

UC Santa Barbara

UC Santa Barbara Previously Published Works

Title

Dimensionality and Measurement Invariance of a School Readiness Screener by Gender and Parent Education Levels

Permalink

<https://escholarship.org/uc/item/9wt0d823>

Journal

International Journal of School & Educational Psychology, 3(3)

Author

Furlong, Michael James

Publication Date

2015

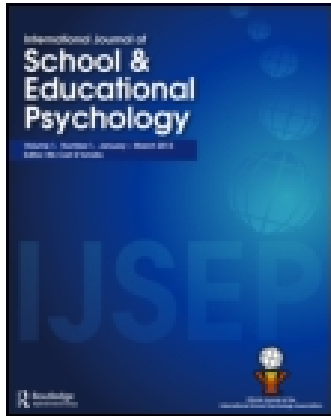
Peer reviewed

This article was downloaded by: [University of California Santa Barbara]

On: 24 August 2015, At: 11:26

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: 5 Howick Place, London, SW1P 1WG



International Journal of School & Educational Psychology

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/usep20>

Dimensionality and Measurement Invariance of a School Readiness Screener by Gender and Parent Education Levels

Matthew Quirk^a, Ashley Mayworm^a, Michael J. Furlong^a, Ryan Grimm^a & Jennica Rebelez^a

^a Department of Counseling, Clinical, and School Psychology, University of California, Santa Barbara, California, USA

Published online: 17 Aug 2015.



[Click for updates](#)

To cite this article: Matthew Quirk, Ashley Mayworm, Michael J. Furlong, Ryan Grimm & Jennica Rebelez (2015) Dimensionality and Measurement Invariance of a School Readiness Screener by Gender and Parent Education Levels, *International Journal of School & Educational Psychology*, 3:3, 167-177, DOI: [10.1080/21683603.2015.1053644](https://doi.org/10.1080/21683603.2015.1053644)

To link to this article: <http://dx.doi.org/10.1080/21683603.2015.1053644>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

Dimensionality and Measurement Invariance of a School Readiness Screener by Gender and Parent Education Levels

Matthew Quirk, Ashley Mayworm, Michael J. Furlong, Ryan Grimm,
and Jennica Rebelez

*Department of Counseling, Clinical, and School Psychology,
University of California, Santa Barbara, California, USA*

This study examined the dimensionality and measurement invariance of the Kindergarten Student Entrance Profile (KSEP), a brief screening tool designed for use as a universal school readiness assessment. Teachers rated the readiness of 10,031 children during the first month of kindergarten in four ethnically diverse, medium-sized school districts in central California. From the total sample, two random, independent subsamples were identified. First, S1 ($n = 5,050$) was utilized to conduct a CFA. Results yielded evidence supporting a two-factor structure encompassing children's social-emotional and cognitive readiness that is similar to previous studies examining the KSEP with less diverse student samples. Next, a series of confirmatory factor analyses (CFAs) were conducted using S2 ($n = 4,981$) to replicate the structure identified in the first CFA with an independent subsample and to test the measurement invariance of the KSEP across two sets of categorical variables: (a) gender and (b) parent education levels. Results from both sets of CFAs showed adequate fit to the two-factor structure, with the KSEP exhibiting measurement invariance across both gender and parent education levels. The results of this study provide additional psychometric evidence supporting the validity of the KSEP as a universal-level school readiness screener. Practical implications are also discussed.

Keywords: School readiness, measurement invariance, Kindergarten Student Entrance Profile, factor analysis, screening

There is renewed awareness of the importance of preschool opportunities for young children in the United States as a matter of public policy interest (Obama, 2013, 2014). Many states have initiated or established efforts to expand preschool opportunities, with the goal of providing all families options for sending their children to preschool prior to entering into kindergarten (e.g., Georgia, Florida, Illinois, and Oklahoma, among others). Although this is a relatively recent movement in the United States, many European countries have a long history of offering universal preschool programs. In a recent keynote address at the National Association of School Psychologists

convention, Darling-Hammond (2011) provided a persuasive argument regarding the importance of universal preschool while discussing the effects of poverty on student achievement in many American communities. Public policy that supports preschool for all is predicated on a growing body of research evidence linking school readiness and subsequent academic success (Matthews, Kizzie, Rowley, & Cortina, 2010; Pianta, Barnett, Burchinal, & Thornburg, 2009; Quirk, Nylund-Gibson, & Furlong 2013) and the realization that children's life course achievement trajectories stabilize very early in their academic careers (Entwisle & Hayduk, 1988; Hulslander, Olson, Willcutt, & Wadsworth, 2010; Torgesen & Burgess, 1998). Some students' long-term academic trajectories are formed even before standardized achievement tests are typically first administered in Grades 2 or 3—a status quo approach that misses the critical opportunity to provide

Correspondence should be addressed to Matthew Quirk, University of California, Santa Barbara, Gevirtz Graduate School of Education, CCSP Department, Santa Barbara, CA 93106, USA. E-mail: mquirk@education.ucsb.edu

targeted services and supports in Grades K and 1 when they are most needed.

