Title
The kindergarten Early Development Instrument predicts third grade academic proficiency

Authors
Duncan, Robert J
Duncan, Greg J
Stanley, Lisa
et al.

Publication Date
2020

DOI
10.1016/j.ecresq.2020.05.009

Peer reviewed

Early Childhood Research Quarterly 53 (2020) 287–300

Contents lists available at ScienceDirect

Early Childhood Research Quarterly

The kindergarten Early Development Instrument predicts third grade academic proficiency

Robert J. Duncan[a,*], Greg J. Duncan[b], Lisa Stanley[c], Efren Aguilar[c], Neal Halfon[c,d,e]

[a] Human Development and Family Studies, Purdue University, United States
[b] School of Education, University of California, Irvine, United States
[c] Department of Pediatrics, Geffen School of Medicine, University of California, Los Angeles, United States
[d] Department of Health Policy and Management, Fielding School of Public Health, University of California, Los Angeles, United States
[e] Department of Public Policy, Luskin School of Public Affairs, Center for Healthier Children Families and Communities, University of California, Los Angeles, United States

ARTICLE INFO

Article history:
Received 24 October 2018
Received in revised form
10 December 2019
Accepted 18 May 2020

Keywords:
School readiness
Achievement proficiency
Early Development Instrument

ABSTRACT

School readiness skills predict later educational achievement, health, and social-emotional outcomes. Measures of school readiness can provide valuable information to assess both the impact of strategies and policies that prepare children for school as well as informing strategies for improving children's educational trajectories across their school years. The Early Development Instrument (EDI) is a measure of school readiness skills based on teacher-reported observational recall. It has been used extensively in Canada and Australia and is in the early stages of adoption in a number of U.S. cities. The current study uses data from roughly 3000 children followed longitudinally from kindergarten through third grade from 7 school districts in Orange County, California. The study assesses whether EDI ratings in kindergarten predict third grade proficiency in mathematics and English Language Arts on state assessments. Ratings on the EDI were strongly associated with proficiency in both academic areas, even in the presence of controls for child-level factors and neighborhood fixed effects. Among its components, ratings on the language and cognitive development, communication skills and general knowledge, and social competence domains strongly differentiated children's likelihood of later proficiency in both academic areas. Implications for improving comprehensive early childhood education and schooling policies based on indicators of school readiness are discussed.

1. Introduction

Early school performance has been linked to subsequent education attainment (Duncan et al., 2007), which, in turn, predicts adult social and health outcomes (Cunha & Heckman, 2007). Although early mathematics skills are one of the strongest predictors of later achievement, literacy, communication, social-emotional (e.g., self-control), and physical skills (e.g., fine motor abilities) are predictive of later achievement as well (e.g., Blair & Raver, 2015; Cameron et al., 2012; Davies, Janus, Duku, & Gaskin, 2016; Duncan et al., 2007; Grissmer, Grimm, Aiyer, Murrah, & Steele, 2010). Improving children's cognitive, physical, and socio-emotional skills around the point of school entry is a key objective of many early childhood education programs (e.g., Head Start). However, despite general agreement on the importance of an inclusive set of school readiness skills (e.g., Ackerman & Barnett, 2005), there is little consensus on valid teacher-reported measures of these skills for populations of young children that predict later performance on school success as indicated by scores on high-stakes standardized tests.

The current study evaluates the predictive utility of the Early Development Instrument (EDI), a tool designed to provide information on five key domains of children's development – physical health and well-being, social competence, emotional maturity, language and cognitive development, and communication skills and general knowledge (Davies et al., 2016; Janus & Offord, 2007). Predictive utility is examined by estimating associations between child-level EDI scores reported by kindergarten teachers and scores on standardized tests administered in third grade. Although these associations speak to only a portion of the overall validity of the EDI, extensive psychometric work has already gone into EDI instrument development (e.g., Brinkman et al., 2013). However, none of this past work has examined the predictive validity of the instrument in U.S. children or fully unpacked its association with later perfor-

* Corresponding author at: Purdue University, 1202 W. State St., Rm 247, West Lafayette, IN 47907-2055, United States.
E-mail address: duncan99@purdue.edu (R.J. Duncan).

https://doi.org/10.1016/j.ecresq.2020.05.009
0885-2006/

mance on high-stakes standardized tests (a key indicator of success for many schools). Thus, our work fills an important gap in the literature on the EDI for researchers, policymakers, and practitioners, especially in the United States.

The EDI was developed to be a population-level tool because it can be used across entire communities, cities, and counties to assess how community-level risk and protective factors influence kindergarten readiness. Thus, it is intended to help community partners and stakeholders understand which of their strategies are working and where additional supports are most needed. It was not intended or developed to be a diagnostic measure of individual children's abilities. However, despite a growing use of the EDI in a number of U.S. cities, school districts, and communities, no research studies have linked it with educational outcomes in a U.S. sample.

The availability of EDI data across a number of school districts in southern California enabled us to engage in a careful examination of the ability of the EDI administered during the kindergarten year to predict key school administrative outcomes measured in third grade. Our work can also be viewed as part of broader efforts to understand how teacher-rated assessments relate to later educational outcomes. In a diverse sample of U.S. children followed longitudinally, we examine whether ratings based on teacher responses on the EDI during kindergarten predict third grade proficiency on state standardized assessments of mathematics and English Language Arts.

Although the EDI was designed to be used for population-level purposes, the current study examines its predictive validity by adopting a more traditional developmental framework of linking individual children's early skills to later achievement outcomes (e.g., Duncan et al., 2007). Our approach aligns with the *skills beget skills* theoretical perspective on human capital (Heckman, 2000). In essence, our study assesses to what extent the skills children have in kindergarten, as assessed by EDI-based teacher reports, predict their likelihood of later academic proficiency (and overall performance), with and without controlling for other possible explanations within the limits of the data. Although we have EDI and third-grade test scores for nearly all kindergarten students attending schools in a number of school districts, our ability to estimate population-level associations is limited by the fact that children were not randomly selected from the aggregate levels (e.g., neighborhoods, schools) at either study entry or at the point of standardized testing. Skills assessed by the EDI are thought to be useful targets for early intervention because they are skills that children need to be successful in schooling contexts (Ackerman & Barnett, 2005).

1.1. School readiness and later achievement

Growing recognition of the importance of children's school readiness has led to a wide range of efforts to measure and intervene to improve these skills. Assessments that measure children's skills can be used in a variety of ways and for a variety of purposes. For example, assessments could be used to consider the supports needed for specific children who enter school with less developed skills than their peers or for understanding populations of children who are the most vulnerable (e.g., based on geographic disparities). Although information on the EDI is specific to the child, when collected at the population level it can be used to inform how well communities (i.e., neighborhoods, cities, school districts, states) are doing in supporting children's overall development. A key advantage to focusing on populations is that communities can collectively consider what investments may help promote children's school readiness, rather than only emphasizing individual level interventions targeting an individual child's skills. This might include a city level strategy of providing early literacy programs in neighborhood libraries, or the provision of mommy and me groups designed to address parental social isolation and enhance social emotional development. The EDI has been extensively used in Canada and Australia for these purposes (D'Angiulli, Warburton, Dahinten, & Hertzman, 2009; Davies et al., 2016; Forget-Dubois et al., 2007; Guhn, Gadermann, Almas, Schonert-Reichl, & Hertzman, 2016; Janus, Brinkman, & Duku, 2011; Lloyd, Irwin, & Hertzman, 2009).

Community-based research can identify and address the multifaceted ways that disparities develop by engaging multiple stakeholders in the process (Israel, Schulz, Parker, & Becker, 1998). Thus, population level indicators of school readiness can provide opportunities for communities to target interventions, fill in services gaps, and gauge the effectiveness of new programs and policies that attempt to improve children's early wellbeing. For example, increased access to high quality childcare can promote school readiness skills and improve the conditions for young children in their community (Campbell, Pungello, Miller-Johnson, Burchinal, & Ramey, 2001). As place-based and community-wide early childhood initiatives spread, early childhood stakeholders are using indicators of school readiness to assess how well strategies and policies are preparing children for school (e.g., Goldfeld et al., 2016; Guthridge et al., 2016; Patel, Corter, Pelletier, & Bertrand, 2016). Similarly, schools and school districts use indicators of school readiness to anticipate and plan for the supports required to enhance educational trajectories as children progress through the school years. For example, understanding early factors that are associated with later academic risk can allow for school administrators to consider which interventions are most likely to be effective for their student population (i.e., providing additional resources for where there is the greatest need). Assessing school readiness at the population level may be most effective for better understanding these multifaceted and ongoing efforts (i.e., trends in school readiness data based on place and time).

In the United States, there is no single, commonly-adopted measure of school readiness that can be used for comparisons across communities and over time. Additionally, there is little consensus on the best content or method for collecting population-level school readiness data. Instead, a variety of tools have been developed which vary in purpose, functional skill areas assessed, processes used to collect information, and levels of aggregation reported. This study informs these ongoing efforts by assessing how well the EDI, when administered in kindergarten, predicts academic proficiency in third grade, a key outcome to school administrators and practitioners.

1.2. Teacher-reported measures of school readiness

Given the limitations of direct child assessments (e.g., training data collectors and/or coders, time, costs) for measuring entire populations of children, some researchers and practitioners have turned to teacher-reported measure of children's abilities on school readiness domains. Many teacher-reported scales exist, ranging from the more brief and targeted assessment (e.g., Child Behavior Rating Scale; Schmitt, Pratt, & McClelland, 2014), to the broader tools designed to track ongoing development in a number of domains (e.g., Teaching Strategies GOLD; Lambert, Kim, & Burts, 2015).

Issues with teacher-rated measures often revolve around their ability to provide unbiased measures of children's skills. For example, teacher-reported data can suffer from biases due to the subjectivity of different raters (Waterman, McDermott, Fantuzzo, & Gadsden, 2011). The benefits of teacher-rated assessments relative to direct assessments would also be undermined if they do not relate to later achievement. The current study contributes to a growing body of evidence evaluating whether teacher-ratings around the point of school entry relate to direct assessments of child performance (e.g., Lambert et al., 2015; Schmitt et al.,

2014). The current study extends this prior research by examining how a teacher-rated measure in kindergarten relates to high-stakes standardized testing three years later rather than researcher administered direct assessments.

