Assessing Approaches to Learning in School Readiness: Comparing the Devereux Early Childhood Assessment to an Early Learning Standards-Based Measure

Otilia C. Barbu
David B. Yaden Jr.
Deborah Levine-Donnerstein
Ronald W. Marx
University of Arizona

This study examines the psychometric properties of two assessments of children's approaches to learning: the Devereux Early Childhood Assessment (DECA) and a 13-item approaches to learning rating scale (AtL) derived from the Arizona Early Learning Standards (AELS). First, we administered questionnaires to 1,145 randomly selected parents/guardians of first-time kindergarteners. Second, we employed confirmatory factor analysis (CFA) with parceling for DECA to reduce errors due to item specificity and prevent convergence difficulties when simultaneously estimating DECA and AtL models. Results indicated an overlap of 55% to 72% variance between the domains of the two instruments and suggested that the new AtL instrument is an easily administered alternative to the DECA for measuring children's approaches to learning. This is one of the first studies that investigated DECA's approaches to learning dimension and explored the measurement properties of an instrument purposely derived from a state's early learning guidelines.

Keywords: *readiness, approaches to learning, DECA, EFA/CFA, parceling*

A substantial body of recent research suggests that the level of *school readiness* is an important and even a critical indicator of whether an entering kindergarten child will succeed in school (Haskins & Barnett, 2010; Magnuson, Rohm, & Waldfogel, 2007; K. Snow, 2006, 2011). Several longitudinal analyses of high-quality early childhood programs over the past decade demonstrate that considerable benefits accrue to both individuals and society when children have positive experiences with well-trained caregivers prior to school entry (Heckman, Moon, Pinto, Savelyev, & Yavitz, 2010; Rolnick & Grunewald, 2003). Conversely, when children's high-quality early experiences are absent, lower school performance in the early grades is likely, reducing the possibility of greater success as children advance in school (Pianta, Barnett, Burchinal, & Thornburg, 2009). Thus, educators, child development specialists, researchers, and policy makers could gain knowledge about children at risk at the start of kindergarten (K. Snow, 2011) by examining school readiness factors, and subsequently, they could design policies and programs to ameliorate these risks (cf. also Center on the Developing Child at Harvard University, 2011).

Approaches to learning is a key construct in many policy-based definitions of school readiness, although insufficient research exists on how to measure this construct efficiently and accurately. We address this gap in the research by evaluating the psychometric properties of two assessments of approaches to learning using data from parent and guardian ratings. We compare findings from the Devereux Early Childhood Assessment, a widely used instrument for assessing emotional intelligence in preschool settings (Chain, Dopp, Smith, Woodland, & LeBuffe, 2010) with findings from a researcher-developed, 13-item instrument designed specifically to assess approaches to learning behaviors based on the Arizona State Early Learning Standards (AELS). In this comparison, we examined the psychometric validity of this shorter instrument and its potential efficiency and usefulness for classroom assessment.

Development of State Standards for Early Childhood

Despite definitional differences regarding the nature of school readiness (Graue, 1993, 2006; Kagan, 1990; Scott-Little, Kagan, & Frelow, 2006), an effort to develop early learning standards for pre-kindergarten to third-grade children has arisen among regional and national policy makers (Brito, 2012; Neuman & Roskos, 2005; Scott-Little et al., 2006). In the United States, outcomes from the National

Education Goals Panel (NEGP) in 1995 and the Goals 2000 legislation passed in 1994 (C. E. Snow & Van Hemel, 2008) led most states to create standards for assisting preschool professionals to guide the development of infants, toddlers, and preschoolers (K. Snow, 2011). These standards included the following domains: (a) physical well-being and motor development, (b) social and emotional development, (c) *approaches to learning* (our emphasis), (d) language development, and (e) cognition and general knowledge (see Kagan, Moore, & Bredekamp, 1995, for the foundational definitions). These dimensions also are included in the United Nations Children's Fund (UNICEF) conceptual frame (Brito, 2012). Additional specificity in outlining the primary domains of readiness has been provided in the United States in the School Readiness Act of 2007, which includes language, literacy, mathematics, science, social and emotional functioning, creative arts, physical skills, and approaches to learning.

In an analysis of 46 early learning standards documents published after 1999, Scott-Little et al. (2006) noted the limited amount of attention given in many standards, in particular, to socioemotional development and approaches to learning (see also Chen, Masur, & McNamee, 2010; Niemeyer & Scott-Little, 2001; Thigpen, 2014, for similar statements), two components identified as key to early success in school (NEGP, 1995; C. E. Snow & Van Hemel, 2008). In addition to the uneven discussion among the standards themselves, to our knowledge, no state has developed assessments tailored to any of their standards. Rather, extant measures are used whose norming samples may have different characteristics than their state's population of children, for whom the standards are targeted. Similar limitations exist in the measurement of school readiness internationally (Brito & Limlingan, 2012).

Moreover, concerns with instruments' measurement accuracy with young children (e.g., Isquith, Gioia, & Espy, 2004; Meisels & Fenichel, 1996; K. Snow, 2006, 2011), as well as their ease of use and interpretation (Diamond, Justice, Siegler, & Snyder, 2013; National Association for the Education of Young Children, 2002), raise issues in school readiness assessments. As these assessments were meant to support instruction, identify children at risk and in need of special services, and facilitate program evaluation and accountability, accuracy of measurement is paramount. Particularly, in program evaluation and accountability, important and politically sensitive program evaluations often are conducted by request of funders, such as federal agencies, state and local governments, nongovernment organizations, and private philanthropy that require reliable and valid data for meaningful investment decisions (Center on the Developing Child at Harvard University, 2011). An additional concern is that the high-stakes testing movement (Books, 2004; Kim & Sunderman, 2005; Woodside-Jiron & Gehsmann, 2009) coupled with oversimplified and one-time-used assessments has produced deleterious effects on vulnerable populations (Gehsmann & Templeton, 2011–2012). This concern was also reflected by UNICEF research (Brito & Lumlingan, 2012).

In light of these concerns (e.g., limited, reliable instrumentation based on specific learning criteria), we describe in this article an instrument designed to measure the domain of approaches to learning (AtL) based on one state's early learning standards that is both a valid indication of the construct and an easily administered instrument. In the sections below, we provide additional information on the approaches to learning domain, a related literature on children's executive functioning (EF), and the context of the large-scale study where the new AtL assessment was administered.

Assessing Approaches to Learning

School Readiness and Components of Approaches to Learning

The approaches to learning construct, introduced by Kagan et al. (1995) as a component of school readiness, has been identified as an important domain related to children's positive early achievement outcomes in math, reading, and socioemotional development (Fantuzzo, Bulotsky-Shearer, Fusco, & McWayne, 2005; Ziv, 2013). According to Kagan et al. (1995), this construct comprises a combination of traits such as gender and temperament, predispositions and attitudes conditioned by culture, and learning styles. However, in contrast to predispositions, Kagan et al. suggested that learning styles are malleable and include variables that affect how children attitudinally address the learning process: their openness to and curiosity about new tasks and challenges; their initiative, task persistence, and attentiveness; their approach to reflection and interpretation; their capacity for invention and imagination; and their cognitive approaches to tasks (p. 23).

Given that certain styles of learning are favored over others in the U.S. educational system, Kagan et al. (1995) urged that further research is needed to understand approaches to learning so that all children, despite their diversity of learning styles, could have equal opportunities to learn with appropriate pedagogical adjustments. Two decades after Kagan et al.'s assertion that approaches to learning was an underresearched area, their claim is echoed in more recent surveys of the research literature (e.g., Scott-Little et al., 2006; C. E. Snow & Van Hemel, 2008; K. Snow, 2011), although efforts have been made to develop instruments to measure, in particular, the malleable aspects of this domain.

Using varied definitions of approaches to learning, researchers have studied characteristics such as initiative, curiosity, persistence, engagement, and problem solving (e.g., Bulotsky-Shearer, Fernandez, Dominguez, & Rouse, 2011; Li-Grining, Votruba-Drzal, Maldonado-Carreño, & Hass, 2010). In addition, researchers also consistently find

connections to children's early cognitive and developmental growth with characteristics defined as attentiveness, flexibility, and organization (Ziv, 2013), as well as resourcefulness, goal orientation, and planfulness (Chen et al., 2010; Chen & McNamee, 2007). For the most part, studies consistently demonstrate that variables associated with the approaches to learning domain have unique contributions to children's achievement beyond other important cognitive and demographic variables such as intelligence, receptive and expressive vocabulary, parental income, and education.