The term *school readiness* has been used broadly to describe factors at multiple levels (e.g., child, family, school, community, etc.) that influence children's transition to formal schooling (Rimm-Kaufman & Pianta, 2000). However, within the context of the current study, and following Snow's (2006) perspective, school readiness is defined as the state of children's skills and dispositions at the time of kindergarten entry that are associated with later success. Although there is no unanimity on the specific configuration of subcomponents or dimensions that comprise children's school readiness, there is increasing consensus that readiness is comprised of a combination of cognitive and social-emotional elements (Blair, 2002). This perspective is supported by a growing body of research that has demonstrated associations between children's cognitive and social-emotional readiness skills and dispositions, and later academic success (Duncan et al., 2007; Galindo & Fuller, 2010; Romano, Babchishin, Pagani, & Kohen, 2010; Sabol & Pianta, 2012).

Given the influence of school readiness and early development on students' long-term academic outcomes, screening for readiness at the time of kindergarten entry should be a critical component of school districts' policies and practices. Universal school readiness screening provides educators with an opportunity to gather information on all incoming students to discern who might benefit from early supplemental supports. In recognition of this unique and critical opportunity, many states and districts have developed and implemented universal school readiness screening procedures (e.g., Florida Kindergarten Readiness Screener [FLKRS], Georgia Kindergarten Inventory of Developing Skills [GKIDS]); however, few of these assessments have been researched extensively to examine specific psychometric characteristics and relations with children's longitudinal achievement levels.

One exception is the Kindergarten Student Entrance Profile (KSEP; Santa Maria–Bonita School District, First 5 of Santa Barbara County, & University of California Santa Barbara, 2005), which is a brief school readiness screening assessment that evaluates social-emotional and cognitive elements of children's readiness during the first month of kindergarten. Previous KSEP research has found evidence supporting its ability to predict children's longitudinal literacy achievement through Grades K and 1 (Quirk, Furlong, Lilles, Felix, & Chin, 2011) as well as its ability to yield differentiable readiness profiles that are predictive of children's reading and math achievement at the end of Grade 2 (Quirk et al., 2013). In addition, previous studies have linked known correlates of readiness (e.g., age, preschool experience, gender) to KSEP ratings at the time of kindergarten entry (Furlong & Quirk, 2011). Finally, a recent study (Quirk, Rebelez, & Furlong, 2014) found evidence supporting a dual-factor structure for the KSEP,

suggesting that it measures two distinct, yet related aspects of children's school readiness (*social-emotional* and *cognitive*) at the time that they enter kindergarten. In each of these studies the KSEP has yielded adequate reliability estimates, with Cronbach's alpha coefficients ranging from .81 to .92.

Despite these promising results, much of the KSEP research to date has drawn upon samples comprised almost exclusively of Latino/a children, many of whom were from families experiencing low economic circumstances or who were designated as English Language Learners. While these studies have provided important evidence supporting the KSEP's use with these populations of students, the demands placed on a universal screening instrument require broader validity evidence to support its use with diverse student populations.

PURPOSE OF THE CURRENT INVESTIGATION

Although previous studies provide a substantial body of evidence supporting the importance of school readiness in predicting longitudinal outcomes for many different student subpopulations, few studies have specifically examined the measurement invariance of school readiness measures across children from various demographic backgrounds. Although some studies examined the invariance of specific cognitive or academic skills tests (Denham, Warren-Khot, Bassett, Wyatt, & Perna, 2012; Fuhs & Day, 2011), only one previous study (Csapó, Molnár, & Nagy, 2014) examined a school readiness measure, specifically, by testing its invariance across modes of delivery (computer-based versus paper-and-pencil). The lack of empirical evidence supporting the invariance of specific readiness measures across student demographic characteristics is particularly significant given the broad use of school readiness assessments with students from diverse backgrounds and in relation to findings from previous research that have identified differences in readiness across subpopulations using these measures. For example, research has identified that socioeconomic circumstances are associated with children's school readiness (Bulotsky-Shearer & Fantuzzo, 2011; Foster, Lambert, Abbott-Shim, McCarty, & Franze, 2005; Stipek & Ryan, 1997) and that, on average, females enter school with more advanced social and behavioral skills that contribute to early academic advantages (DiPrete & Jennings, 2012); however, studies have not yet examined whether specific school readiness assessments function similarly across these subgroups (socioeconomic and gender) rendering previous cross-group comparisons tenuous at best.