1.2.1. The Early Development Instrument (EDI)

The EDI is designed for young children ages four to six years (Janus & Offord, 2007). Kindergarten teachers complete the EDI based on observational recall for each child in their classroom. In the United States, the EDI is most typically implemented in kindergarten because it is usually the first point to universally access children to measure school readiness. For this reason, the EDI is often implemented by all schools serving kindergarten-aged children living in a specified target geographical area. Based on children's home address information, children's EDI data are geo-coded to the census tract of their home residence. In this way, the EDI data can be used to reflect differences in children's wellbeing and school readiness at multiple levels (e.g., child, neighborhood, schools). To remove concerns of using the EDI as a diagnostic measure, EDI scores are typically only reported at aggregate levels.

The EDI comprises 103 core questions grouped into five developmental domains: 1) physical health and well-being, 2) social competence, 3) emotional maturity, 4) language and cognitive development, and 5) communication skills and general knowledge (Janus & Offord, 2007). These five domains are further divided into 16 subdomains. After receiving a one-hour orientation, kindergarten teachers complete the EDI using an online data entry portal on each child in the classroom. Data are typically collected no earlier than three months after the start of the school year to ensure that teachers have ample time to get to know their students. School districts most typically elect to collect EDI in the winter (sometime between January and March).

The current study builds off prior EDI work in Canada, Australia, and other nations that have examined its psychometric validity and reliability and how it relates to educational outcomes (e.g., Brinkman et al., 2007; Forer & Zumbo, 2011; Guhn, Gadermann, & Zumbo, 2007; Hymel, LeMare, & McKee, 2011; Janus & Offord, 2007). Much of the previous work on the EDI has examined its psychometric properties and validity, contributing to its development and rationale for its use in the current study. For instance, the original measurement development study found that, for over 16,000 kindergarteners in Canada, the five domains used in the current study best fit the data (Janus & Offord, 2007).

In the time since the original measurement development, a number of studies have replicated and extended the psychometric work. Work using Rasch analyses have supported the original five domains identified by the developers (Curtin, Browne, Staines, & Perry, 2016). Additional studies have found support for between group reliability (Guhn et al., 2007), construct validity (Forer & Zumbo, 2011; Hymel et al., 2011; Janus, Zeraatkar, Duku, & Bennett, 2018), cross-cultural validity (Brinkman et al., 2007; Janus et al., 2011), and internal consistency (Curtin, Madden, Staines, & Perry, 2013). Building on the extensive work on the psychometric properties, we focus our study on its use as a tool for predicting later educational outcomes.

Prior work has focused on the EDI has a predictor of educational outcomes in a number of countries and with different analytic techniques than the ones used in the current study. For instance, the EDI scores were found to explain 34% of children's school achievement a year later in first grade (i.e., in Canada; Forget-Dubois et al., 2007), third and fourth grade academic achievement in Canada (Davies et al., 2016; Guhn et al., 2016), and found to have equally strong associations with children's test scores at third, fifth, and seventh grade in Australia (Brinkman et al., 2013). The current study builds on this international program of research by examining the associations between the EDI and children's standardized testing performance in third grade in a U.S. population. Importantly, the current study evaluates these associations with several types of models (linear and logistic), multiple types of clustering (neighborhood and classroom), and multiple ways of coding both the EDI and achievement scores (categorical and continuously).

Since 1998, the EDI has been used in Canada across many jurisdictions in different provinces, it is the national indicator of child wellbeing for all of Australia (Goldfeld, Sayers, Brinkman, Silburn, & Oberklaid, 2009), and it has been used in over 25 other countries. In the United States, the use of the EDI has grown from an initial pilot site in 2009 in Southern California to now over 70 communities in 16 states. However, no research study to date has linked the EDI with later academic outcomes in a U.S. population of children. Because the EDI is a multidimensional measure of healthy child development, it is being used to engage a broad cross-sector group of organizations in multidisciplinary strategies to improve children's developmental outcomes. This has included cross-sector strategies to engage health, K-12 education, early childhood education, and family and community support sectors in what are now commonly referred to as collective impact and Cradle to Career initiatives designed to improve school readiness and health and wellbeing over the life course (Janus, 2013; Jutte, Miller, & Erickson, 2015).

1.3. Current study

The current study represents the first examination of whether the EDI has predictive utility for assessing academic proficiency in a U.S. population of children. This study addresses whether the EDI ratings in kindergarten predict reaching proficiency standards on third grade mathematics and English Language Arts in a large, diverse sample of children in Orange County, California. In particular, the study asks to what extent each of the five domains of the EDI (i.e., physical health and well-being, social competence, emotional maturity, language and cognitive development, and communication skills and general knowledge) relate to third grade academic proficiency? Based on prior research on the EDI as well as research on school readiness (e.g., Brinkman et al., 2013; Duncan et al., 2007), we hypothesize the language and cognitive development domain will be most closely related to academic proficiency, though all domains will significantly relate to varying degrees. The current study controls for multiple characteristics of the child and adjusts for neighborhood characteristics that could account for the observed associations between school readiness as measured by the EDI and academic proficiency.

2. Method

2.1. Procedure

The study sample comes from school districts in Orange County, California that collected EDI data on their 2011–2012 school year kindergarten cohort. This cohort was part of a larger, multi-year effort to collect EDI data from all school districts in the county. The 2011–2012 cohort was the first to receive the new state-wide testing instruments – the Smarter Balanced Assessments (SBA) – during the child's third grade academic year in 2014–2015. Of the 11 school districts that conducted the EDI in Orange County during the 2011–2012 year, 7 provided administrative records that allowed for linking the EDI data with SBA performance in 2014–2015.

2.1.1. Collecting EDI data

At least three months after the start of the school year, teachers received a one-hour training at their respective schools. Trainings used a standardized set of materials including a Teacher Guide that offered supplemental information on how to interpret each EDI

question and how to access and use the online data entry system. For students transferring in to the classroom mid-year, teachers were instructed to exclude them from the EDI if they had been in their class for less than one month. Teachers were also instructed to exclude children whose parents opted out of the EDI.

2.2. Participants

The total resulting sample size for the study was 2976 children (54% male). At the time of the EDI data collection in kindergarten, the average age was 5.79 years (*SD* = 0.30). The sample represents the diversity of students in the participating school districts, though differs from the county overall, as the sample had a higher proportion of Hispanic/Latinx children and English Language Learners (ELLs). Specifically, the sample was 74% Hispanic/Latinx, 11% White, and 15% other; and 67% of kindergarteners were considered ELLs (county average is 39%). Two particularly large school districts contributed to 62% of the study sample (i.e., 33% and 29%), with the other five school districts contributing 5% to 10% of the total sample.

2.3. Measures

2.3.1. The Early Development Instrument (EDI)

The independent variables of interest were the five EDI domains: *physical health and well-being, social competence, emotional maturity, language and cognitive development*, and *communication skills and general knowledge*. The EDI has undergone extensive validation work, with the current version of the EDI and its items made available by the developers at https://edi.offordcentre.com/about/what-is-the-edi/.

For each child's EDI record, an average score on each of the five domains was calculated by adding up the scores for all of the core items in that domain and dividing by the total number of core items comprising the domain. This average score then allowed each record to be compared with the normative sample cutoffs, specifically "vulnerable," "at risk," "on track middle," and "on track top" cutoffs. The normative cutoffs were determined using an independent convenience sample of children from school districts in the United States that collected EDI data in the 2009–2010 school year (*N* = 10,244). To establish these cutoffs, an average score for each domain was first developed per child with data valid for analysis. Using the averages for all valid records, the 10th, the 25th, and 75th percentile cutoff scores were determined for each domain.

Children were categorized as "vulnerable" in a domain if the mean score of their EDI items for that domain fell at or below the 10th percentile population cutoff. Children were categorized as "at risk (for becoming vulnerable)" in a domain if the mean of their EDI items for that domain was above the 10th percentile cutoff but fell at or below the 25th percentile cutoff. Children were categorized as "on track middle" in a domain if the mean of their EDI items for that domain fell above the 25th percentile cutoff but fell at or below the 75th percentile. Children were categorized as "on track top" if they fell above the 75th percentile. These normed reference categorical scores were used with community partners and in the current study for two reasons: 1) in order to make the data more translatable than the continuous scores, and 2) because continuous scores were highly skewed and bunched towards the top of the distribution. The linear and logistic regression models were run with neighborhood fixed effects and the EDI domains separately and together with continuous EDI predictors instead of categorical in additional analyses to the ones presented. Although not the focus of the current study, these model results are available by the first author by request.

2.3.2. Physical health and well-being

Three subdomains comprised the physical health and well-being domain of the EDI: 1) physical readiness for school work (4 items), 2) physical independence (4 items), and 3) gross and fine motor skills (5 items). Two different types of responses were used depending on the items: 1) 'very good/good', 'average', or 'poor/very poor', or 2) 'yes' or 'no'. An example item is: the child has 'proficiency at holding a pen, crayons, or a paintbrush' and 'is well coordinated (i.e., moves without running into or tripping over things)'. The alpha for this scale was 0.78 in our sample, which is consistent with other work finding an alpha of 0.84 for this scale (Janus & Offord, 2007).

2.3.3. Social competence

Four subdomains made up social competence domain of the EDI: 1) overall social competence with peers (5 items), 2) respect and responsibility (8 items), 3) approaches to learning (9 items), and 4) readiness to explore new things (4 items). Responses were 'often or very true', 'sometimes or somewhat true', 'never or not true', or 'don't know' to the items. Examples items are: the child . . . 'listens attentively', 'demonstrates respect for adults', 'works independently', and 'is able to follow one-step instructions'. The alpha for this scale was 0.97 in our sample, which is consistent with other work finding an alpha of 0.96 for this scale (Janus & Offord, 2007).