Related Definitions of Executive Functioning, Self-Control, and "Everyday Behaviors"

The approaches to learning construct described above shares a great deal in common with the research and literature on executive functions, the latter defined as involving such metacognitive processes as found in working memory, attention shifting, inhibitory control processes, planning, error correction, resistance to interference, and memory updating, to name a few (Blair, Zelazo, & Greenberg, 2005; Carlson, 2005; Schmeichel & Tang, 2015, for additional terminology). As with the definitions of approaches to learning, various terms are associated with executive functioning processes, which include such designations as "impulsivity, conscientiousness, self-regulation, delay of gratification, inattention-hyperactivity, executive function, willpower, and intertemporal choice" (Moffit et al., 2011, p. 2693), descriptions that primarily reflect their disciplinary origin in the medically oriented fields.

Attempting to bring some "ecological validity" or real-life applications to these constructs and behaviors, which have been researched regularly in neuropsychological fields, Galinsky (2010) has focused on the *manifestation* or applications of these executive functions in real life and in the classroom as being critical *life skills* and has described them as follows: focus and self-control, perspective taking, communicating, making connections, critical thinking, taking on challenges, and self-directed and engaged learning. Similarly, Isquith et al. (2004) have designated the practical outcome of such executive function processes in working memory or inhibitory self-control as "everyday behaviors" with the claim that "the child's everyday environments, both at home and at school or day care, are important venues for observing routine manifestations of the executive functions" (p. 406). Other recent work to bring the esoteric, neurological, and clinical terminology into the public awareness is the concept of "grit" described by Duckworth, Peterson, Matthews, and Kelly (2007) as "perseverance and passion for long-term goals" (p. 1087).

Thus, in comparing the executive function and approaches to learning literatures (with few overlapping citations), the latter body of research, although complementary to the executive function work, aims to focus on the malleability, learning, and practical manifestations of behaviors related to executive functioning rather than its actual nature, as related to particular neurological and brain structures (see the special issue of *Developmental Neuropsychology* edited by Blair et al., 2005, for a comprehensive review of this work). It is clear from the rapid expansion of work on early learning standards across the nation (Scott-Little et al., 2006; K. Snow, 2011) that states have made dedicated efforts to outline the behaviors of young children, which will eventually lead to successful achievement in kindergarten and beyond (for recent examples, see Arizona Department of Education, 2005; Connecticut Early Learning and Development Standards, 2014; North Carolina Division of Child Development, 2008).

However, common to both areas (EF and approaches to learning) is the limited research on standardized, reliable, easily administered assessments that practitioners can use to gain information about very young children, in particular, in this important area. Although there are a number of clinical instruments or protocols to measure various aspects of EF (Carlson, 2005), most of these are not designed for normal classroom use, usually requiring one-on-one administration, needing some props, and having limited psychometric stability. Thus, a consensus across all researchers examining EF, approaches to learning, or other self-regulatory behavior (e.g., McClelland, Acock, Piccinin, Rhea, & Stallings, 2013; Poropat, 2014) is that there are few easily administered, psychometrically robust instruments available for teachers or parents to use.

Approaches to Learning Instrumentation

Studies measuring the approaches to learning domain have followed either one of two approaches by (a) using items initially developed for other instruments and purposes or (b) developing dedicated instruments focused on approaches to learning itself. An example of the first are secondary analyses (e.g., Li-Grining et al., 2010) of the Early Childhood Longitudinal Study–Kindergarten Cohort (1998–1999). This study used a small subset of items taken originally from the Social Skills Rating System (Gresham & Elliot, 1990) in which parents and teachers rated a child's behavior on aspects of persistence in a task, curiosity, creativity, ability to concentrate, ability to work independently, and paying attention (National Center for Educational Statistics, 2010).

The second approach includes instruments developed to measure aspects of approaches to learning directly (e.g., Chen & McNamee, 2007; McDermott, Leigh, & Perry, 2002) and comprise rating scales of children's behaviors in a preschool classroom based on the frequency of various activities (e.g., individual engagement with books or peer interaction during play), as occurring *very often, sometimes*, or *never*. For example, in their Bridging assessment, Chen and McNamee (2007) measured approaches to learning

across different school tasks (e.g., moving to music, block play, and counting) while rating children's "Initial Engagement," "Focused Attention, Planfulness," "Goal Orientation," and "Resourcefulness." This assessment is based on Vygotsky's (1931/2012) sociocultural theory, which infers that children's behaviors may vary in the function of different task demands.

In contrast, McDermott et al.'s (2002) Preschool Learning Behavior Scale (PLBS) is a 29-item instrument designed to capture children's overall behaviors across various activities according to their "Competence/Motivation," "Attention/ Persistence," and "Attitude Toward Learning." Developers of both instruments designed them for early childhood and kindergarten teachers to administer to children, but McDermott et al.'s AtL instrument is the sole one that is normed (see Fantuzzo, Perry, & McDermott, 2004).

Each of these instruments represents one of two major philosophical views on how children's learning may be assessed. Bridging (Chen & McNamee, 2007) is a portfolio-type assessment, assessing a range of performances across many activities over time and requiring an integrative, qualitative judgment related to the child's development. The Preschool Learning Behavior Scale (McDermott et al., 2002), on the other hand, uses a quantitative rating scale to index degrees of performance relative to previously identified latent factor structures. The merits of each view have been debated extensively among early childhood researchers and educators (Meisels & Fenichel, 1996; Pianta, Barnett, Justice, & Sheridan, 2012). For example, Gilliam and Frede (2012) stated that early childhood development is too "variable and rapid," and children are "not consistent" in demonstrating their abilities for any test to capture a child's ability accurately in the brief assessment time. Nonetheless, the selection of either a qualitative portfolio assessment or rating scale measures seems largely based on the user's philosophical preferences, with those claiming a sociocultural orientation leaning toward portfolio assessment and those with more of a cognitive or psychometric perspective leaning toward rating scale measures. Nonetheless, it is clear in the research literature that the majority of studies examining approaches to learning in young children favor the use of rating scales that are examined via statistical analyses.

A standards-based measure of approaches to learning. The new AtL assessment we describe below, as well as the Devereux Early Childhood Assessment (DECA), fall into the latter category and share the general advantages of teacher and guardian rating scales in terms of amount of information obtained, ease of administration, scoring, and prior reliability and validity testing (see Campbell & James, 2007). Although the DECA was designed primarily for early identification of social and emotional problems in young children, the Devereux Foundation (2003) has reported that the instrument measures an approaches-to-learning construct. However, we

could not find published literature or research studies examining the psychometric properties of this aspect of the DECA instrument or any descriptions from the Devereux Foundation of items dedicated to assessing approaches to learning behaviors.

Related to the development of the shorter AtL instrument described here, a recent review of research (Diamond et al., 2013), funded by the Institute for Education Science (IES) on early intervention and education, highlighted the need for efficient, easily administered assessments with strong evidence of reliability and validity for use in research and educational applications. Diamond et al. (2013) suggested that the research is inadequate in early assessment regarding important characteristics, such as approaches to learning, and additional research is needed to add to the knowledge base in this arena.

The Current Study

In this article, we describe research regarding the psychometric properties of two assessments of approaches to learning, an important domain of school readiness. The context for this study was an Arizona early childhood initiative, First Things First (FTF), a comprehensive, statewide early childhood program intended to increase access to health and educational services for families and children, birth to age 5 years, and to improve preschoolers' readiness for school. The project was designed to include a researcher-developed instrument created to assess approaches to learning as defined by the AELS (Arizona Department of Education, 2005) and the widely used and standardized DECA (LeBuffe & Naglieri, 1999a, 1999b).

The AtL scale was part of a demographic questionnaire in a larger battery of readiness assessments that included direct measures of language, literacy, math, and demographics (e.g., height and weight). The researcher-designed items rating kindergarteners' AtL were derived from the state's early learning standards, which specified that children should be developing attitudes and behaviors characterizing initiative, curiosity, engagement, persistence, reasoning, and problem-solving skills. The AtL instrument also was designed to reduce participant response burden by decreasing time to complete the assessment portfolio for each child and to reduce overall cost in the next stage of the project.

The purpose of this study is to examine the psychometric properties of the DECA and AtL assessments and to ascertain whether these instruments measure the same or different approaches to learning domains based on parent/guardian perceptions of their children, as defined by the AELS (Arizona Department of Education, 2005) and the DECA (LeBuffe & Naglieri, 1999a, 1999b). We accomplish these goals by (a) providing evidence for validity and reliability of the DECA and AtL instruments and presenting recent results based on contemporary statistical techniques concerning

DECA's validity, (b) testing whether the parent/guardian data from the DECA rating scale fit the hypothesized DECA measurement model to determine the degree to which the DECA's model as a whole is consistent with the empirical data, (c) developing confirmatory factor analysis (CFA) models for DECA data by using the *parceling* procedure to reduce the errors due to each item's specificity and to prevent convergence difficulties when DECA and AtL models were estimated simultaneously, and (d) evaluating the combined DECA-AtL model and examining the relationship between DECA and AtL questionnaires.