Establishing measurement invariance for specific assessments is a necessary condition prior to conducting research that compares means and relational differences among different groups (Borsboom, 2006; Milfont & Fischer,

2010) and is essential for measures such as the KSEP when used at the individual student level (Meredith, 1993). Measurement invariance needs to be established before researchers can assume that group differences in observed scores are indicative of actual differences between groups on latent constructs. Establishing measurement invariance across groups increases confidence that a scale is measuring the same construct for each of the groups being examined or compared and provides information to assess any cross-group measurement bias.

The current study extends previous research in multiple ways. First, it replicated KSEP analyses previously conducted with Latino/a student samples (Quirk et al., 2014) to determine whether its two-dimensional structure held when examined with data from a more diverse sample of students. In addition, the current study addressed a critical gap in the school readiness assessment literature by testing measurement invariance of the KSEP across gender and parent education variables. Evaluating measurement invariance for these two variables is particularly critical for the validation of any universal measure used in schools because almost all classrooms are comprised of students from both genders and from varying socioeconomic backgrounds. Establishing the dimensionality and invariance of the KSEP across these groups would allow for schools and researchers to assume that comparisons made across students from different genders and socioeconomic backgrounds are valid.

METHOD

Participants

Participants in this study included all entering kindergarten students ($N = 10,031$) from four medium-sized school districts in California. Of the participating students, 47% were identified as female, and the average age of students at the time of assessment was 5 years, 4 months old. A diverse range of ethnicities was represented in this sample, with 45% identified as Hispanic/Latino, 33% White, 11% Asian, 6% Black, 1% Pacific Islander, 1% Alaskan/Native American, and 2% Mixed Background. Over half of the sample (71%) had English reported as their primary language, followed by Spanish (22%), Hmong (4%), and other (4%). Across the four participating districts, the proportion of children qualifying for free or reduced-price lunch services ranged from 38% to 86% ($M = 66%$), indicating that a majority of students were from families experiencing low socioeconomic circumstances.

Using SPSS version 21, the total sample was split into two random samples with approximately 50% of the participants in each sample, with $n = 5,050$ in sample 1 (S1) and $n = 4,981$ in sample 2 (S2). The two samples were

TABLE 1
Participant Demographics for Each Independent Random Sample

Demographic variable	Sample 1 ($n = 5,050$)		Sample 2 ($n = 4,981$)	
	<i>n</i>	%	<i>n</i>	%
<i>Gender</i>				
Female	2,385	47.2%	2,364	47.5%
Male	2,665	52.8%	2,617	52.5%
<i>Primary language</i>				
English	3,534	70.0%	3,535	71.0%
Spanish	1,104	21.8%	1,061	21.3%
Hmong	200	4.0%	190	3.8%
Other	212	4.2%	195	3.9%
<i>Parent Education</i>				
High school diploma or less	1,409	27.9%	1,442	29.0%
Some college/AA degree	1,065	21.1%	1,045	21.0%
Bachelor's degree or higher	1,455	28.8%	1,415	28.4%
Unknown or declined	1,121	22.2%	1,079	21.7%
<i>Sociocultural background</i>				
Hispanic/Latino	2,286	45.3%	2,210	44.4%
Asian	552	10.9%	566	11.4%
Black	315	6.2%	311	6.2%
Pacific Islander	63	1.2%	62	1.2%
White	1,657	32.8%	1,634	32.8%
American Indian or Alaskan Native	40	0.8%	49	1.0%
Mixed	117	2.3%	127	2.5%
Missing	20	0.4%	22	0.4%

selected from the total sample so that unique subsamples of data could be used in the different stages of the primary data analysis plan. Table 1 provides a summary of demographic information for the two independent samples.

Measure

Kindergarten Student Entrance Profile

The Kindergarten Student Entrance Profile (KSEP; Quirk et al., 2014) was used as a universal screening tool to assess social-emotional and behavioral and cognitive aspects of children's school readiness during the first month of kindergarten. The KSEP is not a direct assessment of the child. Rather, it is a rating scale completed by teachers on the basis of their observations and professional judgments regarding the readiness of children whom they have had the opportunity to observe in the natural classroom environment over at least a three-week period. All of the teachers who completed ratings for this study participated in an online training session focused on procedures for administering and scoring the KSEP. In particular, these training sessions focused significant time on learning the KSEP rubric, which is described in greater detail in a later section.