2.3.4. Emotional maturity

Four subdomains made up emotional maturity domain of the EDI: 1) prosocial and helping behavior (8 items), 2) anxious and fearful behavior (8 items), 3) aggressive behavior (7 items), and 4) hyperactive and inattentive behavior (7 items). Two different types of responses were used depending on the items: 1) 'very good/good', 'average', or 'poor/very poor', or 2) 'often or very true', 'sometimes or somewhat true', 'never or not true', or 'don't know'. Examples items are: the child . . . 'helps other children who are feeling sick', 'kicks, bites, hits other children or adults', 'has temper tantrums', and 'bullies or is mean to others'. The alpha for this scale was 0.92 in our sample, which is consistent with other work finding an alpha of 0.92 for this scale (Janus & Offord, 2007).

2.3.5. Language and cognitive development

Four subdomains made up the language and cognitive development domain of the EDI: 1) basic literacy skills (8 items), 2) basic numeracy skills (7 items), 3) interest in literacy/numeracy and memory (5 items), and 4) advanced literacy skills (6 items). Responses were yes/no to the items. Example items are: the child . . . 'is generally interested in books (pictures and print)', 'is able to write simple words', 'is interested in mathematics', and 'is able to count to 20'. The alpha for this scale was 0.91 in our sample, which is consistent with other work finding an alpha of 0.93 for this scale (Janus & Offord, 2007).

2.3.6. Communication skills and general knowledge

Only one subdomain comprised the communication skills and general knowledge domain of the EDI (8 items). Responses were 'very good/good', 'average', or 'poor/very poor' to items. Example items are: child has the . . . 'ability to understand on first try what is being said to him/her', 'ability to listen', 'ability to tell a story', and 'ability to take part in imaginative play'. The alpha for this scale was 0.94 in our sample, which is consistent with other work finding an alpha of 0.95 for this scale (Janus & Offord, 2007).

2.3.7. Smarter Balanced Assessments

The dependent variables of interest were measures of third grade academic performance from the California Smarter Balanced Assessments (SBA). The SBA mathematics in third grade assesses four components, which were developed in alignment

with the common core standards (Smarter Balanced Assessment Consortium, 2016). These four components (called claims in the manual) include: concepts and procedures, problem solving, communicating reasoning, and modeling and data analysis. Detailed information on the SBA mathematics assessment is available at http://www.smarterbalanced.org/wp-content/uploads/2015/08/Mathematics-Content-Specifications.pdf. The SBA English Language Arts in third grade also assesses four components, which were developed in alignment with the common core standards. These four components include: reading, writing, speaking/listening, and research. Detailed information on the SBA English Language Arts assessment is available at https://portal.smarterbalanced.org/library/en/english-language-artsliteracy-content-specifications.pdf.

The current study used the mathematics and English Language Arts assessments of the SBA to measure academic proficiency/risk. The continuous scores were coded into four levels based on the SBA instrument norms: Level 4 (standard exceeded), Level 3 (standard met), Level 2 (standard nearly met), and Level 1 (standard not met). For the purposes of this study, we recoded the four levels into two categories: Levels 2–4 (proficient) and Level 1 (academic risk, or not proficient). Thus, our distinction is primarily between children who are most at risk for academic difficulties and all other students. This distinction is considered to be of high interest because children at the lowest levels of achievement in third grade are the most likely to incur future schooling costs due to grade retention or additional educational services, or higher societal costs through failure to complete High School. For the mathematics assessment of the SBA in our data (N = 2975), 64.1% of children were coded as proficient. For the English Language Arts assessment of the SBA (N = 2977), 57.6% of children were coded as proficient. Although not the primary focus of the current study, models using the continuous outcomes are reported as well (for mathematics, N = 2,418, M = 2403.8, SD = 77.9; for ELA, N = 2,419, M = 2381.4, SD = 84.2). We focus primarily on the dichotomized academic proficiency/risk in our results and discussion because of their high interest to community partners and school administrators.

2.3.8. Covariates

To minimize endogeneity bias, data on our control variables were drawn from the time EDI data were collected (kindergarten) rather than from administrative records from third grade. Available control variables included child age (grand mean centered in analyses), gender, race/ethnicity, and ELL status. No information on the mothers and families were available in our administrative data.

In order to account for potential associations due to socio-economic status in the results we controlled for a child's kindergarten neighborhood of residence in the analyses. The geographic boundaries of the neighborhoods were developed in facilitated community meetings incorporating feedback from residents and community leaders across multiple sectors (e.g., Health, Education, Law Enforcement). Using U.S. census tracts as the smallest geographic units of analysis, these facilitated discussions were designed to draw on local expertise to determine how to cluster census tracts into community defined neighborhoods. This helped to ensure that neighborhoods used in this analysis reflected: 1) distinct geographic areas that residents, community leaders, and local policy makers recognize as a neighborhood; 2) a shared resident experience in terms of some level of social interaction and shared institutions; and 3) a large enough number of EDI records to provide statistical stability while also being small enough to detect variability in child development and socio-economic factors. Given these considerations, there is a high expectation that these neighborhoods are indicative of a child's socio-economic status (Krieger, Chen, Waterman, Rehkopf, & Subramanian, 2003; Morrissey & Vinopal, 2018).

2.4. Analytic strategy

All analyses were conducted in Stata 15.1 (StataCorp, 2017). Our primary analyses are based on linear regression models that control for time-invariant neighborhood-level factors by using neighborhood fixed effects (i.e., which essentially transforms all dependent and independent variables in the analysis into deviations from their neighborhood-specific mean values). Neighborhood fixed effects insures that all estimates are based on within-neighborhood variation, which effectively controls for persistent factors involved in neighborhood selection. Our data included 103 neighborhoods in which the number of children ranged from 1 to 206, with an average size of 29. The neighborhood intraclass correlations (ICC) were 0.19 for SBA Mathematics proficiency and 0.19 for SBA English Language Arts proficiency. These were calculated using Stata's multilevel logistic regression command and requesting the corresponding ICC for each outcome.

In our primary models, each outcome was examined independently and each EDI domain was entered one at a time. Our rationale for this approach was that the categorical codes on the domains were highly collinear and school administrators and community partners often times consider each domain on its own, rather than the association of one while holding all others constant.

In order to guard against misinterpretation of the data due to limitations of examining dichotomized outcomes, secondary models reported in the paper include the same approach but predict continuous SBA achievement scores rather than the dichotomized ones. Although we used a conceptually meaningful dichotomization in the outcomes (i.e., Level 1 [not proficient] versus Levels 2–4 [nearly proficient or better]), our results could be biased if the EDI is only differentiating children at the bottom of SBA performance. The continuous outcomes avoid this limitation by showing how the EDI relates to achievement across the entire range of scores.

We run a number of additional models that are included in the appendix. We report on the primary models with all EDI domains entered simultaneously because this is a more common approach taken in research (i.e., holding constant the effects of the other domains). We also report on the linear regression models that control for time-invariant kindergarten classroom-level factors by using a classroom fixed effects approach (ICC was 0.17 for SBA Mathematics and 0.19 for SBA English Language Arts) and logistic regression models with both the neighborhood and kindergarten classroom fixed effect approach. Our data included 309 kindergarten classrooms in which teachers reported on between 1 and 30 children, with an average size of 9. Considering kindergarten teachers were rating children on the EDI, using kindergarten classroom fixed effects focuses on variation across students for each teacher and thus adjusts for any biases in the ratings arising from between-teacher difference in how they understood their rating instructions and how they viewed the response scales. Logistic regression models are reported in the appendix with the corresponding odds ratios because linear regression models can predict values outside the range of dichotomous outcomes. Comparing the model results of the linear versus logistic regression, and models with neighborhood versus classroom fixed effects demonstrated no appreciable differences in the conclusions (discussed below).

All of the models included child gender, child race/ethnicity, ELL status, and child age (grand mean centered) at the time of the EDI assessment as control variables. The regression models estimated the predicted proportion proficient on SBA mathematics and English Language Arts holding constant all demographic variables. Age was grand mean centered in all analyses to aid in interpretation of other effects. The reference group for the categorical variables in

all models was: White, non-ELL, and male (i.e., this is the group for the predicted probabilities presented in the figures, but other groups could be easily estimated by adding or subtracting the coefficient for that variable from those predicted values).

It should be noted our models were run with listwise deletion (primarily due to a loss of observations when linking administrative data of the state standardized testing). A model with multiple imputation data was run for the primary models, though results did not differ in any substantial or meaningful way. These results are available by request from the first author.

3. Results

Teachers' ratings of their students' development on the five domains of the EDI are reported in Table 1. For each domain, the majority of the sample was in the 'on track middle' classification, generally followed by 'on track top', 'at risk', and then 'vulnerable'. The last two columns of Table 1 show the chi-squared test statistic for whether EDI ratings were associated with proficiency on mathematics or English Language Arts. In all cases, the null hypothesis of no relation between the variables was rejected. The full breakdown of all EDI domains and levels with academic proficiency can be found in Table 2. All domains were observed to have increases in the percentage proficient for each stepwise progression from 'vulnerable' to 'on track top,' with the most notable increases observed for the language and cognitive development domain.

The Spearman's correlations between the EDI domains and academic proficiency are reported in Table 3. All correlations were statistically significant between the EDI domains, with the strongest correlation between the social competence and emotional maturity domains ($r = 0.67$, $p < .001$). Although all correlations were statistically significant, the domains appear to be tapping different constructs as correlations among the EDI domains ranged from .37 to .67. In terms of academic proficiency, the language and cognitive development domain was most correlated with both mathematics (r = 0.41, p < .001) and English Language Arts (r = 0.43, p < .001), although all EDI domains were significantly correlated with both academic proficiency variables. It should be noted that these bivariate associations for the overall sample could be more reflective of the true relations if our primary models are over-controlling for proxies of determinants for children's outcomes (e.g., socio-economic status relates to later achievement due to its association with earlier skills).