Method

Sample

We used a proportional, stratified random sampling approach that included 1,145 kindergarten children attending 82 schools drawn from 48 districts across the state of Arizona. We used this sampling design to ensure that the sample was randomly selected and representative of the population. We randomly selected the sample at the child and school level that included three types of schools (public, private, and charter), from three different regions (northern, central, and southern). This sample distribution, derived across the type of schools and regions, was in close agreement with the state proportions (see Barbu, Levine-Donnerstein, Marx, & Yaden, 2012; Barbu, Marx, Yaden, & Levine-Donnerstein, 2015; Yaden et al., 2011).

The average age of the children in the sample was 5 years, 8 months with an age range from 5 to 6 years, 7 months. Fifty-one percent of the children were male, with 49.1% of all children identified as Hispanic. In the present sample, 5.5% of the children had an individualized educational plan, which is required for children with identified special needs, 51.3% were eligible for free or reduced-price lunch, and 73.2% had attended an out-of-home child care, nursery school, or pre-kindergarten program, including Head Start (Barbu et al., 2012). For the analyses, the rating scales from 1,025 families with 1,145 children were examined (some families had multiple children in the sample).

Procedures

For each participating school, a directory of first-time kindergarten students was obtained. From these directories, 14 first-time kindergarten children were randomly selected per school or from one class if only one class existed. From schools with two kindergarten classrooms, seven children were randomly selected from each classroom; from schools with three classrooms, five children per classroom were randomly drawn; and from schools with more than three classrooms, first, three classes were randomly selected, and then the targeted number of children (five per classroom) was randomly drawn from each. After children were identified,

the teacher sent home the child with a packet containing a letter explaining the purpose of the study, a parent consent form, and a child assent form. The research staff then called parents to confirm they received the packet, further explained the study, responded to any questions, and invited them to participate.

Instrumentation

The parents/guardians of participating students completed two questionnaires: the newly developed AtL instrument and the DECA.

AtL rating scale. The newly developed AtL rating scale included 13 items derived from the state's early learning guidelines. To illustrate its alignment with state standards and the DECA, Table 1 provides a comparison of the AtL items, the learning indicators from the AELS (Arizona Department of Education, 2005), and the selected items from the DECA with similarly worded items. The AtL instrument employs a 4-point scale: (a) *proficient* (coded as 3), child demonstrates skill, knowledge, or behavior consistency (i.e., regularly); (b) *in progress* (coded as 2), child demonstrates skill, knowledge, or behavior with some regularity; (c) *not yet* (coded as 1), child cannot perform the skill, knowledge, or behavior; and (d) *don't know* (coded as 0).

In a previous analysis, we investigated the psychometric properties of the AtL instrument (Barbu et al., 2015) and found a one-factor structure via exploratory factor analysis (EFA) and CFA. Moreover, early childhood educators and researchers from a tri-university research consortium (University of Arizona, Arizona State University, and Northern Arizona University) assessed the face validity and content validity of this rating scale. These results, combined with evidence of reliability (0.83 for latent variables and ranging between 0.52 and 0.79 for manifest variables) and structural validity of the instrument (Barbu et al., 2015), supported the educational utility of the AtL as a tool for measuring school readiness among kindergarteners with a population similar to children in Arizona.

DECA. For comparison purposes with the researcher-developed AtL instrument and to index broader socioemotional characteristics in the larger study, we administered the DECA, a standardized, norm-referenced assessment of within-child protective factors, designed originally to identify resilience in children ages 2 to 6 (LeBuffe & Naglieri, 1999a, 1999b). The DECA's 37 items are measured on a 5-point Likert scale indexing behaviors as occurring *never, rarely, occasionally, frequently*, or *very frequently*. They are designed to measure four distinguishable latent variables: initiative (IN—11 items), self-control (SC—8 items), attachment (AT—8 items), and behavioral concerns (BC—10 items). Used in practice, DECA provides an option for

TABLE 1

Comparisons Between the 13-Item Approaches to Learning (AtL) Scale, Statements From the Social Emotional Domain of the Arizona Early Learning Standards,[a] and Selected Items of the Devereux Early Childhood Assessment.

AtL Scale Item	Arizona Early Learning Standards Indicator/Strand/Concept	Devereux Early Childhood Assessment Item: Subscale
1. Sustains positive interactions with other children (e.g., When doing a puzzle, child asks if he can help. The children finish the puzzle together).	Initiates and sustains positive interactions with adults and friends. **Strand 2:** Social interactions with others *Concept 2:* **Cooperation**	33. Cooperate with others? *Subscale:* **Self-Control** 9. Touch children/adults inappropriately? *Subscale:* **Behavioral Concern**
2. Sustains positive interactions with familiar adults.	Initiates and sustains positive interactions with adults and friends. **Strand 2:** Social interactions with others *Concept 2:* **Cooperation**	10. Show affection for familiar adults? *Subscale:* **Attachment** 9. Touch children/adults inappropriately? *Subscale:* **Behavioral Concern**
3. Has friends.	Child sees his friend crying and then gives her a hug.[b] **Strand 1:** Knowledge of self *Concept 2:* **Recognition and expression of feelings** Child trades toys with a friend. **Strand 2:** Social interactions with others Concept 2*:* **Cooperation** Child reminds friends that running is for outside. **Strand 3:** Responsibility for self and others *Concept:* **Self-control** Child inquires why his friend is not at school. **Strand 4:** Approaches to learning *Concept:* **Curiosity**	20. Start or organize play with other children? *Subscale:* **Initiative** 25. Share with other children? *Subscale:* **Self-Control**
4. Adjusts behavior to correspond to different settings (e.g., child knows when to use a "quiet voice").	Adjusts behavior for alternate activities and in different settings of the learning environment. **Strand 3:** Responsibility for self and others *Concept 1:* **Self-control**	30. Accept another choice when his or her first choice was unavailable. *Subscale:* **Self-Control**
5. Follows rules.	Understands and follows rules in the learning environment. **Strand 3:** Responsibility for self and others *Concept 1:* **Self-control**	4. Listen to or respect others? *Subscale:* **Self-Control**
6. Manages transitions (e.g., When it is time for a story, child puts away the blocks and goes to hear the story).	Manages transitions, daily routines, and unexpected events. **Strand 3:** Responsibility for self and others *Concept 1:* **Self-control**	30. Accept another choice when his or her first choice was unavailable. *Subscale:* **Self-Control**
7. Shows curiosity as a learner.	Shows interest in learning new things and trying new experiences. **Strand 4:** Approaches to learning *Concept 1:* **Curiosity**	19. Try or ask to try new things or activities? *Subscale:* **Initiative**
8. Makes independent decisions (e.g., instead of playing with friends, the child decides to read a story).	Makes decisions independently. **Strand 4:** Approaches to learning *Concept 2:* **Initiative**	36. Make decisions for himself or herself? *Subscale:* **Initiative** 2. Does things for himself or herself? *Subscale:* **Initiative**
9. Attends to tasks (e.g., child works on building a Lego structure throughout the course of the day).	Continuously attends to a task. **Strand 4:** Approaches to learning *Concept 3:* **Persistence**	24. Focus his or her attention or concentrate on a task or activity? *Subscale:* **Initiative**
10. Seeks help when encountering a problem (e.g., child tells adult, "He took my toy.").	Seeks adult assistance when support is required. **Strand 4:** Approaches to learning *Concept 5:* **Problem solving**	31. Seeks help from adults/children when necessary? *Subscale:* **Attachment**
11. Copes with frustration (e.g., child says, "We have to go inside, it's raining. We can come back out when it stops.").	Copes with frustration or disappointment. **Strand 4:** Approaches to learning *Concept 3:* **Persistence**	13. Handle frustration well? *Subscale:* **Self-Control**
12. Takes risks during learning situations.	Is willing to take risks and consider a variety of alternatives. **Strand 4:** Approaches to learning *Concept 6:* **Confidence**	12. Takes risks during learning situations.
13. Shows respect for toys.	Shows respect for learning materials and toys. **Strand 3:** Responsibility for self and others *Concept 2:* **Respect**	13. Shows respect for toys.