The KSEP protocol used in this study included 12 items that previous research has linked to social-emotional and

behavioral (6 items) and cognitive (6 items) elements of children's school readiness (Quirk et al., 2014). As part of an ongoing scale refinement process, Item 13 (children's letter name knowledge) was added to the cognitive readiness scale, as research has consistently yielded evidence of its importance as an emergent literacy skill (Adams, 1990; Drouin, Horner, & Sondergeld, 2012). Each item on the KSEP protocol is accompanied by a four-point rating rubric that provides an operational definition and an example of the type of behaviors that would be indicative of a child who exhibits various levels of mastery. For example, the rubric for the KSEP item assessing impulse control contains the following descriptions of behavioral markers at each level: (a) *not yet*—unable to delay having wants and needs met; (b) *emerging*—distracted by getting wants and needs met, yet able to be redirected by others; (c) *almost mastered*—distracted by getting wants and needs met but redirects self; and (d) *mastered*—able to delay wants and needs until appropriate time. Total ratings on the 13-item scale range from 13 to 52, with a rating total of 52 indicating that the child demonstrates mastery across all of the areas measured. Because the KSEP is not an assessment of English language proficiency, teachers are directed to consider observational evidence of mastery in any language or mode of communication. In fact, many of the items from the social-emotional and behavioral domain of the KSEP do not require language in any form to demonstrate mastery.

In previous research, the KSEP has demonstrated strong reliability, with total internal consistency coefficients ranging from .91 to .92 (Lilles et al., 2009; Quirk et al., 2011) and individual subscale reliability coefficients of .88 for the social-emotional and behavioral subscale and .81 for the cognitive subscale (Quirk et al., 2014). These and other studies have also found evidence to support the validity of the KSEP, with results indicating that ratings are associated with variables known to influence children's school readiness at kindergarten entry (Furlong & Quirk, 2011) and are also predictive of children's subsequent academic achievement (Nylund-Gibson, Grimm, Quirk, & Furlong, 2014; Quirk et al., 2013). For the current sample, the overall internal consistency (Cronbach's alpha) of the 13-item scale was .92, with subscale reliability coefficients of .90 (social-emotional and behavioral) and .89 (cognitive). Total scores for the full sample ranged from 13 to 52 ($M = 41.62$, $SD = 9.10$).

Procedure

Data collection for this study was conducted within the context of a broader cross-district initiative facilitated by a county office of education in the first year of a two-year process implementing a universal school readiness screening procedure. Prior to the first week of the school year, all of the kindergarten teachers in the participating districts were required to complete an online training session that

provided explicit guidance on using the KSEP rubric to complete ratings for all entering students during the fall of 2012. Each student's primary kindergarten teacher completed all KSEP ratings during the first month of the school year. District personnel entered these data into each district's primary student information database, where KSEP data were also linked with other demographic variables (e.g., gender, parent education, etc.) and stripped of unique identifying information before being shared with the primary investigators for evaluation purposes, per the requirements of the university's human subjects review board.

Due to the increased complexity of interpreting invariance test results according to the number of groups tested, the researchers collapsed the parent education variable from five categories to three. The original data included the following parent education level categories: (a) not high school graduate, (b) high school graduate, (c) some college or associate degree, (d) college graduate, and (e) graduate school. The data that were analyzed collapsed the original distinctions into the following three groups: (a) parents with a high school diploma or less, (b) parents with some college or associate degree, and (c) parents with a college degree or higher. The researchers chose these groupings for two reasons. First, the parent education variable was examined as a proxy for socioeconomic circumstances; therefore, groups were collapsed to reflect commonly used markers of socioeconomic differences by education level (e.g., high school, some college, college graduate). Second, collapsing the groups in this configuration resulted in similar group sizes, which was preferable for data analysis purposes.

Data Analysis Plan

To investigate the measurement invariance of the KSEP, the fit of a CFA model was tested, followed by the imposition of a series of parameter constraints using *Mplus* (version 7.1) statistical software (Muthén & Muthén, 2008–2013). First, data were screened to ensure the variables did not violate assumptions of multivariate normality. Next, using SPSS version 21, a random split was applied to the overall sample, resulting in two random samples. These were used to confirm the factor structure and, subsequently, to test measurement invariance of the KSEP. The first random sample (S1, 50.3%) was used to conduct a confirmatory factor analysis (CFA), and the second random sample (S2, 49.7%) was used to test for measurement invariance based on gender and parent education level.