3.1. Does the EDI predict third grade proficiency in mathematics and English Language Arts?

The top portion of Table 4 reports on the effects for each EDI domain predicting SBA mathematics proficiency and the bottom portion reports on the effects for each EDI domain predicting SBA English Language Arts. Of primary interest was how the five EDI domains related to proficiency in both academic areas. In general, each stepwise progression (i.e., vulnerable to at risk; at risk to on track middle; and, on track middle to on track top) was associated with statistically significant increases in the likelihood of third grade proficiency. Across all five EDI domains, the only comparisons that were not statistically significant were between vulnerable and at risk on the emotional maturity domain predicting third grade SBA mathematics and SBA English Language Arts proficiency. Across all models, the most within-neighborhood variance was always explained by the language and cognitive development domain, followed by the communication and general knowledge domain, social competence domain, and then the emotional maturity and physical health and well-being domains. We compared across models using the within neighborhood variance explained and did not directly compare specific coefficients across models because within each of the five models there were six possible contrasts (i.e., vulnerable to at risk; vulnerable to on track middle; vulnerable to on track top; at risk to on track middle; at risk to on track top; and on track middle to on track top). Thus, for the five models (for each outcome), there would be many potential contrasts that could be compared in total if comparing the coefficients as opposed to the within neighborhood variance explained.

In general, the pattern of results indicate that children coded as on track top in any EDI domain were likely to reach proficiency in both academic outcomes, whereas children coded as vulnerable were relatively unlikely to reach proficiency. Fig. 1 shows this pattern of results for each EDI domain and rating for predicting SBA mathematics proficiency. Holding all other covariates constant and adjusting for neighborhood effects, roughly 26% of children coded as vulnerable reached proficiency compared with 90% coded as on track top for the language and cognitive development domain. The communication and general knowledge domain showed a similar pattern overall, but a slightly weaker contrast (32% compared with 84%). Children coded as on track top on the social competence domain were particularly likely to reach proficiency compared with children coded as vulnerable (i.e., 93% versus 43%).

The results for SBA English Language Arts are shown in Fig. 2 and are very similar to the SBA mathematics pattern of results. Holding all other covariates constant, children coded as vulnerable on the language and cognitive development domain were relatively unlikely to be proficient, 24%, compared with children coded as on track top, 82%. Again, a similar pattern to language and cognitive development was observed for communication skills and general knowledge, though a slightly weaker contrast (26% versus 76%). As with mathematics, children coded as on track top in social competence were particularly likely to reach proficiency in SBA English Language Arts at 84%, compared with 38% for children coded as vulnerable.

3.1.1. Covariates

The largest differences in proficiency outcomes were due to the EDI ratings and not to any control variable. Holding constant kindergarten EDI scores, control variables were weakly associated with proficiency outcomes. For example, only for SBA English Language Arts were covariates statistically significant for each model. Specifically, females did better than males (bs ranges from 0.05 to 0.08, ps < .01) and children coded as white did better than children coded as Hispanic/Latinx (bs ranges from -0.07 to -0.10, ps < .05). We consider these contrasts as relatively weak compared with the bs associated with the changes in the EDI levels across domains.

3.1.2. Continuous SBA achievement outcomes

In Table 5, we report on the linear regression models for the continuous achievement outcomes that included all EDI domains entered independently with demographic control variables and neighborhood fixed effects. All significant effects for the EDI domains were consistent with the models that had the dichotomous proficiency outcomes. Additionally, the language and cognitive development domain and the communication skills and general knowledge domain tended to best differentiate children's continuous achievement scores (based on the within-neighborhood R^2 values).

3.1.3. Logistic regression models

The results for the logistic regression model (see Appendix Table A1) are similar to the results reported from the primary models that used linear regression in terms of significant associations and most closely related constructs. For SBA mathematics, children coded as on track top compared with vulnerable in the language and cognitive development domain had 34.15 (95% confidence interval:

Table 1
Vulnerability classifications for the five EDI domains in kindergarten and their associations with third grade proficiency.

	N	Vulnerable	At Risk	On Track Mid	On Track Top	$\chi^2(3)$ Math	$\chi^2(3)$ ELA
Physical Health & Well-Being	2975	8.27%	14.89%	44.67%	32.17%	181.87***	207.53***
Social Competence	2976	9.64%	14.62%	48.49%	27.25%	318.85***	317.81***
Emotional Maturity	2939	7.66%	12.49%	49.37%	30.49%	191.70***	189.96***
Lang. & Cog. Develop.	2976	9.41%	22.11%	48.12%	20.36%	546.34***	547.39***
General Know. & Comm.	2976	12.53%	20.06%	40.69%	26.71%	443.75***	435.64***

Note. Chi-square tests show the bivariate association between EDI domain and the SBA Mathematics or English Language Arts (ELA) outcomes with 3 degrees of freedom.

Table 2
Breakdown of proficiency on math and ELA by each EDI domain and classification.

	Vulnerable		At Risk		On Track Mid		On Track Top	
	Math	*ELA*	*Math*	*ELA*	*Math*	*ELA*	*Math*	*ELA*
Physical Health & Well-Being	40.24%	33.74%	52.49%	45.15%	61.78%	53.95%	78.97%	74.71%
Social Competence	35.19%	28.57%	45.52%	38.62%	63.94%	57.03%	84.69%	79.16%
Emotional Maturity	40.44%	34.67%	43.87%	36.51%	64.39%	57.55%	77.68%	71.88%
Lang. & Cog. Develop.	18.57%	15.71%	44.68%	35.11%	71.96%	64.80%	87.79%	84.49%
General Know. & Comm.	28.95%	21.72%	45.39%	39.20%	71.98%	64.57%	82.75%	77.74%

Note. ELA is English Language Arts.

Table 3
Spearman's correlations between EDI domains and academic proficiency.

	1.	2.	3.	4.	5.	6.
1 Physical Health & Well-Being						
2 Social Competence	.52***					
3 Emotional Maturity	.43***	.67***				
4 Lang. & Cog. Develop.	.37***	.48***	.37***			
5 General Know. & Comm.	.48***	.61***	.50***	.55***		
6 Math	.24***	.33***	.25***	.41***	.37***	
7 ELA	.26***	.33***	.25***	.43***	.37***	.60***

Note. ELA is English Language Arts.

Table 4
Results from neighborhood fixed-effects regression models predicting 3rd grade math and ELA proficiency with the kindergarten EDI and control variables (Ns = 2905 - 2944).

	3rd Grade Math				
	Lang. & Cog. Develop.	General Know. & Comm.	Physical Health & Well-Being	Social Competence	Emotional Maturity
	B (SE)	*B (SE)*	*B (SE)*	*B (SE)*	*B (SE)*
Vulnerable	Reference	Reference	Reference	Reference	Reference
At Risk	0.25 (0.03)***	0.17 (0.03)***	0.14 (0.04)***	0.13 (0.03)***	0.04 (0.04)
On Track Mid	0.51 (0.03)***	0.41 (0.03)***	0.21 (0.03)***	0.30 (0.03)***	0.25 (0.03)***
On Track Top	0.64 (0.03)***	0.52 (0.03)***	0.36 (0.03)***	0.50 (0.03)***	0.38 (0.03)***
ELL	−0.01 (0.02)	0.04 (0.02)	−0.04 (0.02)	−0.02 (0.02)	−0.02 (0.02)
Hispanic	−0.05 (0.03)	−0.05 (0.04)	−0.08 (0.04)*	−0.08 (0.04)*	−0.08 (0.04)*
Other	0.03 (0.03)	0.07 (0.03)*	0.05 (0.04)	0.06 (0.04)	0.08 (0.04)*
Female	−0.03 (0.02)	−0.04 (0.02)**	−0.02 (0.02)	−0.06 (0.02)***	−0.06 (0.02)***
Age	0.01 (0.03)	0.04 (0.03)	0.07 (0.03)*	0.05 (0.03)	0.08 (0.03)**
Intercept	0.26 (0.04)***	0.32 (0.04)***	0.50 (0.04)***	0.43 (0.04)***	0.48 (0.04)***
Within R^2	0.17	0.15	0.07	0.12	0.08
	3rd Grade ELA				
Vulnerable	reference	reference	reference	reference	reference
At Risk	0.18 (0.03)***	0.18 (0.03)***	0.13 (0.04)***	0.10 (0.03)**	0.01 (0.04)
On Track Mid	0.44 (0.03)***	0.38 (0.03)***	0.19 (0.03)***	0.26 (0.03)***	0.20 (0.03)***
On Track Top	0.58 (0.03)***	0.50 (0.03)***	0.36 (0.03)***	0.46 (0.03)***	0.32 (0.04)***
ELL	−0.03 (0.02)	0.01 (0.02)	−0.06 (0.02)**	−0.04 (0.02)*	−0.04 (0.02)*
Hispanic	−0.07 (0.04)*	−0.07 (0.04)*	−0.10 (0.04)**	−0.10 (0.04)**	−0.10 (0.04)**
Other	0.05 (0.04)	0.09 (0.04)**	0.07 (0.04)	0.08 (0.04)*	0.10 (0.04)**
Female	0.08 (0.02)***	0.07 (0.02)***	0.08 (0.02)***	0.06 (0.02)***	0.05 (0.02)**
Age	0.05 (0.03)	0.07 (0.03)**	0.10 (0.03)**	0.08 (0.03)**	0.11 (0.03)***
Intercept	0.24 (0.04)***	0.26 (0.04)***	0.42 (0.04)***	0.38 (0.04)***	0.45 (0.04)***
Within R^2	0.17	0.15	0.09	0.13	0.09

Note. Outcomes are dichotomous (0 = not proficient, 1 = proficient). ELA is English Language Arts.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

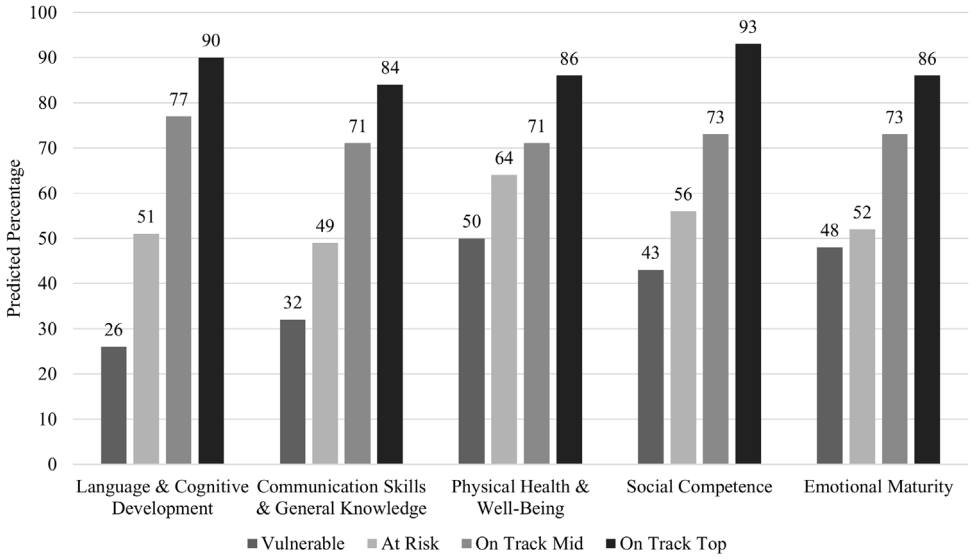

Fig. 1. Predicted Percentage Proficient on Third Grade SBA Mathematics Based on EDI Classification and Domains Controlling for Child-level Covariates and Neighborhood.