[a]During 2009 when the study was conducted, the second edition of the Arizona Early Learning Standards (Arizona Department of Education, 2005) was in force. In this document, approaches to learning was one of four strands in the social and emotional standard. In the third edition (2013) of the standards, the approaches to learning strand was recast as a standard itself. No new wording or descriptions were added in 2013.
[b]There is no specific wording in the social-emotional domain about "having friends" per se, but multiple statements for all indicators occur across all strands indicating that the child interacts with friends in various settings and situations.

researchers to further group together the first three factors (i.e., IN, SC, and AT) into a total protective factor (TPF) that evaluates the frequency of positive behaviors. The TPF measures positive child outcomes and includes two domains: social and emotional development and approaches to learning (Devereux Foundation, 2003). The instrument is used for screening, progress monitoring (e.g., DECA can be administer two to three times per year), intervention planning, and research (Henderson & Strain, 2009).

The DECA is considered a reliable instrument for assessing children's protective factors (LeBuffe & Naglieri, 1999b), based on research findings from infants to preschoolers (Barbu et al., 2012; Brinkman, Wigent, Tomac, Pham, & Carlson, 2007; Buhs, 2003; Bulotsky-Shearer, Fernandez, & Rainelli, 2013; Chittooran, 2003; Jaberg, Dixon, & Weis, 2009; Lien & Carlson, 2009; Meyer, 2008; Ogg, Brinkman, Dedrick, & Carlson, 2010).

Although the DECA measures a range of important constructs, recent research using structural equation modeling techniques has found issues related to DECA's validity that may limit its use; therefore, researchers have recommended further investigation of DECA's psychometric properties. For example, Ogg et al. (2010) found 10 problematic item pairs as a source of poor model fit of the DECA instrument in their sample of 1,344 participants recruited from 25 Head Start centers across a four-county region in the Midwest. In addition, Barbu et al. (2012) found insufficient discriminant validity of the DECA instrument based on samples of parents' and teachers' ratings of 1,145 entering kindergartners in the Southwest. More recently, Bulotsky-Shearer et al. (2013) found that DECA's factor structure was not adequate using a large sample of culturally and linguistically diverse Head Start children ($N = 5,197$) in the Southeast.

A possible explanation for the departure of these recent conclusions from the results reported in DECA's technical manual (LeBuffe & Naglieri, 1999a) and obtained with classical statistical analyses is that the traditional multivariate procedures are incapable of assessing and correcting for measurement errors (Byrne, 2009). Therefore, correlations between two latent variables in structural equation modeling (SEM) could be more than twice that for the individual observed variables analyzed with classical statistics, and hidden effects among latent variables (i.e., multicollinearity) could remain undetected (Bollen, 1989).

Data Analysis

We approached the data analysis in three phases. First, we developed CFA models for our sample DECA data based on the theoretical structure explained above. The purpose of assessing a model's overall fit is to determine the degree to which the model as a whole is consistent with the empirical data. The null hypothesis states that the model fits the population data perfectly, and the aim is not to reject this null hypothesis. Although a wide range of goodness-of-fit indices

have been developed to provide measures of a model's overall fit, one is not superior to the others in all situations (Brown, 2006). In addressing the ordinal nature of the observed variables, Byrne (2009) suggested five goodness-of-fit indices to test and respecify the hypothesized model: chi-square (χ^2), root mean square error of approximation (RMSEA), goodness-of-fit index (GFI), comparative fit index (CFI), and expected cross-validation index (ECVI). The following cutoff criteria (Hu & Bentler, 1999) were used as guidelines for goodness-of-fit indices between the target model and the observed data: (a) GFI values close to 0.95 or greater, (b) RMSEA values close to 0.06 or below, and (c) CFI values close to 0.95 or greater. In addition, the model with the smallest ECVI value suggests that the hypothesized model is well fitted and represents a reasonable approximation to the population (Browne & Cudeck, 1993).

Second, we developed CFA models for our sample DECA data by constructing *parcels* based on T. D. Little, Cunningham, Shahar, and Widaman's (2002) technique and confirmed the correctness of each parcel by calculating their internal consistency. A parcel is an aggregated indicator composed of the sum or average of several items that measure the same construct. Our decision of how to select these indicators for each parcel was based on the loadings obtained for each factor, such that the coefficients' center of mass would be preserved. First, we identified the three items with the highest loadings to anchor the three-parcel solution for each factor (Nasser & Wisenbaker, 2003). Second, we added three items with the next highest item-to-factor loadings to the anchors in an inverted order such as the highest loaded item from among the anchor items would be matched with the lowest loaded item from among the second selections. Finally, we placed lower loaded items with higher loaded parcels. Thus, parcels resulting from this technique had a different number of items, and they were selected to achieve a reasonable balance, in agreement with the method presented above (T. D. Little et al., 2002).

In the present analysis, a parceling approach for the DECA instrument was appropriate for two reasons. First, the errors were reduced in the final variance-covariance matrix due to each item's specificity, in addition to the randomness of each item's errors. Therefore, the value of GFI was expected to improve with a correct selection of factors forming each parcel (T. D. Little et al., 2002). Second, factor structures can be difficult to determine when analyzing individual items from a lengthy questionnaire. Practical convergence difficulties could arise in LISREL 8.80 when analyzing a larger number of manifest variables. In our case, this occurred when we added 13 more items from AtL to our sample DECA data, which already contained 37 measured variables.

The potential advantages of using parcels include the following: (a) Parcels better approximate normality than individual items (Brown, 2006), (b) they improve reliability and relationships with other variables (Kishton & Widaman, 1994), and (c) models based on parcels may be less complex

than models based on individual items (i.e., fewer parameters, smaller input matrix) (Brown, 2006). However, a variety of disadvantages could exist in the following cases: (a) The assumption of unidimensionality is not met (i.e., each indicator loads on a single factor and the error terms are independent), (b) the likelihood of improper or nonconvergent solutions increases as the number of parcels decreases (Nasser & Wisenbaker, 2003), and (c) the use of parcels may not be feasible in situations in which too few items form a sufficient number of parcels (Brown, 2006). In our case, the *parceling* was a vital solution in solving the problem of convergence when DECA and AtL models were estimated simultaneously.

Third, we examined the relationship between DECA and AtL questionnaires to ascertain the extent to which these instruments measured the same or different approach-to-learning domains. In this phase, the parameters for both DECA and AtL models were estimated simultaneously, and the hypothesized structure was analyzed as a function of the overall goodness of fit for the combined model, followed by an examination of standardized residuals. We investigated the correlations among latent variables for the completely standardized solution, because the two tests (i.e., DECA and AtL) were measured using two different Likert scales.

Results

Data Missingness

For the current study, less than 3% of the data for guardians were missing, and we concluded that these results raised no concerns regarding data missingness. Data in this study were ordinal based on DECA's five rating categories and AtL's three rating categories for measurement. As a result, we generated polychoric correlation matrices using PRELIS 2.0. Two possible approaches are implemented in PRELIS 2.0 when data are missing: (a) the expectation maximization algorithm (EM; Dempster, Laird, & Rubin, 1977) and (b) Markov chain Monte Carlo method (MCMC; Gilks, Richardson, & Spiegelhalter, 1995). Collectively known as multiple imputation (MI), these procedures replace the missing values across multiple variables under the assumptions of either data missing at random (MAR) or data missing completely at random (MCAR) and multivariate normality. The assumption of multivariate normality of observed data is usually violated when the outcomes are rated on Likert-type scales (Lubke & Muthén, 2004), but the weighted least squares (WLS) estimator adjusts for this violation in large samples with categorical outcomes (Brown, 2006).

We used MCMC imputation technique and full-information maximum likelihood (FIML) estimation to correct for missing data and evaluate the consistency between the missing data treatment methods. The MCMC imputations were based on generation of five imputed data sets. The parameter estimates (i.e., factor loadings, available model fit statistics, etc.) for MCMC imputation and FIML were nearly identical across both missing data treatment methods; therefore, we reported the MCMC imputation results in agreement with data analyses in the larger project from which these data were drawn (Yaden et al., 2011). In our case, Little's MCAR test, $\chi^2(535, 984) = 951.11$, $p < .001$, based on the evaluation of the homogeneity of the available means for different patterns of incomplete data (R. J. A. Little, 1988), was statistically significant, indicating that the data for DECA and AtL instruments were not missing completely at random. However, the Separate Variance Test indicated no relationship between missingness and variables, and thus we proceeded under MAR assumptions (Tabachnick & Fidell, 2013).

