more subdomains (social competence, emotional competence, behavior problems, self-regulation, and executive function) 5. No purchase required	1. Moderate reliability 2. Moderate validity 3. Normed on a non-representative sample (race) 4. Long administration time (1 hour) 5. Limited age range (18-36 months)
Rothbart's Infant Behavior Questionnaire-Revised (IBQ-R)	1. Available in multiple languages 2. No training required 3. Parent form 4. Covers 2 or more subdomains (social competence, emotional competence, behavior problems, self-regulation, and executive function) 5. No purchase required	1. Moderate reliability 2. Moderate validity 3. Non-diverse norming/validation sample

Summary of Review of Measures of Executive Function

Of the 23 measures reviewed, four measures, the PLBS, CBQ, CBQ-VSF, and DCCS, could be considered for use as a one-time special feature in *America's Children*. In addition, the BASC-2, BRIEF-P, CHEXI, CCTI, IBQ-R, and ECBQ are appropriate for consideration for new data collections.

Longer administration times are concerns with the BASC-2, BRIEF-P, CBQ, CBQ-VSF, ECBQ, and IBQ-R. Although abbreviated versions could be developed, the psychometric properties of the shortened versions would still need to be assessed, and this is not a small task. However, it is one that could be undertaken if a high priority were placed on inclusion of a measure with this limitation.

Additionally, the CHEXI, CCTI, CBQ-VSF, ECBQ, and IBQ-R suffer from a small and/or non-representative norming sample and the CCTI is available in English only. While these concerns can be remedied, as noted earlier, such remedies are no small undertaking.