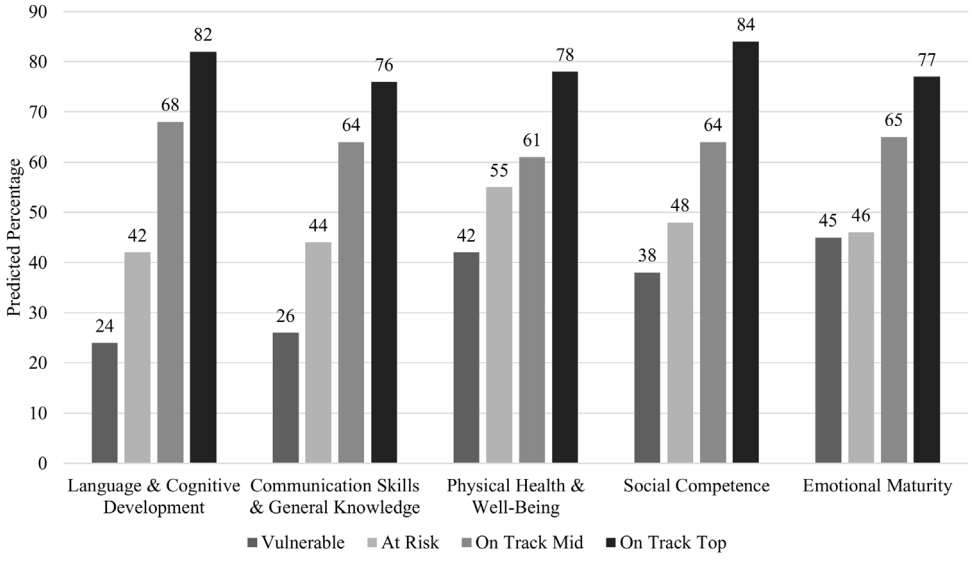

Fig. 2. Predicted Percentage Proficient on Third Grade SBA ELA Based on EDI Classification and Domains Controlling for Child-level Covariates and Neighborhood. ELA is English Language Arts.

[22.08, 52.84], $p < .001$) times greater odds of proficiency. For SBA English Language Arts, the odds of proficiency were 23.44 (95% confidence interval: [15.30, 35.91], $p < 0.001$) times greater for children coded as on track top compared with vulnerable for the language and cognitive development domain.

3.1.4. Kindergarten classroom fixed effects models

The strong associations between EDI and third grade academic proficiency held while using a kindergarten classroom fixed effects approach instead of neighborhood fixed effects (see Appendix Tables A2 and A3). Therefore, there is no evidence that differences between kindergarten classrooms (and teachers) are inflating the results because these models only examined within classroom variation and the decisions related to neighborhood versus classroom clustering did not appear to alter conclusions.

3.1.5. Including all EDI domains simultaneously

In Appendix Table A4, we report on the linear regression models for proficiency outcomes that included all EDI domains simultaneously, demographic control variables, and neighborhood fixed effects. When holding other EDI domains constant, generally significant effects for the progressions on the language and cognitive development domain and the communication skills and general knowledge domain were observed, with an additional significant association between 'vulnerable' and 'on track top' for the social competence domain. The contrast between the 'vulnerable' and 'on track top' for the language and cognitive development domain (difference of 42%, $p < .001$) remained particularly noteworthy even when holding other EDI domains constant. Additionally, when all the domains were modeled simultaneously, there was little statistical evidence that physical health and well-being and emotional maturity were related to academic proficiency above and beyond the other EDI domains.

4. Discussion

The current study demonstrated that the kindergarten assessment of the EDI strongly predicted third grade proficiency in

Table 5
Results from neighborhood fixed-effects linear regression models predicting 3rd grade math and ELA continuous scores with the kindergarten EDI and control variables (Ns = 2358 - 2391).

	3rd Grade Math				
	Lang. & Cog. Develop.	General Know. & Comm.	Physical Health & Well-Being	Social Competence	Emotional Maturity
	B (SE)	*B (SE)*	*B (SE)*	*B (SE)*	*B (SE)*
Vulnerable	reference	reference	reference	reference	reference
At Risk	46.89 (4.76)***	30.37 (4.62)***	23.57 (5.88)***	25.51 (5.48)***	12.22 (6.30)
On Track Mid	88.40 (4.41)***	69.59 (4.25)***	42.51 (5.19)***	56.20 (4.75)***	46.28 (5.33)***
On Track Top	118.79 (5.14)***	95.28 (4.63)***	70.59 (5.40)***	95.70 (5.15)***	71.46 (5.69)***
ELL	0.02 (3.30)	8.14 (3.42)*	−6.97 (3.58)	−4.04 (3.42)	−3.58 (3.61)
Hispanic	−7.72 (6.19)	−8.63 (6.34)	−14.19 (6.73)*	−14.56 (6.43)	−15.26 (6.74)*
Other	25.66 (6.14)***	32.07 (6.27)***	26.83 (6.66)***	28.54 (6.37)***	32.80 (6.65)***
Female	−5.27 (2.52)*	−8.47 (2.61)**	−5.36 (2.76)	−11.39 (2.67)***	−11.23 (2.85)***
Age	4.75 (4.22)	9.14 (4.31)*	15.71 (4.54)**	12.27 (4.36)**	16.77 (4.58)***
Intercept	2333.51 (6.69)***	2345.52 (6.73)***	2374.53 (7.16)***	2363.16 (6.96)***	2373.80 (7.46)***
Within R^2	0.26	0.22	0.13	0.20	0.13
	3rd Grade ELA				
Vulnerable	reference	reference	reference	reference	reference
At Risk	34.48 (5.19)***	29.35 (5.02)***	20.26 (6.38)**	20.33 (5.96)***	9.55 (6.83)
On Track Mid	79.47 (4.81)***	67.70 (4.61)***	34.91 (5.63)***	48.22 (5.16)***	40.21 (5.77)***
On Track Top	115.33 (5.61)***	95.49 (5.04)***	64.01 (5.85)***	91.52 (5.60)***	65.54 (6.17)***
ELL	−12.28 (3.60)**	−4.30 (3.72)	−19.15 (3.88)***	−16.46 (3.72)***	−16.69 (3.91)***
Hispanic	−7.05 (6.75)	−8.22 (6.89)	−13.24 (7.30)	−14.41 (6.99)*	−14.23 (7.30)
Other	23.12 (6.69)***	29.75 (6.81)***	25.10 (7.21)**	25.96 (6.92)***	30.63 (7.20)***
Female	14.98 (2.75)***	11.69 (2.83)***	15.32 (2.99)***	9.20 (2.90)**	9.70 (3.08)**
Age	0.51 (4.60)	4.84 (4.68)	11.97 (4.92)*	8.24 (4.73)	13.24 (4.96)**
Intercept	2317.85 (7.29)***	2323.27 (7.32)***	2356.46 (7.76)***	2345.82 (7.56)***	2355.53 (8.09)***
Within R^2	0.24	0.21	0.12	0.19	0.13

Note. Outcomes are continuous. ELA is English Language Arts. For Math, M = 2403.82, SD = 77.94. For ELA, M = 2381.42, SD = 84.20.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

mathematics and English Language Arts. The results showed that all five EDI domains have strong associations with third grade proficiency in both academic achievement areas, though the strongest associations were found for the three domains most consistently linked to later academic achievement: language and cognitive development, communication skills and general knowledge, and social competence. This study expands on previous work on teacher assessments of school readiness and children's achievement (Lambert et al., 2015; Schmitt et al., 2014). Specifically, the findings highlight that not only can teacher ratings be related to research administered direct assessments (evidence of concurrent predictive validity), but these types of instruments may be strong predictors of standardized educational testing multiple years later. Teacher ratings may be one way of collecting information on populations of kindergarteners that are cost-effective and closely relate to later achievement.

This study also extends previous research in Canada and Australia on the predictive utility of the EDI in a diverse, U.S. sample of children with multiple analytic models that control for child-level covariates and neighborhood effects (as well as classroom effects in a robustness check). Although extensive psychometric work on the EDI has been conducted in prior studies (e.g., Forer & Zumbo, 2011; Hymel et al., 2011; Janus & Offord, 2007; Janus et al., 2018), the current study is the first to detail the longitudinal associations between kindergarten teacher ratings on the EDI and third grade standardized testing performance in a U.S. population. In general, the empirical results support the use of the EDI ratings as a strong indicator of which children are most likely to reach third grade proficiency in mathematics and English Language Arts. Additionally, these findings are consistent with previous research on the EDI as a valid measure of school readiness that can be used to assess the population impact of a variety of school readiness efforts during the early childhood years on school readiness (e.g., Goldfeld et al., 2016; Guthridge et al., 2016; Patel et al., 2016) and formal schooling years.