The original sample of 2,290 DECA and AtL questionnaires completed by the parents/guardians of 1,145 participating students was reduced to 2,075 (1,053 for DECA and 1,022 for AtL) by eliminating the participants without records. Consequently, data from participants with at least one response were imputed to complete the missing data, using the MI procedure (Rubin, 1987) under the MAR assumption, with a WLS estimator to correct the violation of multivariate normality. Finally, we identified 984 parent/guardian questionnaires for both DECA and AtL after eliminating the participants with only one questionnaire (i.e., one DECA or one AtL questionnaire).

DECA Model Analysis

In the first model (M1) for the DECA (see Table 2), we followed the official prescription in assigning the indicators for each of the four latent variables considered. The variance of each latent variable was fixed to 1. The minimum fit function value was $\chi^2(627, n = 1,053) = 2,764.16$. The values of RMSEA (.057), ECVI (2.78), and CFI (.97) indicated a reasonable fit of the model. Still, the value of the GFI (.74) was low. Further attempts to improve the model by freeing the covariances among the measured indicator errors, as suggested by the modification indices, did not provide a satisfactory improvement of the GFI measure. Consequently, we used item parcels instead of individual items.

In parceling the DECA model, we averaged two, three, or four original indicators together to generate a new indicator (Bollen, 1989) and reduced the number of indicators for each of the four factors to three. Table 3 shows a detailed diagram of this item parceling selection process. Internal consistency reliability of the parcels was acceptable, ranging from .64 for the DECA 2 parcel to .87 for the DECA 12 parcel. In addition, the normality of the parcels indicated that the values for skewness were less than an absolute value of $sk = 1.92$, and the values for kurtosis were less than absolute value of $kur = 6.13$. Neither skewness nor kurtosis exceeded recommended cutoffs, |2.00| and |7.00| (Curran, West, & Finch, 1996), respectively, indicating adequate univariate normality.

We formed and investigated two parcel models. The first model represented a parceled model consisting of four

TABLE 2

Fit Statistics of the Devereux Early Childhood Assessment Instrument for Guardians.

Model	Goodness of Fit						
	χ^2	*df*	RMSEA	CFI	GFI	ECVI	$\Delta\chi^2$
M1	2764.16	627	.057	.97	.74	2.78	
M2	352.09	48	.078	.97	.98	.39	
M3	352.09	48	.078	.97	.98	.39	
M4	281.19	47	.069	.97	.99	.33	70.90
M5	200.66	46	.057	.98	.99	.25	80.53

Note: χ^2 = chi-square fit statistic; RMSEA = root mean square error of approximation; CFI = comparative fit index; GFI = goodness-of-fit index; ECVI = expected cross-validation index; M1 = original model; M2 = model with four exogenous latent variables; M3 = model with two exogenous and three endogenous latent factors; M4 = model with item parcel DECA_7 loaded on attachment (AT) and self-control (SC); M5 = model with item parcels DECA_7 and DECA_8 loaded on AT and SC.

exogenous latent variables with three indicators per variable (M2). The second model included a structural model with two exogenous and three endogenous latent factors (M3). The fitting parameters for both models were similar: RMSEA (.078), ECVI (.39), CFI (.97), GFI (.98), and χ^2(48, *n* = 1,053) = 352.09. Since including a structural relationship did not lead to an improvement in the fitting of the original model, this more complex model (M3) was abandoned, and the next step consisted of improving the fitting of model M2. Further examination of the modification indices suggested that some pairs of parcels shared similar content. Therefore, we inspected two nested models M4 and M5 (see Table 2). The results indicated that the item parcels DECA_7 and DECA_8 in M5 were dependent not only on the AT latent variable as initially proposed but also on the SC variable. Inspection of the goodness-of-fit indices RMSEA (.057), ECVI (.25), CFI (.99), and GFI (.99) indicated a good model fit (see Figure 1). In addition, inspection of the stem-leaf plot of the standardized residuals revealed a symmetrically clustered distribution around zero, accounting for a reasonably well-fitted model.

Therefore, on the basis of the fit statistics, in conjunction with the following evidence—(a) a decrease of the ECVI value, (b) the distribution of the standardized residual values, and (c) a significant reduction in χ^2 between models M4 and M5, $\Delta\chi^2$(1, *n* = 1,053) = 281.19 – 200.66 = 80.53—we concluded that the parent/guardian data fit the hypothesized DECA measurement model and considered model M5 to represent the best model fit of the DECA instrument.

Combined DECA-AtL Model Analysis

Figure 2 presents the standardized factor intercorrelations, factor loadings, and residuals for the combined AtL-DECA model for parent/guardian data. The overall goodness-of-fit indices for this model, χ^2(261, *n* = 984) = 925.08, RMSEA = .051, CFI = .97, and GFI =.98, indicated

a good model fit. This outcome was supported by standardized residuals below 2.00 and standardized factor loadings below 1.00 (Brown, 2006). Further attempts of improving the model fit by allowing the measurement errors to correlate did not result in a considerable change, and thus, the nested models were discarded.

First, the BC measure correlated negatively with all the other factors from both tests (i.e., IN, SC, AT, and AtL). This result was consistent with the way in which the instruments were constructed. The items measuring the BC factor were designed to address developmental factors that are contraindicative of the factors in the other scales. Therefore, items indicating large behavioral concerns were designed as indicators of a child's low socioemotional development.

Second, correlations between the latent factor of the AtL instrument and latent factors of the DECA instrument ranged between .74 and .85 in absolute value, *r* (AtL, BC) = –.74, *r* (AtL, IN) = .79, *r* (AtL, AT) = .83, and *r* (AtL, SC) = .85. These results indicated that the AtL explained between 55% and 72% of the variance in testing the DECA, and thus, the four latent factors of the DECA instrument were strongly related to the factor of the AtL instrument.

Discussion

This is one of the first studies that examined an assessment developed specifically to reflect children's performance on a particular state standard related to approaches to learning. Thus, aligned with the need for valid and reliable instruments that are easy for parents/guardians to complete (Diamond et al., 2013), we developed a 13-item instrument based on the AELS (Arizona Department of Education, 2005) that examined approaches-to-learning behaviors among kindergarteners. Additionally, this is one of the first research efforts of which we know that investigated the DECA's approaches-to-learning dimension. Although DECA is a widely used tool that measures the

TABLE 3
Item Parcel Selection for the DECA Data.

Factor Name	Old Item	New Item
IN	DECA3	DECA_1
	DECA12	
	DECA19	
	DECA2	DECA_2
	DECA7	
	DECA16	
	DECA20	
	DECA24	DECA_3
	DECA28	
	DECA32	
	DECA36	
SC	DECA4	DECA_4
	DECA5	
	DECA34	
	DECA13	DECA_5
	DECA25	
	DECA21	DECA_6
	DECA30	
	DECA33	
AT	DECA1	DECA_7
	DECA10	
	DECA6	DECA_8
	DECA17	
	DECA29	
	DECA22	DECA_9
	DECA31	
	DECA37	
BC	DECA8	DECA_10
	DECA11	
	DECA18	
	DECA9	DECA_11
	DECA15	
	DECA23	
	DECA14	DECA_12
	DECA26	
	DECA27	
	DECA35	

Note: DECA = Devereux Early Childhood Assessment; IN = initiative; SC = self-control; AT = attachment; BC = behavioral concerns.

socioemotional resilience of young children, we have found no studies in the literature that explore the approaches-to-learning component of the instrument.

In previous analyses (Barbu et al., 2015), we found that the 13-item AtL rating scale has a valid and reliable one-factor structure, and it can be reliably used to assess children's approaches-to-learning behaviors. In this study, our purpose was to ascertain the extent to which AtL and DECA
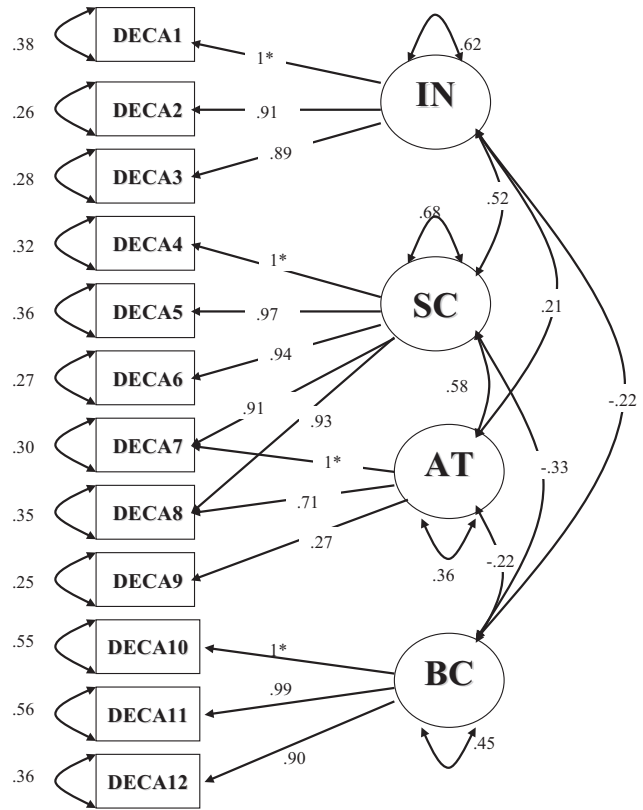