Confirmatory factor analysis

A previous study explored and confirmed the factor structure of the KSEP as a screener for use with an entirely Latino/a kindergarten sample (Quirk et al., 2014). Results of

that study supported a two-factor model (*social-emotional* and *cognitive*) with correlated residuals on three pairs of items. The current study aimed to confirm this two-factor structure with a larger and more diverse sample of students and test for measurement invariance across gender and parent education level. Thus, using random S1, a CFA was run using maximum likelihood (ML) estimator. ML estimates the parameters of a statistical model by maximizing the likelihood of obtaining the observed values given the model selected (Brown, 2006). Unit Loading Identification (ULI) was used to determine the scale of the factors; the unstandardized factor loading of the observed reference variable for each factor was fixed to 1.0 and the remaining loadings were allowed to be freely estimated. To evaluate model fit for the baseline CFA, the following criteria were considered as indicative of good fit: a nonsignificant chi-square value, CFI > .90, SRMR ≤ .08, and RMSEA values ≤ .08 (Hu & Bentler, 1999). Also, factor loadings exceeding .30–.35 were considered adequate (Brown, 2006).

Multiple-groups analysis: Statistical methodology

A second CFA was run with S2 to confirm the factor structure of the items in the second sample and to serve as a baseline of model fit for the multiple-groups analysis. After confirming the factor structure of the items for Sample 2, multiple-group analysis was conducted to examine whether measurement invariance existed across: (a) male and female participants and (b) parent education levels.

To establish measurement invariance for gender, a series of steps was implemented in a hierarchical fashion (Brown, 2006). The first step involved testing the CFA model separately for males and females to determine whether the model fit well for both groups. Step 2, testing configural invariance, involved running another CFA on the groups of interest (males and females) simultaneously. The primary question at this step is to examine the patterns of factor loadings across groups as well as the number and consistency of latent factors (Cheung & Rensvold, 2002). To accomplish this, it is necessary to fix other parameters to equality while allowing factor loadings to freely vary. As such, the loadings of the indicators were freely estimated for both factors, the factor means were fixed to zero, and intercepts were constrained to equality. Configural invariance indicates whether males and females have the same number of factors and if the same indicators load onto the same factors for both groups (Cheung & Rensvold, 2002). This model was used as a comparison testing the fit of this model to subsequent models with additional parameter constraints. The third step, examining metric invariance, involved constraining the factor loadings to equality across groups, allowing an examination of whether the values of the factor loadings were equal for males and females. The fourth and final step tested scalar invariance, which required

constraining the item intercepts to equality in addition to the factor loadings. Scalar invariance indicates that regardless of one's group membership (male or female), those with the same score on the latent construct will have the same score on the observed variable (Tsaousis & Kazi, 2013). These steps were repeated a second time to examine whether or not measurement invariance existed across participants' parent education levels. The parent education grouping variable that was used in these analyses included: (a) parents with a high school diploma or less, (b) parents with some college or associate degree, and (c) parents with a college degree or higher.

Determining whether or not parameters were invariant for both gender and parent education level was accomplished by evaluating the differences in model fit between subsequent models (e.g., configural invariance to metric invariance and metric invariance to scalar invariance). Both the chi-square difference test (significant difference at $p < .05$) and changes in the CFI index (noninvariant if $\Delta\text{CFI} > .01$) were used to evaluate measurement invariance. Cheung and Rensvold (2002) have recommended the use of change in CFI ($\Delta\text{CFI} > .01$) for measurement invariance testing because chi-square testing is sensitive to large sample sizes. Because of the particularly large sample size in the current study, change in CFI was the primary metric used to examine invariance as it was determined to be a more appropriate indicator of significant differences between nested models. CFI differences greater than .01 were interpreted as measurement noninvariance across the groups being tested.

RESULTS

Data Screening

Prior to the primary analyses, data were screened to determine whether the variables included in the CFAs met all prerequisite statistical assumptions. Histograms, box plots, and Q-Q plots were used to examine univariate and multivariate normality. Results from these analyses indicated no significant multivariate outliers and all items had unimodal distributions. Skewness and kurtosis values for all 13 KSEP items did not exceed critical limits (|2.0| for skewness and |7.0| for kurtosis; Chou & Bentler, 1995; Curran, West, & Finch, 1996), suggesting no major violations to normality. Bivariate correlations for all 13 items showed significant ($p < .05$) but moderate correlations across items, indicating no major concerns of multicollinearity. Table 2 provides a summary of bivariate correlations, means, standard deviations, and skewness and kurtosis for each of the 13 KSEP items. Overall, these preliminary results suggest that the KSEP data met all prerequisite assumptions for use in the subsequent analyses.