Findings from the current study only speak to how the EDI relates to academic success in mathematics and English Language Arts three years later. Although our results consistently find the tightest coupling between early language and cognitive development with these outcomes, other key outcomes (e.g., school suspensions, grade retention) may be better predicted by different EDI domains (e.g., emotional maturity, social competence). These other key educational outcomes remain key questions worth addressing in future research.

4.1. The EDI and school readiness skills as predictors of future achievement

Building on the skills beget skills theory of human capital (Heckman, 2000), our findings provide additional evidence that skills early in school are strong indicators of children's academic performance years later. Our findings not only support the predictive utility of the EDI specifically, but are also consistent with results from a broader field of research linking early childhood skills to later academic achievement (e.g., Blair & Raver, 2015; Cameron et al., 2012; Davies et al., 2016; Duncan et al., 2007; Grissmer et al., 2010). Largely consistent with prior research, our models found the language and cognitive development domain and the communication skills and general knowledge domain to be most strongly related to both academic proficiency and a continuous measure of academic performance. These findings held when examining the EDI domains independently and simultaneously. This suggests that while each of the EDI domains were related to later achievement, regardless of what was shared across the five EDI domains, the language and cognitive development domain and the communication skills and general knowledge domain remained significantly

associated with later educational outcomes. Children in the highest group of social competence were also highly likely to do well academically, even when controlling for other EDI domains. This is consistent with theories on the importance of self-regulation during the early schooling years for fostering academic performance (Blair & Raver, 2015). Overall, these findings continue to underscore the longstanding evidence that these early skills are key predictors of future academic performance.

Unlike prior many research studies that have systematically designed longitudinal studies that assess skills around kindergarten entry and follow-up with children for their later achievement (e.g., Duncan et al., 2007), this study is rather novel in its combination of data sources. It capitalizes on efforts by researchers and community partners to understand geographic disparities in school readiness, with population-wide EDI assessments every three years, and school district collaboration by allowing the linking of children's standardized testing performance to EDI records. These researcher-practitioner partnerships yield opportunities for increasing the translation of research and practice to support children's healthy development.

4.2. Implications

In the United States, there is currently no universally accepted way of measuring school readiness across communities. Although local, state, and national policymakers are developing strategies and policies for promoting the health and well-being of young children (e.g., universal prekindergarten), data across entire populations of children are needed to inform these efforts and to compare and evaluate them over time.

Communities in the U.S. are currently using of the EDI results in a variety of ways to reduce risks and improve the services and supports available to families that support school readiness. These include the creation of county- and city-wide early childhood coalitions, data driven planning efforts aimed at improving the coordination and alignment of services across the community. EDI data has also been used to inform the allocation and/or redistribution of local resources within a community to better target specific geographic locales and areas of need, including targeted campaigns to improve the quantity or quality of school readiness services. Policymakers have utilized data from the EDI to inform the creation of city-level early childhood policies, including creating new administrative structures that coordinate efforts for young children and their families across all municipal departments. Because the EDI data are often mapped at a neighborhood level, some cities have used the EDI data to target specific interventions and policies with greater geographic specificity.

The EDI has also been used as a civic engagement tool to build capacity of residents to understand the needs of their children and to advocate and participate in specific change efforts in their communities. Residents have been trained to interpret their local EDI scores and identify areas of need based on the neighborhood level EDI results so that they can use this information to create community change projects with community partners to bolster school readiness.

In the Canadian, Australian and now in several U.S. settings, EDI data are typically collected once every three years so that population level changes can be tracked over time, and in relation to programmatic and policy changes instituted in specific communities. This every three-year administration limits the burden placed on teachers, and potentially provides enough time for the effects of new policies and programs to be detectable through subsequent waves of EDI measurement. It also underscores that data are not intended to be an individual child-level diagnostic assessment. Understanding how potential early childhood systems improvements are translating into measurable effects on the EDI could be used as one way for determining which strategies are most effective at promoting children's successful development.

4.3. Limitations and future directions

Given that these estimates are correlational and not causal, efforts to improve these skills alone should not be assumed to translate into the large effects reported in this study, as other broader family and community factors could be contributing to both the EDI ratings/scores and third grade achievement (Bailey, Duncan, Watts, Clements, & Sarama, 2018). Despite cautions against interpreting the correlational results from this study as causal, our findings are consistent with a large literature identifying earlier skills as both 1) important targets of interventions, and 2) important signals for future educational outcomes. The EDI was not developed as an individual diagnostic but rather a tool that can be used to understand populations of children's development across geographic areas and over time. No identifiable, child-level data is accessible. Only de-identified data are shared back to participating school districts. Aggregated, group-level data are shared widely to facilitate discussions around collective efforts to boost group-level performance. Identifying geographic areas of heightened vulnerability can lead to discussions of the types of resources needed to help support children's healthy development. Our results are consistent with using the EDI as an indicator of future academic success which can be utilized by multiple stakeholders to better understand which policies and programs are promoting school readiness skills as well as areas still in need of additional supports. In this way the EDI community level data can inform ways that a "village can help raise the child."

Additionally, future replication studies to test whether these findings hold for the EDI in other regions and populations in the United States are needed. Although our sample had a large proportion of Latinx children and English Language Learners, it had a notable underrepresentation of African American children. Though the EDI has been validated with racially diverse populations in other countries (e.g., Muhajarine, Puchala, & Janus, 2011), future research in the United States should specifically focus on expanding the examination of predictive validity for African American children. Additionally, the current study did not have individual child level indicators for socio-economic status (e.g., family income, maternal education), thus was limited in terms of controlling for these potential influences. Although we used children's neighborhood as a proxy for their socio-economic status, individual-level data would have been better able to capture the heterogeneity that exists for socio-economic status within neighborhoods. However, if these socio-economic factors (e.g., neighborhood) matter for children's achievement due to their effects on children's early skills, they may be over-controlling in the models.

Finally, it would be helpful to develop a better understanding of the impact that local community contexts (e.g., socio-economic factors, health and education systems, residential civic engagement) have on neighborhood-level EDI results and third grade achievement. In the current study, we attempted to control for neighborhood-level influences by using a neighborhood fixed effect model. Teasing out the roles of community contexts and city specific policies and resources, however, could provide a fruitful avenue of future research. For example, which community- or city-specific investments in early childhood social or educational services are translating to impacts on the EDI at the child-level or the neighborhood-aggregate level? Additionally, can school districts use the information gained from the EDI to inform early investments and programs to increase children's likelihood of later academic success?

5. Conclusions

This study has provided the first demonstration of the predictive utility of the EDI in a large and diverse sample of young children, in one county, in the United States. The findings are very consistent with conclusions from previous validation studies done in Canada and Australia – namely, the EDI provides useful information for predicting children's later academic success. With many communities attempting to implement placed based comprehensive early childhood strategies, through efforts like the federal Early Childhood Comprehensive Systems initiative, or United Ways Success by Six, the value of a measure that can used for an entire population, like the EDI, can better capture the impact of community level interventions and become all the more important. Population level data will also become important at a city and county level, as policies are implemented to improve the odds of optimal school readiness and academic success.

Funding

This research was made possible by the significant support of the Children and Families Commission of Orange County which in partnership with all elementary school districts in the county has established the first full countywide snapshot in the U.S. of children's health and development using the EDI. We are very thankful to the Commission for the countless hours spent recruiting school districts, training teachers, managing data collection efforts, and facilitating partnerships. Work on this paper was also supported by grants from the Institute of Education Sciences, U.S. Department of Education, Grant R305B120013 to the University of California, Irvine (PI: Greg Duncan) and the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number P01HD065704 (PI: Greg Duncan). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the Institute of Education Sciences, or the U.S. Department of Education.

CRediT authorship contribution statement

Robert J. Duncan: Writing - original draft, Conceptualization, Methodology, Formal analysis. **Greg J. Duncan:** Writing - review & editing, Conceptualization, Methodology. **Lisa Stanley:** Writing - review & editing, Resources, Project administration. **Efren Aguilar:** Writing - review & editing, Resources, Project administration. **Neal Halfon:** Writing - review & editing, Conceptualization, Methodology, Supervision.

Acknowledgement

We would like to acknowledge and thank the developer of the EDI, Magdalena Janus at McMaster University, for her international leadership with research and support for the use of the EDI. We also thank the school district superintendents, principals, teachers, kindergarten children and their parents who participated in this study.

Appendix A

Table A1
Results from neighborhood fixed-effects logistic regression models predicting 3rd grade math and ELA proficiency with the kindergarten EDI and control variables (Ns = 2763 - 2853).

	3rd Grade Math				
	Lang. & Cog. Develop.	General Know. & Comm.	Physical Health & Well-Being	Social Competence	Emotional Maturity
	OR (SE)	*OR (SE)*	*OR (SE)*	*OR (SE)*	*OR (SE)*
Vulnerable	reference	reference	reference	reference	reference
At Risk	3.90 (0.74)***	2.29 (0.36)***	2.01 (0.36)***	1.89 (0.33)***	1.25 (0.24)
On Track Mid	12.87 (2.38)***	6.95 (1.03)***	2.74 (0.44)***	4.23 (0.66)***	3.41 (0.57)***
On Track Top	34.15 (7.60)***	14.22 (2.41)***	6.24 (1.07)***	14.37 (2.61)***	6.94 (1.28)***
ELL	0.96 (0.11)	1.24 (0.14)	0.81 (0.09)	0.90 (0.10)	0.88 (0.10)
Hispanic	0.86 (0.19)	0.87 (0.19)	0.71 (0.15)*	0.68 (0.15)	0.69 (0.15)
Other	1.49 (0.38)	1.92 (0.47)**	1.54 (0.37)	1.60 (0.39)	1.83 (0.44)*
Female	0.86 (0.08)	0.78 (0.07)**	0.89 (0.08)	0.73 (0.07)***	0.73 (0.07)***
Age	1.03 (0.17)	1.24 (0.19)	1.43 (0.21)*	1.33 (0.20)	1.48 (0.22)**
	3rd Grade ELA				
Vulnerable	reference	reference	reference	reference	reference
At Risk	2.92 (0.58)***	2.60 (0.43)***	2.01 (0.38)***	1.72 (0.32)**	1.05 (0.21)
On Track Mid	9.44 (1.77)***	6.61 (1.03)***	2.60 (0.43)***	3.58 (0.58)***	2.69 (0.46)***
On Track Top	23.44 (5.10)***	12.73 (2.19)***	6.27 (1.10)***	10.51 (1.89)***	4.94 (0.91)***
ELL	0.85 (0.10)	1.07 (0.12)	0.73 (0.08)**	0.80 (0.09)	0.80 (0.09)*
Hispanic	0.79 (0.17)	0.77 (0.16)	0.65 (0.13)*	0.64 (0.13)*	0.64 (0.13)*
Other	1.72 (0.41)	2.10 (0.49)***	1.72 (0.39)*	0.64 (0.13)*	1.97 (0.44)**
Female	1.58 (0.14)***	1.44 (0.13)***	1.54 (0.13)***	1.82 (0.41)***	1.32 (0.12)**
Age	1.31 (0.20)	1.47 (0.22)**	1.63 (0.23)***	1.57 (0.23)**	1.74 (0.25)***

Note. Outcomes are dichotomous (0 = not proficient, 1 = proficient). ELA is English Language Arts.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

Table A2
Results from kindergarten classroom fixed-effects regression models predicting 3rd grade math and ELA proficiency with the kindergarten EDI and control variables (Ns = 2549 - 2584).