FIGURE 1. *Standardized factor intercorrelations, factor loadings, and residuals of the Devereux Early Childhood Assessment (DECA) model after parceling (M5) for parent/guardian samples. Overall fit of the M5 model: χ^2(46, n = 1,053) = 200.66, comparative fit index = .98, goodness-of-fit index = .99, and root mean square error of approximation = .057. IN = initiative; SC = self-control; AT = attachment; BC = behavioral concerns.*

instruments measured similar approaches-to-learning domains (Arizona Department of Education, 2005; LeBuffe & Naglieri, 1999a, 1999b) based on parent/guardian's perceptions of their children. Through an analysis of a combined DECA-AtL model of the two instruments, we found that although their Likert rating scales differed, they measured a similar approaches-to-learning domain, and thus, used simultaneously, these instruments are redundant. Our findings also supported the content validity of the AtL instrument, being based on a comparison of the AtL items with those of the AELS and DECA (see Table 1) and an analysis at the construct level between DECA and AtL instruments.

Unfortunately, an analysis at the item level between these two instruments was not possible because DECA's technical manual (LeBuffe & Naglieri, 1999a, 1999b) does not identify which of its items is included in the approaches-to-learning domain. Moreover, our findings that the DECA subscales (i.e., self-control, initiative, attachment, and behavioral concerns) strongly correlated with AtL calls into question their
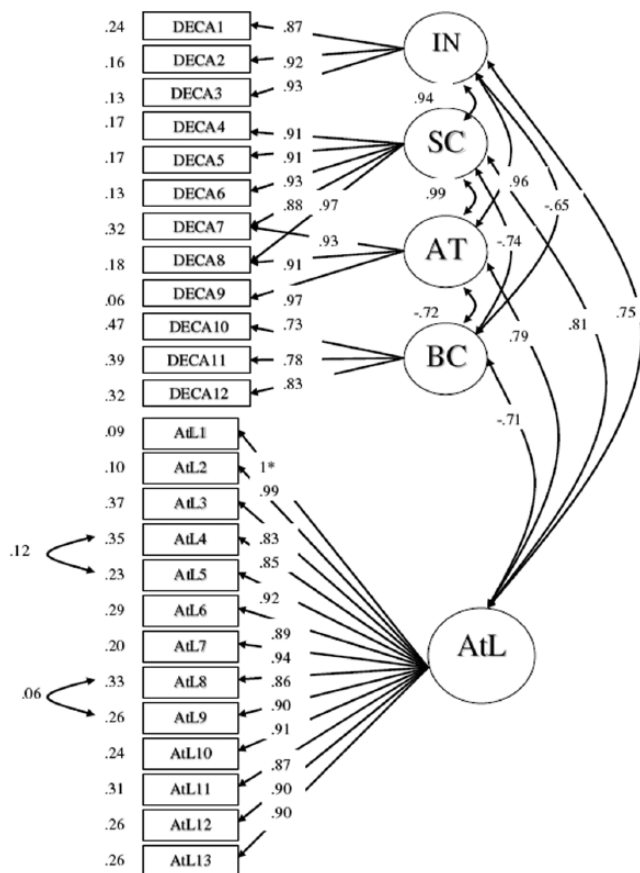
FIGURE 2. *Standardized factor intercorrelations, factor loadings, and residuals of the Devereux Early Childhood Assessment (DECA)–approaches to learning (AtL) combined model for parent/guardian samples. Overall fit of the DECA and AtL combined model: $\chi^2(261$, n = 984) = 925.08, comparative fit index = .97, goodness-of-fit index =.98, and root mean square error of approximation = .051. IN = initiative; SC = self-control; AT = attachment; BC = behavioral concerns.*

uniqueness in measuring broader aspects of social and emotional development, in addition to the more narrow approaches-to-learning construct, which it states to measure (Devereux Foundation, 2003). In a previous analysis (Barbu et al., 2012), we found that the DECA subscales of self-control, initiative, and attachment lacked discriminant validity; therefore, scores on these subscales should be viewed with great caution regarding making decisions about children's performances or their resilience. Thus, we invite the DECA's authors to test and reevaluate the content and structural validity of their instrument and encourage research to explore the approaches-to-learning dimension of the DECA instrument in behavior and psychoeducational studies.

The statistical procedures used in this study provided an improved set of techniques: (a) to appropriately mitigate missing parent/guardian data and increase sample size; (b) to obtain the most accurate estimated model fit between

DECA's parent/guardian empirical data and its hypothesized measurement model, which indicated that DECA's model as a whole was consistent with these data; and (c) to demonstrate evidence that AtL and DECA questionnaires correlate and both measure the approaches-to-learning domain. Moreover, using CFA with the *parceling* technique when simultaneously estimating the DECA and AtL models presents a robust method of convergence. This analytical approach provides an approach for researchers to find new dimensions of the approaches-to-learning domain in existent questionnaires in which this construct is not apparent.

Limitations

Three potential sources of limitations are highlighted in this study. First, the use of the item parceling procedure could affect the structural parameter estimates and might lead to a biased estimate of model parameters in some situations, while allowing for a slight improvement of the model fitting. However, researchers have demonstrated that parcels do not outperform models based on individual items (Hau & Marsh, 2004). Therefore, we concluded that this technique was appropriate for our study.

A second limitation consists in data composed of ordinal measurement scales. Although Likert scales are widely used methods of capturing ratings from respondents in the social and behavior sciences, they produce imprecise response measures from a restrained number of categories (e.g., 5-point rating scales); thus, information might be lost due to the limited resolution of categories (i.e., less precision) (Neibecker, 1984). In our case, considering that DECA had five measurement categories and AtL included four, the presence or absence of any additional categories in Likert scales could lead to different results.

Finally, the *Don't know* and *Not yet* categories of the AtL rating scale could lead to misinterpretations of the results if the difference between categories is not distinguished or the categories are not coded appropriately. The *Don't know* category indicated that the observer did not have enough information about the child or relevant settings to determine whether the target behavior had occurred, whereas *Not yet* indicated that the behavior was not observed in situations where the rater had enough information about the child to recognize it. In a preliminary analysis, we eliminated the *Don't know* category to investigate the coder effect, and we found no significant difference between results. However, this aspect should be examined empirically in future studies.

Conclusion

Researchers and program evaluators charged with assessing and evaluating early childhood programs have little guidance in choosing what assessments might be appropriate for measuring early learning standards for particular populations of children in different geographical areas. Also,

while standardized instruments aimed to measure various abilities in the socioemotional domain exist (Denham, 2006; Denham, Wyatt, Bassett, Echeverria, & Knox, 2009), the main purpose of these instruments is primarily clinical. They have not been designed for large-scale administration or to address children's readiness for kindergarten, as indicated by their progression on specific early learning standards.

In a recent review of research, funded by the Institute of Education Sciences from 2002 to 2006, Diamond et al. (2013) strongly recommended that "research identifying valid and reliable ways to measure children's skills and capture their learning over time is greatly needed," and furthermore, "there is a need to develop tools that can be readily used within everyday educational settings by teachers and other practitioners" (p. 29). Similarly, early childhood researchers associated with the Center on the Developing Child at Harvard University (2011) in their report, *Building the Brain's "Air Traffic" Control System*, stress the need for policy makers, parents, caregivers, and researchers to have access to accurate, reliable, and practical measurements of these emergent self-regulatory behaviors such that the planning of intervention programs, normal curricula, and research is based on valid, sensitive instrumentation that captures the variability in these early manifestations of approaches to learning. We heartily concur with such recommendations and suggest that researchers norm assessments for young children with diverse backgrounds in specific geographical areas as past findings (Diamond et al., 2013) indicated inconsistent results for socioemotional and related domains, such as approaches to learning.

Our goal was to create a statistically robust, easily administered assessment aimed to effectively measure the approaches-to-learning domain based on Arizona's early learning standards, which are similar to the standards adopted by many other states. The question of generalizability to populations in other states and regions, in addition to assessments derived from other state standards, remains open. The response to these questions requires further research to ascertain AtL's external validity beyond our sample, based on teacher and parent ratings. Moreover, our intent was to examine and determine whether the DECA and AtL instruments measured similar approaches-to-learning domains based on parent/guardian perceptions of children. In our next analysis, data from teachers' responses to both instruments will be examined to ascertain whether the relationship between DECA and AtL's instruments is retained.