TABLE 2
Correlation Matrix, Means, Standard Deviations, and Normality for All 13 KSEP Items ($N = 10,031$)

KSEP item	1	2	3	4	5	6	7	8	9	10	11	12	13
1	—												
2	.62	—											
3	.46	.47	—										
4	.53	.46	.75	—									
5	.63	.62	.53	.59	—								
6	.59	.54	.63	.70	.63	—							
7	.47	.43	.38	.44	.47	.47	—						
8	.63	.54	.33	.40	.52	.46	.47	—					
9	.50	.40	.36	.43	.45	.47	.55	.52	—				
10	.46	.42	.38	.44	.44	.47	.72	.45	.57	—			
11	.43	.36	.30	.35	.39	.37	.48	.50	.52	.49	—		
12	.39	.31	.24	.31	.34	.33	.42	.44	.51	.45	.58	—	
13	.42	.32	.28	.36	.37	.38	.50	.44	.57	.53	.54	.58	—
<i>M</i>	3.15	3.51	3.10	2.92	3.34	3.21	3.48	3.09	3.16	3.41	3.36	3.18	2.92
<i>SD</i>	0.92	0.76	0.95	0.91	0.83	0.81	0.88	0.93	1.04	0.96	0.89	1.04	1.23
<i>Skewness</i>	-0.59	-1.46	-0.62	-0.43	-1.03	-0.91	-1.64	-0.67	-0.93	-1.46	-1.24	-0.95	-0.56
<i>Kurtosis</i>	-0.91	1.33	-0.80	-0.69	0.15	0.45	1.65	-0.55	-0.46	0.82	0.57	-0.46	-1.34

Note. All correlations significant at $p < .01$.

Confirmatory Factor Analysis

Using S1 ($n = 5,050$), a CFA was run to confirm the two-factor solution for the KSEP that was supported in previous studies (Quirk et al., 2014). The factor structure tested included two factors: a *Social-Emotional* factor (with KSEP items 1–6) and a *Cognitive* factor (with KSEP items 7–13) with three theoretically supported correlated residuals (item 3 with item 4, item 7 with item 10, and item 11 with item 12). All item pairs with correlated residuals requested ratings on theoretically and practically related skills or dispositions (e.g., recognizes written name and writes name). Additionally, the correlated residuals appeared to reflect method effects related to item wording. Thus, associations among error terms were considered pertinent a priori as a result of previous studies and the current administration of the KSEP.

The fit statistics for the two-factor solution with correlated residuals were satisfactory when considered together (Brown, 2006), $\chi^2(64) = 2284.723$, $p < .001$, SRMR = 0.046, CFI = 0.940, and RMSEA = 0.085 90% CI = [0.082, 0.088]. Although the RMSEA was slightly above the recommended cutoff, the CFI and SRMR were well within acceptable limits provided by Hu and Bentler (1999). In fact, Hu and Bentler also found that the SRMR was the most sensitive index when used as the sole criterion in evaluating the fit of CFA models, correctly rejecting 99% to 100% of misspecified models when a < 0.05 cutoff was applied and sample size was $N > 250$, as was the case in this analysis. In addition, all standardized factor loadings were above .60 (large magnitude; Brown, 2006) and were significant at $p < .001$ (see Figure 1 and Table 3 for all standardized factor loadings). Thus, the two-factor model with three correlated residuals was retained.

Multiple-Group Analyses

A second CFA was rerun with S2 ($N = 4,981$) to confirm the fit of the two-factor model in the second sample and to serve as the baseline model for the subsequent multiple-group analyses. All obtained fit indices were similar to those for the CFA tested previously in S1 and in previous studies (Quirk et al., 2014), indicating overall adequate model fit for S2 (see Table 3 for factor loadings and Tables 4 and 5 for fit indices).

Gender

First, a multiple-group CFA was conducted to determine whether the two-factor model was invariant across males and females. Table 4 shows the fit indices for all steps of the multiple-group CFA for gender. In the first step, a CFA was run separately for each group of interest (males and females) and overall fit indices indicated that each model fit adequately for both males and females. Another model was then fit to males and females simultaneously to test configural invariance. Again, fit indices indicated adequate fit to the observed data. This served as the baseline result for subsequent, more stringent tests of measurement invariance. Results of the third step, metric invariance (factor loadings were constrained to equality), showed there was no significant increase in model misfit as compared to the configural invariance model when using the ΔCFI ($< .01$) as the indicator of significant change (see Table 4 for $\Delta\chi^2$ scores, statistical significance, and ΔCFI). Cheung and Rensvold (2002) recommend that the ΔCFI statistic be the primary metric examined when determining whether measurement invariance exists, particularly with large samples. Last, the model testing scalar invariance (loadings