	3rd Grade Math				
	Lang. & Cog. Develop.	General Know. & Comm.	Physical Health & Well-Being	Social Competence	Emotional Maturity
	B (SE)	*B (SE)*	*B (SE)*	*B (SE)*	*B (SE)*
Vulnerable	reference	reference	reference	reference	reference
At Risk	0.24 (0.03)***	0.16 (0.03)***	0.14 (0.04)**	0.15 (0.04)***	0.02 (0.04)
On Track Mid	0.54 (0.03)***	0.43 (0.03)***	0.24 (0.04)***	0.37 (0.03)***	0.25 (0.04)***
On Track Top	0.72 (0.03)***	0.58 (0.03)***	0.43 (0.04)***	0.61 (0.04)***	0.42 (0.04)***
ELL	0.01 (0.02)	0.05 (0.02)	−0.04 (0.03)	−0.02 (0.02)	−0.04 (0.03)
Hispanic	−0.08 (0.04)	−0.08 (0.04)	−0.13 (0.05)**	−0.13 (0.04)**	−0.13 (0.05)**
Other	0.02 (0.04)	0.07 (0.04)	0.03 (0.04)	0.04 (0.04)	0.07 (0.04)
Female	−0.04 (0.02)*	−0.06 (0.02)**	−0.04 (0.02)*	−0.08 (0.02)***	−0.08 (0.02)***
Age	0.01 (0.03)	0.04 (0.03)	0.08 (0.03)*	0.06 (0.03)	0.08 (0.03)**
Intercept	0.24 (0.05)***	0.30 (0.05)***	0.49 (0.05)***	0.41 (0.05)***	0.52 (0.05)***
	3rd Grade ELA				
Vulnerable	reference	reference	reference	reference	reference
At Risk	0.14 (0.03)***	0.16 (0.03)***	0.13 (0.04)**	0.11 (0.04)**	−0.02 (0.04)
On Track Mid	0.45 (0.03)***	0.40 (0.03)***	0.21 (0.04)***	0.31 (0.03)***	0.19 (0.04)***
On Track Top	0.66 (0.04)***	0.55 (0.03)***	0.42 (0.04)***	0.55 (0.04)***	0.34 (0.04)***
ELL	0.00 (0.02)	0.04 (0.02)	−0.05 (0.03)	−0.03 (0.02)	−0.04 (0.03)
Hispanic	−0.08 (0.04)	−0.08 (0.04)	−0.12 (0.05)**	−0.13 (0.04)**	−0.12 (0.05)**
Other	0.06 (0.04)	0.11 (0.04)*	0.07 (0.04)	0.08 (0.04)	0.11 (0.04)*
Female	0.08 (0.02)***	0.07 (0.02)***	0.08 (0.02)***	0.05 (0.02)*	0.05 (0.02)**
Age	0.03 (0.03)	0.06 (0.03)	0.09 (0.03)**	0.07 (0.03)*	0.10 (0.03)**
Intercept	0.18 (0.05)***	0.20 (0.05)***	0.38 (0.05)***	0.32 (0.05)***	0.44 (0.05)***

Note. Outcomes are dichotomous (0 = not proficient, 1 = proficient). ELA is English Language Arts.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

Table A3
Results from kindergarten classroom fixed-effects logistic regression models predicting 3rd grade math and ELA proficiency with the kindergarten EDI and control variables (Ns = 2241 - 2348).

	3rd Grade Math				
	Lang. & Cog. Develop.	General Know. & Comm.	Physical Health & Well-Being	Social Competence	Emotional Maturity
	OR (SE)	*OR (SE)*	*OR (SE)*	*OR (SE)*	*OR (SE)*
Vulnerable	reference	reference	reference	reference	reference
At Risk	3.79 (0.82)***	2.30 (0.43)***	2.00 (0.41)**	2.31 (0.51)***	1.14 (0.25)
On Track Mid	16.42 (3.54)***	8.78 (1.59)***	3.21 (0.61)***	6.75 (1.39)***	3.45 (0.67)***
On Track Top	67.41(18.99)***	20.91 (4.37)***	8.75 (1.85)***	28.83 (6.88)***	8.46 (1.89)***
ELL	1.03 (0.15)	1.31 (0.18)	0.79 (0.11)	0.89 (0.13)	0.85 (0.11)
Hispanic	0.72 (0.22)	0.70 (0.20)	0.54 (0.15)*	0.47 (0.13)**	0.51 (0.14)*
Other	1.41 (0.46)	1.78 (0.56)	1.24 (0.38)	1.29 (0.40)	1.54 (0.46)
Female	0.77 (0.08)*	0.71 (0.07)**	0.81 (0.08)*	0.61 (0.06)***	0.65 (0.07)***
Age	1.07 (0.19)	1.23 (0.21)	1.46 (0.23)*	1.36 (0.23)	1.51 (0.25)*
	3rd Grade ELA				
Vulnerable	reference	reference	reference	reference	reference
At Risk	2.54 (0.58)***	2.64 (0.53)***	2.06 (0.45)**	2.05 (0.48)**	0.94 (0.22)
On Track Mid	11.65 (2.57)***	8.15 (1.55)***	3.06 (0.61)***	5.60 (1.19)***	2.68 (0.54)***
On Track Top	54.31(15.38)***	19.28 (4.11)***	8.77 (1.91)***	19.90 (4.73)***	5.73 (1.29)***
ELL	1.00 (0.14)	1.23 (0.17)	0.77 (0.10)*	0.87 (0.12)	0.80 (0.11)*
Hispanic	0.80 (0.23)	0.74 (0.20)	0.60 (0.16)	0.54 (0.15)*	0.60 (0.16)
Other	1.90 (0.58)*	2.27 (0.67)**	1.64 (0.47)	1.77 (0.52)	2.04 (0.58)*
Female	1.59 (0.16)***	1.45 (0.15)***	1.55 (0.15)***	1.29 (0.13)*	1.31 (0.13)**
Age	1.15 (0.20)	1.34 (0.23)	1.54 (0.25)**	1.50 (0.25)*	1.65 (0.27)**

Note. Outcomes are dichotomous (0 = not proficient, 1 = proficient). ELA is English Language Arts.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

Table A4
Results from neighborhood fixed-effects linear regression models predicting 3rd grade math and ELA proficiency with the kindergarten EDI domains entered simultaneously with control variables (Ns = 2905 - 2907).

		3rd Grade Math *B (SE)*	3rd Grade ELA *B (SE)*
	Vulnerable	reference	reference
	At Risk	0.17 (0.03)***	0.10 (0.03)**
Lang. & Cog. Develop.	On Track Mid	0.34 (0.03)***	0.27 (0.03)***
	On Track Top	0.42 (0.04)***	0.37 (0.04)***
	At Risk	0.04 (0.03)	0.08 (0.03)*
General Know. & Comm.	On Track Mid	0.18 (0.03)***	0.18 (0.03)***
	On Track Top	0.20 (0.04)***	0.20 (0.04)***
	At Risk	0.01 (0.03)	0.01 (0.04)
Physical Health & Well-Being	On Track Mid	−0.03 (0.03)	−0.02 (0.03)
	On Track Top	−0.01 (0.04)	0.04 (0.04)
	At Risk	−0.00 (0.04)	−0.00 (0.04)
Social Competence	On Track Mid	0.04 (0.04)	0.03 (0.04)
	On Track Top	0.12 (0.04)**	0.12 (0.04)*
	At Risk	−0.02 (0.04)	−0.05 (0.04)
Emotional Maturity	On Track Mid	0.04 (0.04)	−0.00 (0.04)
	On Track Top	0.03 (0.04)	−0.03 (0.04)
	ELL	0.02 (0.02)	−0.01 (0.02)
	Hispanic	−0.05 (0.03)	−0.08 (0.03)*
Covariates	Other	0.04 (0.03)	0.05 (0.03)
	Female	−0.06 (0.02)***	0.05 (0.02)**
	Age	0.00 (0.03)	0.04 (0.03)
	Intercept	0.20 (0.05)***	0.20 (0.05)***

Note. Outcomes are continuous. ELA is English Language Arts.
* $p < .05$.
** $p < .01$.
*** $p < .001$

References

Ackerman, D. J., & Barnett, W. S. (2005). *Prepared for kindergarten: What does "readiness" mean? NIEER policy report.*. Retrieved Online at: https://www.researchgate.net/profile/Debra_Ackerman/publication/255617499_Prepared_for_Kindergarten_What_Does_Readiness_Mean/links/0f31753b2d9f34f6f5000000/Prepared-for-Kindergarten-What-Does-Readiness-Mean.pdf.