We suggest that the AtL questionnaire, with its 13 items, serves as a more efficient and accurate tool for data collection of the approaches-to-learning domain than the DECA. However, we need to test this instrument among populations in other states with learning standards that similarly align with those from Arizona.

In addition, future research should evaluate AtL's efficacy against other instruments (i.e., Bridging assessment and PLBS) that assess children's learning behavior and to explore its generalizability in measuring the approaches-to-learning domain as an important component of school readiness. At the same time, we invite researchers and practitioners to test the AtL instrument with populations different from or similar to children in Arizona and to evaluate the psychometric properties of this promising tool in behavior and psychoeducational studies.

References

Arizona Department of Education. (2005). *Arizona early learning standards*. Phoenix, AZ: Author. Retrieved from http://www.azed.gov/wpcontent/uploads/PDF/EarlyLearningStandards.pdf

Barbu, O.C., Levine-Donnerstein, D., Marx R.W., Yaden D.B. (2012). Reliability and Validity of the Devereux Early Childhood Assessment (DECA) as a Function of Parent and Teacher Ratings. *Journal of Psychoeducational Assessment, 31*(5), 469–481.

Barbu, O.C., Marx R.W, Yaden D.B., & Levine-Donnerstein, D. (in press). Measuring Approaches to Learning in Preschoolers: Validating the Structure of an Instrument for Teachers and Parents. *Education,* 3–13.

Blair, C., Zelazo, P. D., & Greenberg, M. T. (2005). The measurement of executive functioning in early childhood. *Developmental Neuropsychology*, *28*, 561–571.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY: John Wiley.

Books, S. (2004). *Poverty and schooling in the U.S.: Contexts and consequences*. Mahwah, NJ: Lawrence Erlbaum.

Brinkman, T. M., Wigent, C. A., Tomac, R. A., Pham, A. V., & Carlson, J. S. (2007). Using the Devereux Early Childhood Assessment to identify behavioral risk and protective factors within a Head Start population. *Canadian Journal of School Psychology*, *22*(2), 136–151.

Brito, P. R. (2012). *School readiness: A conceptual frame*. New York, NY: United Nations Children's Fund.

Brito, P. R., & Lumlingan, M. C. (2012). *School readiness and transitions*. New York, NY: United Nations Children's Fund.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 445–455). Newbury Park, CA: Sage.

Buhs, E. S. (2003). Devereux Early Childhood Assessment. In B. S. Plake, J. C. Impara, & R. A. Spies (Eds.), *The fifteenth mental measurements yearbook*. Lincoln, NE: Buros Institute of Mental Measurements.

Bulotsky-Shearer, R. J., Fernandez, V., Dominguez, X., & Rouse, H. L. (2011). Behavior problems in learning activities and social interactions in Head Start classrooms and early reading, mathematics, and approaches to learning. *School Psychology Review*, *40*(1), 39–56.

Bulotsky-Shearer, R. J., Fernandez, V., & Rainelli, S. (2013). The validity of the Devereux Early Childhood Assessment for culturally and linguistically diverse Head Start children. *Early Childhood Research Quarterly*, *28*, 794–807.

Byrne, B. M. (2009). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS*. Mahwah, NJ: Lawrence Erlbaum.

Campbell, J. M., & James, C. L. (2007). Assessment of social and emotional development in preschool children. In B. A. Bracken & R. J. Nagle (Eds.), *The psychoeducational assessment of preschool children* (pp. 111–135). Mahwah, NJ: Lawrence Erlbaum.

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology*, *28*, 595–616.

Center on the Developing Child at Harvard University. (2011). *Building the brain's "air Traffic control" system: How early experiences shape the development of executive function* (Working Paper No. 11). Retrieved from www.developingchild.harvard.edu

Chain, J., Dopp, A., Smith, G., Woodland, S., & LeBuffe, P. (2010). *The Devereux Early Childhood Assessment literature review*. Retrieved from http://www.devereux.org/site/PageServer?pagename=deci_research

Chen, J. Q., Masur, A., & McNamee, G. (2010). Young children's approaches to learning: A socio-cultural approach. *Early Child Development and Care*, *181*, 1137–1152.

Chen, J. Q., & McNamee, G. D. (Eds.). (2007). *Bridging: Assessment for teaching and learning in early childhood classrooms, PreK–3*. Thousand Oaks, CA: Corwin Press.

Chittooran, M. M. (2003). Devereux Early Childhood Assessment. In B. S. Plake, J. C. Impara, & R. A. Spies (Eds.), *The fifteenth mental measurements yearbook*. Lincoln, NE: Buros Institute of Mental Measurements.

Connecticutt Early Learning and Development Standards: What children, birth to five, should know and be able to do (2014). Retrieved from www.ct.gov/oec/elds.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods, 1*, 16–29.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society*, *39*, 1–38.

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., Wyatt, T. M., Bassett, H. H., Echeverria, D., & Knox, S. S. (2009). Assessing social-emotional development in children from a longitudinal perspective. *Journal of Epidemiology and Community Health*, *63*(11), 137–152.

Devereux Foundation, DECI. (2003). *The Search Institute's 40 developmental assets in conjunction with the Devereux Early Childhood Assessment (DECA) Program*. Retrieved from http://www.devereux.org/site/DocServer/crosswalks-40Assets.doc?docID=3361

Diamond, K. E., Justice, L. M., Siegler, R. S., & Snyder, P. A. (2013). *Synthesis of IES research on early intervention and early childhood education* (NCSER 2013–3001). Washington, DC: National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long term goals. *Journal of Personality and Social Psychology*, *92*, 1087–1101.

Fantuzzo, J., Perry, M. A., & McDermott, P. (2004). Preschool approaches to learning and their relationship to other relevant classroom competencies for low-income children. *School Psychology Quarterly*, *19*(3), 212–230.

Fantuzzo, J. W., Bulotsky-Shearer, R., Fusco, R. A., & McWayne, C. (2005). An investigation of preschool classroom behavioral adjustment problems and social-emotional school readiness competencies. *Early Childhood Research Quarterly*, *20*, 259–275.

Galinsky, E. (2010). *Mind in the making: The seven essential life skills that every child needs*. New York, NY: William and Morrow.

Gehsmann, K., & Templeton, S. (2011–2012). Stages and standards in literacy: Teaching developmentally in the age of accountability. *Journal of Education*, *192*, 5–16.

Gilks, W., Richardson, S., & Spiegelhalter, D. (1995). *Markov chain Monte Carlo in practice*. New York, NY: Chapman & Hall.

Gilliam, W. S., & Frede, E. (2012). Accountability and program evaluation in early education. In R. C. Pianta, W. S. Barnett, L. M. Justice, & S. M. Sheridan (Eds.), *Handbook of early childhood education* (pp. 73–91). New York, NY: Guilford.

Graue, M. E. (1993). *Ready for what? Constructing meanings of readiness for kindergarten*. Albany: State University of New York Press.

Graue, M. E. (2006). The answer is readiness—Now what is the question? *Early Education and Development*, *17*, 43–56.

Gresham, F. M., & Elliot, S. N. (1990). *Social skills rating system*. Circle Pines, MN: Pearson.

Haskins, R., & Barnett, W. S. (Eds.). (2010). *Investing in young children: New directions in federal preschool and early childhood policy*. Washington, DC: The Center on Children and Families and the National Institute for Early Education Research.

Hau, K.-T., & Marsh, H. W. (2004). The use of item parcels in structural equation modeling: Nonnormal data and small sample sizes. *British Journal of Mathematical and Statistical Psychology*, *57*, 327–351.

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., & Yavitz, A. (2010). The rate of return to the High Scope Perry Preschool Program. *Journal of Public Economics*, *94*, 114–128.

Henderson, J., & Strain, P. S. (2009). *Screening for delays and problem behavior (Roadmap to Effective Intervention Practices)*. Tampa: University of South Florida.

Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.

Isquith, P. K., Gioia, G. A., & Espy, K. A. (2004). Executive function in preschool children: Examination through everyday behavior. *Developmental Neuropsychology*, *26*, 403–422.

Jaberg, P. E., Dixon, D. J., & Weis, G. M. (2009). Replication evidence in support of the psychometric properties of the Devereux Early Childhood Assessment. *Canadian Journal of School Psychology*, *24*(2), 158–166.