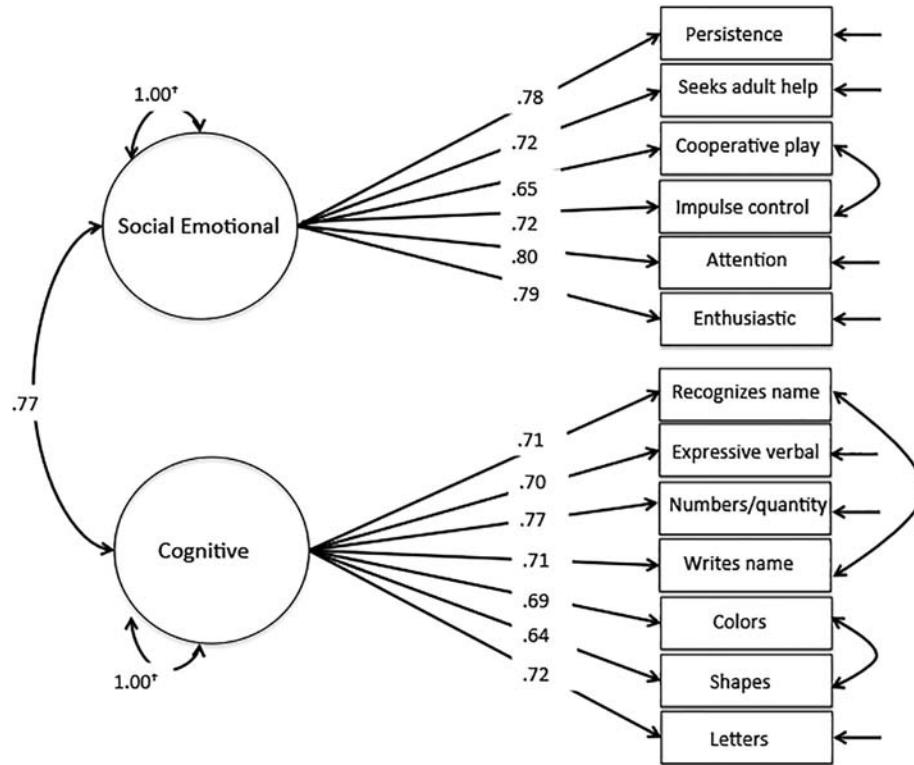


FIGURE 1 Standardized parameter estimates and factor correlations for general two-factor CFA model. All estimates were significant at the $p < .001$ level.

and intercepts held equal across groups) revealed no significant increase in model misfit after including these additional constraints ($\Delta CFI < .01$). Based upon these results, there was sufficient evidence to conclude that the KSEP’s factor structure was consistent across males and

females and that measurement invariance existed across gender.

Parent education

A second multiple-group CFA was conducted to test invariance of KSEP ratings across children with parents of differing education levels (high school diploma or less, some college or associate degree, college graduate or more). Table 5 shows the fit indices for all steps in the multiple-group CFA for parent education. In the first step of the multiple-group CFA, three CFAs were run separately for each group of interest (high school diploma or less, some college or associate degree, and college graduate or more). All three models showed adequate fit to the observed data. Next, configural invariance was tested by fitting a model that included all three parent education groups simultaneously—the fit indices in combination indicated adequate fit to the observed data. A subsequent model testing metric invariance revealed no significant increase in misfit for this model, as indicated by the ΔCFI , when compared with the results of the configural model. Next, scalar invariance was evaluated by testing a model that held loadings and intercepts equal across groups. Results indicated that scalar invariance criteria were met ($\Delta CFI < .01$; Cheung & Rensvold, 2002). Therefore, there was sufficient evidence to conclude that the factor structure of

TABLE 3
Standardized Factor Loadings for CFA of the Two-Factor Model in S1 and S2

Item	Social-emotional factor		Cognitive factor	
	Sample 1	Sample 2	Sample 1	Sample 2
1. Seeks adult help	.78	.79		
2. Cooperative play	.72	.72		
3. Impulse control	.65	.68		
4. Attention	.72	.76		
5. Enthusiastic	.80	.80		
6. Persistence	.79	.81		
7. Recognizes name			.71	.71
8. Expressive verbal			.70	.68
9. Numbers/quantity			.77	.78
10. Writes name			.71	.73
11. Colors			.69	.69
12. Shapes			.64	.67
13. Letters			.72	.72

Note. All factor loadings significant at $p < .001$.