Bailey, D. H., Duncan, G. J., Watts, T., Clements, D., & Sarama, J. (2018). Risky business: Correlation and causation in longitudinal studies of skill development. *The American Psychologist*, *73*(1), 81–94. http://dx.doi.org/10.1037/amp0000146

Blair, C., & Raver, C. C. (2015). School readiness and self-regulation: A developmental psychobiological approach. *Annual Review of Psychology*, *66*, 711–731. http://dx.doi.org/10.1146/annurev-psych-010814-015221

Brinkman, S., Gregory, T., Harris, J., Hart, B., Blackmore, S., & Janus, M. (2013). Associations between the early development instrument at age 5, and reading and numeracy skills at ages 8, 10 and 12: A prospective linked data study. *Child Indicators Research*, *6*(4), 695–708. http://dx.doi.org/10.1007/s12187-013-9189-3

Brinkman, S. A., Silburn, S., Lawrence, D., Goldfeld, S., Sayers, M., & Oberklaid, F. (2007). Investigating the validity of the australian early development index. *Early Education and Development*, *3*, 427–451. http://dx.doi.org/10.1080/10409280701610812

Cameron, C. E., Brock, L. L., Murrah, W. M., Bell, L. H., Worzalla, S. L., Grissmer, D., et al. (2012). Fine motor skills and executive function both contribute to kindergarten achievement. *Child Development*, *83*(4), 1229–1244. http://dx.doi.org/10.1111/j.1467-8624.2012.01768.x

Campbell, F. A., Pungello, E. P., Miller-Johnson, S., Burchinal, M., & Ramey, C. T. (2001). The development of cognitive and academic abilities: Growth curves from an early childhood educational experiment. *Developmental Psychology*, *37*(2), 231–242. http://dx.doi.org/10.1037/0012-1649.37.2.231

Cunha, F., & Heckman, J. (2007). The technology of skill formation. *The American Economic Review*, *97*(2), 31–47. http://dx.doi.org/10.1257/aer.97.2.31

Curtin, M., Browne, J., Staines, A., & Perry, I. J. (2016). The Early Development Instrument: An evaluation of its five domains using Rasch analysis. *BMC Pediatrics*, *16* http://dx.doi.org/10.1186/s12887-016-0543-8

Curtin, M., Madden, J., Staines, A., & Perry, I. J. (2013). Determinants of vulnerability in early childhood development in Ireland: A cross-sectional study. *BMJ Open*, *3*, e002387 http://dx.doi.org/10.1136/bmjopen-2012-002387

D'Angiulli, A., Warburton, W., Dahinten, S., & Hertzman, C. (2009). Population level associations between preschool vulnerability and grade-four basic skills. *PloS One*, *4*(11), e7692. http://dx.doi.org/10.1371/journal.pone.0007692

Davies, S., Janus, M., Duku, E., & Gaskin, A. (2016). Using the Early Development Instrument to examine cognitive and non-cognitive school readiness and elementary student achievement. *Early Childhood Research Quarterly*, *35*, 63–75. http://dx.doi.org/10.1016/j.ecresq.2015.10.002

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., et al. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428–1446. http://dx.doi.org/10.1037/0012-1649.43.6.1428

Forer, B., & Zumbo, B. D. (2011). Validation of multilevel constructs: Validation methods and empirical findings for the EDI. *Social Indicators Research*, *103*(2), 231. http://dx.doi.org/10.1007/s11205-011-9844-3

Forget-Dubois, N., Lemelin, J. P., Boivin, M., Dionne, G., Séguin, J. R., Vitaro, F., et al. (2007). Predicting early school achievement with the EDI: A longitudinal population-based study. *Early Education and Development*, *18*(3), 405–426. http://dx.doi.org/10.1080/10409280701610796

Goldfeld, S., O'Connor, E., O'Connor, M., Sayers, M., Moore, T., Kvalsvig, A., et al. (2016). The role of preschool in promoting children's healthy development: Evidence from an Australian population cohort. *Early Childhood Research Quarterly*, *35*, 40–48. http://dx.doi.org/10.1016/j.ecresq.2015.11.001

Goldfeld, S., Sayers, M., Brinkman, S., Silburn, S., & Oberklaid, F. (2009). The process and policy challenges of adapting and implementing the Early Development Instrument in Australia. *Early Education and Development*, *20*(6), 978–991. http://dx.doi.org/10.1080/10409280903375800

Grissmer, D., Grimm, K. J., Aiyer, S. M., Murrah, W. M., & Steele, J. S. (2010). Fine motor skills and early comprehension of the world: Two new school readiness indicators. *Developmental Psychology*, *46*, 1008–1017. http://dx.doi.org/10.1037/a0020104

Guhn, M., Gadermann, A. M., Almas, A., Schonert-Reichl, K. A., & Hertzman, C. (2016). Associations of teacher-rated social, emotional, and cognitive development in kindergarten to self-reported wellbeing, peer relations, and academic test scores in middle childhood. *Early Childhood Research Quarterly*, *35*, 76–84. http://dx.doi.org/10.1016/j.ecresq.2015.12.027

Guhn, M., Gadermann, A., & Zumbo, B. D. (2007). Does the EDI measure school readiness in the same way across different groups of children? *Early Education and Development*, *18*(3), 453–472. http://dx.doi.org/10.1080/10409280701610838

Guthridge, S., Li, L., Silburn, S., Li, S. Q., McKenzie, J., & Lynch, J. (2016). Early influences on developmental outcomes among children, at age 5, in Australia's Northern Territory. *Early Childhood Research Quarterly*, *35*, 124–134. http://dx.doi.org/10.1016/j.ecresq.2015.12.008

Heckman, J. J. (2000). Policies to foster human capital. *Research in Economics*, *54*(1), 3–56. http://dx.doi.org/10.1006/reec.1999.0225

Hymel, S., LeMare, L., & McKee, W. (2011). The Early Development Instrument: An examination of convergent and discriminant validity. *Social Indicators Research*, *103*(2), 267–282. http://dx.doi.org/10.1007/s11205-011-9845-2

Israel, B. A., Schulz, A. J., Parker, E. A., & Becker, A. B. (1998). Review of community-based research: assessing partnership approaches to improve public health. *Annual Review of Public Health*, *19*(1), 173–202. http://dx.doi.org/10.1146/annurev.publhealth.19.1.173

Janus, M. (2013). *Early Development Instrument: "from results to action survey" report* Retrieved from:. https://edi.offordcentre.com/wp/wp-content/uploads/2015/06/EDI_to_Action_Survey_Report.pdf

Janus, M., & Offord, D. R. (2007). Development and psychometric properties of the Early Development Instrument (EDI): A measure of children's school readiness. *Canadian Journal of Behavioural Science*, *39*, 1–22. http://dx.doi.org/10.1037/cjbs2007001

Janus, M., Brinkman, S. A., & Duku, E. K. (2011). Validity and psychometric properties of the early development instrument in Canada, Australia, United States, and Jamaica. *Social Indicators Research*, *103*(2), 283–297. http://dx.doi.org/10.1007/s11205-011-9846-1

Janus, M., Zeraatkar, D., Duku, E., & Bennett, T. (2018). Validation of the Early Development Instrument for children with special health needs. *Journal of Pediatrics and Child Health*, 1–7. http://dx.doi.org/10.1111/jpc.14264

Jutte, D. P., Miller, J. L., & Erickson, D. J. (2015). Neighborhood adversity, child health, and the role for community development. *Pediatrics*, *135*, S48–S57. http://dx.doi.org/10.1542/peds.2014-3549F

Krieger, N., Chen, J. T., Waterman, P. D., Rehkopf, D. H., & Subramanian, S. V. (2003). Race/ethnicity, gender, and monitoring socioeconomic gradients in health: A comparison of area-based socioeconomic measures—The public health disparities geocoding project. *American Journal of Public Health*, *93*(10), 1655–1671. http://dx.doi.org/10.2105/AJPH.93.10.1655

Lambert, R. G., Kim, D. H., & Burts, D. C. (2015). The measurement properties of the Teaching Strategies GOLD® assessment system. *Early Childhood Research Quarterly*, *33*, 49–63. https://doi.org/10.1016/j.ecresq.2015.05.004

Lloyd, J. E., Irwin, L. G., & Hertzman, C. (2009). Kindergarten school readiness and fourth-grade literacy and numeracy outcomes of children with special needs: A population-based study. *Educational Psychology*, *29*(5), 583–602. http://dx.doi.org/10.1080/01443410903165391

Morrissey, T. W., & Vinopal, K. M. (2018). Neighborhood poverty and children's academic skills and behavior in early elementary school. *Journal of Marriage and the Family*, *80*(1), 182–197. http://dx.doi.org/10.1111/jomf.12430

Muhajarine, N., Puchala, C., & Janus, M. (2011). Does the EDI equivalently measure facets of school readiness for Aboriginal and non-Aboriginal children? *Social Indicators Research*, *103*(2), 299–314. http://dx.doi.org/10.1007/s11205-011-9847-0

Patel, S., Corter, C., Pelletier, J., & Bertrand, J. (2016). 'Dose-response' relations between participation in integrated early childhood services and children's early development. *Early Childhood Research Quarterly*, *35*, 49–62. http://dx.doi.org/10.1016/j.ecresq.2015.12.006

Schmitt, S. A., Pratt, M. E., & McClelland, M. M. (2014). Examining the validity of behavioral self-regulation tools in predicting preschoolers' academic achievement. *Early Education and Development*, *25*(5), 641–660. https://doi.org/10.1080/10409289.2014.850397

Smarter Balanced Assessment Consortium. (2016). *Smarter balanced assessment consortium: 2014-2015 technical report*.

StataCorp. (2017). *Stata statistical software: Release 15*. College Station, TX: StataCorp LP.

Waterman, C., McDermott, P. A., Fantuzzo, J. W., & Gadsden, V. L. (2011). The matter of assessor variance in early childhood education – Or whose score is it anyway? *Early Childhood Research Quarterly*, *27*, 46–54. http://dx.doi.org/10.1016/j.ecresq.2011.06.003