Kagan, S. L. (1990). Readiness 2000: Rethinking rhetoric and responsibility. *Phi Delta Kappan*, *72*, 272–279.

Kagan, S. L., Moore, E., & Bredekamp, S. (1995). *Reconsidering children's early learning and development: Toward shared beliefs and vocabulary*. Washington, DC: National Education Goals Panel.

Kim, J. S., & Sunderman, G. L. (2005). Measuring academic proficiency under the No Child Left Behind Act: Implications for educational equity. *Educational Researcher*, *34*, 3–13.

Kishton, J. M., & Widaman, K. F. (1994). Unidimensional versus domain representative parceling of questionnaire items:

An empirical example. *Educational and Psychological Measurement*, *54*, 757–765.

LeBuffe, P. A., & Naglieri, J. A. (1999a). *Technical manual for the Devereux Early Childhood Assessment (DECA)*. Villanova, PA: Devereux Foundation.

LeBuffe, P. A., & Naglieri, J. A. (1999b). *Devereux Early Childhood Assessment user's guide*. Lewisville, NC: Kaplan Early Learning Company Publishing.

Li-Grining, C. P., Votruba-Drzal, E., Maldonado-Carreño, C., & Haas, K. (2010). Children's early approaches to learning and academic trajectories through fifth grade. *Developmental Psychology*, *46*(5), 1062.

Lien, M. T., & Carlson, J. S. (2009). Psychometric properties of the Devereux Early Childhood Assessment in a Head Start sample. *Journal of Psychoeducational Assessment*, *27*(5), 386–396.

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*, 1198–1202.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling*, *9*(2), 151–173.

Lubke, G. H., & Muthén, B. O. (2004). Applying multi-group confirmatory factor models for continuous outcomes to Likert scale data complicates meaningful group comparisons. *Structural Equation Modeling*, *11*, 514–534.

Magnuson, K. A., Rohm, C., & Waldfogel, J. (2007). Does pre-kindergarten improve school preparation and performance? *Economics of Education Review*, *26*(1), 33–51.

McClelland, M. M., Acock, A. C., Piccinin, A., Rhea, S. A., & Stallings, M. C. (2013). Relations between preschool attention span-persistence and age 25 educational outcomes. *Early Childhood Research Quarterly*, *28*, 314–324.

McDermott, P. A., Leigh, N. M., & Perry, M. A. (2002). Development and validation of the preschool learning behaviors scale. *Psychology in the Schools*, *39*(4), 353–365.

Meisels, S. J., & Fenichel, J. (Eds.). (1996). *New visions for the developmental assessment of infants and young children*. Washington, DC: Zero to Three: National Center for Infants, Toddlers, and Families.

Meyer, L. (2008). *Fremont school readiness: Devereux Early Childhood Assessment (DECA) outcomes 2007–08*. Retrieved July 8, 2011, from http://www.echofamilycenter.com/uploads/child_outcomes/07-08/DECA07-08_School_Readiness-Final.pdf

Moffitt, M. 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. *Proceedings of the National Academy of Sciences*, *108*, 2693–2698.

Nasser, F., & Wisenbaker, J. (2003). A Monte Carlo study investigating the impact of item parceling on measures of fit in confirmatory factor analysis. *Educational and Psychological Measurement*, *63*, 729–757.

National Association for the Education of Young Children (NAEYC). (2002). *Early learning standards: Creating conditions for success*. Washington, DC: Author.

National Center for Educational Statistics. (2010). *Kindergarten through first grade parent Social Rating Scale (SRS)* (Approaches to Learning scale items). Retrieved from nces.ed.gov/pubs2010/data/2010070_atl_readme.pdf-2010-07-28.

National Education Goals Panel (NEGP). (1995). *Reconsidering children's early development and learning: Toward common views and vocabulary*. Reston, VA: Author.

Neibecker, B. (1984). The validity of computer-controlled magnitude scaling to measure emotional impact of stimuli. *Journal of Marketing Research*, *21*(3), 325–331.

Neuman, S. B., & Roskos, K. (2005). The state of pre-kindergarten standards. *Early Childhood Research Quarterly*, *20*, 125–145.

Niemeyer, J., & Scott-Little, C. (2001). *Assessing kindergarten children: A compendium of assessment instruments*. Tallahassee, FL: SERVE. Retrieved from www.serve.org/publications/rdakcc.pdf

North Carolina Division of Child Development. (2008). *Approaches to learning*. Raleigh: North Carolina Department of Health and Human Services, North Carolina Division of Child Development-Infant-Toddler Foundations. Retrieved from http://ncchildcare.nc.gov/pdf_forms/infant_toddler_foundations.pdf

Ogg, J. A., Brinkman, T. M., Dedrick, R. F., & Carlson, J. S. (2010). Factor structure and invariance across gender of the Devereux Early Childhood Assessment protective factor scale. *School Psychology Quarterly*, *25*(2), 107–118.

Pianta, R. C., Barnett, W. S., Burchinal, M., & Thornburg, K. R. (2009). The effects of preschool education: What we know, how public policy is or is not aligned with the evidence base, and what we need to know. *Psychological Science in the Public Interest*, *10*, 49–88.

Pianta, R. C., Barnett, W. S., Justice, L. M., & Sheridan, S. M. (2012). *Handbook of early childhood education*. New York, NY: Guilford.

Poropat, A. E. (2014). Other-related personality and academic performance: Evidence and implications. *Learning and Individual Differences*, *34*, 24–32.

Rolnick, A., & Grunewald, R. (2003). *Early childhood development: Economic development with a high public return*. Minneapolis: The Federal Reserve Bank of Minneapolis. Retrieved from http://www.minneapolis.org/publications_papers/pub_display.pdf

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: John Wiley.

Schmeichel, B. J., & Tang, D. (2015). Individual differences in executive functioning and their relationship to emotional processes and responses. *Current Directions in Psychological Science*, *24*, 93–98.

Scott-Little, C., Kagan, S. L., & Frelow, V. S. (2006). Conceptualization of readiness and the content of early learning standards: The intersection of policy and research? *Early Childhood Research Quarterly*, *21*, 153–173.

Snow, C. E., & Van Hemel, S. B. (2008). *Early childhood assessment: Why, what & how*? Washington, DC: National Academies Press.

Snow, K. (2006). Measuring school readiness: Conceptual and practical considerations. *Early Education and Development*, *17*, 7–41.

Snow, K. (2011). *Developing kindergarten readiness and other large-scale assessment systems: Necessary considerations in the assessment of young children*. Washington, DC: National Association for the Education of Young Children.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Allyn & Bacon.

Thigpen, B. (2014). School readiness: A challenge and opportunity. Retrieved from http://orgs.wku.edu/ttas/docs/resources/ta-paper-122008-school-readiness.pdf

Vygotsky, L. S. (2012). *Thought and language.* Cambridge, MA: MIT Press. (Original work published 1931)

Woodside-Jiron, H., & Gehsmann, K. (2009). Peeling back the layers of policy and school reform: Revealing the structural and social complexities within. *International Journal of Disability, Development, and Education*, *56*, 49–72.

Yaden, D. B., Cimetta, A., Marx, R. W., Taren, D., Perry, N., Enz, B., Swadener, B., Horn, R., & Gallagher, L. (2011, January). Arizona Kindergarten Readiness Study [Technical Report No. 1]. Tucson, AZ: University of Arizona, College of Education, First Things First External Evaluation.

Ziv, Y. (2013). Social information processing patterns, social skills, and school readiness in preschool children. *Journal of Experimental Child Psychology, 114*, 306–320.

Authors

OTILIA C. BARBU, PhD, is scientist at the SPH Analytics and research associate at the University of Arizona, College of Education. Her research interests include mathematical and statistical modeling, measurement, assessment, and evaluation in education and healthcare.

DAVID B. YADEN, Jr., PhD, is professor of language, reading, and culture in the Department of Teaching, Learning and Sociocultural Studies at the University of Arizona, College of Education. His research interests include early childhood development and assessment, the early acquisition reading and writing in multilingual environments and the application of complex adaptive systems theory to the emergence and development of literacy knowledge and ability.

DEBORAH LEVINE-DONNERSTEIN, PhD, is senior lecturer and senior researcher in the Department of Educational Psychology, College of Education, at the University of Arizona. Her research encompasses first generation undergraduates, witnessed race-ethnic prejudice models, effects of sleep among youth, and early childhood education.

RONALD W. MARX, PhD, is dean and professor of educational psychology at the College of Education at the University of Arizona, where he holds the Lindsey/Alexander Chair in Education. His work in learning and instruction has focused on science education in urban school reform, and more recently, measurement issues in early childhood education.