TABLE 4
Tests of Measurement Invariance of the CFA Model in Females and Males (S2)

	χ^2	<i>df</i>	$\Delta\chi^2$	Δdf	<i>SRMR</i>	<i>RMSEA [90% CI]</i>	<i>CFI</i>	ΔCFI
<i>Single-group solutions</i>								
Overall	2419.447 ($p < .001$)	61	—	—	.046	.088 [.085, .091]	.938	—
Female ($n = 2,364$)	1342.814 ($p < .001$)	61	—	—	.048	.094 [.090, .099]	.928	—
Male ($n = 2,617$)	1175.739 ($p < .001$)	61	—	—	.046	.084 [.079, .088]	.943	—
<i>Measurement invariance</i>								
Configural	2518.552 ($p < .001$)	122	—	—	.047	.089 [.086, .092]	.936	—
Metric	2553.163 ($p < .001$)	133	34.611*	11	.049	.085 [.083, .088]	.936	.000
Scalar	2719.301 ($p < .001$)	144	166.138*	11	.054	.085 [.082, .088]	.931	.005

Note. $N = 4,981$. $\Delta\chi^2$, nested χ^2 difference. ΔCFI , nested CFI difference.

* $p < .001$.

TABLE 5
Tests of Measurement Invariance of the CFA Model for Parent Education Level (S2)

	χ^2	<i>df</i>	$\Delta\chi^2$	Δdf	<i>SRMR</i>	<i>RMSEA [90% CI]</i>	<i>CFI</i>	ΔCFI
<i>Single-group solutions</i>								
Overall sample ($N = 4,981$)	2419.447 ($p < .001$)	61	—	—	.046	.088 [.085, .091]	.938	—
HS diploma/less ($n = 1,442$)	732.029 ($p < .001$)	61	—	—	.050	.087 [.082, .093]	.935	—
Some college ($n = 1,045$)	554.718 ($p < .001$)	61	—	—	.046	.088 [.081, .095]	.933	—
College grad/higher ($n = 1,415$)	786.551 ($p < .001$)	61	—	—	.048	.092 [.086, .097]	.937	—
<i>Measurement invariance</i>								
Configural	2073.298 ($p < .001$)	183	—	—	.048	.089 [.086, .093]	.935	—
Metric	2109.759 ($p < .001$)	205	36.461*	22	.052	.085 [.081, .088]	.935	.000
Scalar	2201.051 ($p < .001$)	227	91.292**	22	.055	.082 [.079, .085]	.933	.002

Note. Total $n = 3,902$ for this analysis indicates that 1,079 students from S2 did not have parent education data. $\Delta\chi^2$, nested χ^2 difference. ΔCFI , nested CFI difference.

* $p < .05$.

** $p < .001$.

the KSEP was consistent across parent education groups and that measurement invariance existed across parent education level.

DISCUSSION

In this study, we tested the dimensionality and measurement invariance of a universal-level school readiness screening measure, the KSEP (Quirk et al., 2014). Examining invariance is (a) a necessary condition for a universal screening measure (Borsboom, 2006) and (b) a validation approach that is being increasingly applied to direct measures of cognitive school readiness (e.g., Csapó et al., 2014; Fuhs & Day, 2011). Our study also contributed to the literature by evaluating a universal-level measure that includes a social-emotional component, which is an essential element of children's school readiness (Blair, 2002).

To examine invariance, we tested a series of models with two independent subsamples to discern if a two-factor structure, encompassing children's social-emotional and cognitive readiness, held across multiple replications with

more socioculturally diverse samples than were used in previous studies. More importantly, we examined whether the two-dimensional structure of the KSEP was invariant across the variables of gender and parent education levels. Results provided evidence supporting the KSEP's two-dimensional structure, which is consistent with previous findings that used a sample predominantly comprised of Latino/a children, many of which were English learners (Quirk et al., 2014). The KSEP's two-dimensional structure is also consistent with other research that has indicated that school readiness is a multidimensional construct that includes nonacademic elements (e.g., social-emotional readiness) linked with children's holistic development (McWayne, Hahs-Vaughn, Cheung, & Wright, 2012; Sabol & Pianta, 2012). In addition, the findings of the present study provided psychometric evidence suggesting that the two-dimensional structure of the KSEP is invariant across both gender and parent education levels, allowing for valid interpretations of readiness data across these subpopulations of students, which is the typical circumstance for most school districts. Although not often examined, this form of psychometric evidence is particularly important for universal screening tools that are commonly used to

determine follow-up assessment or intervention procedures (Borsboom, 2006).

Implications for Practice

In addition to the psychometric contributions of this study, there are also multiple implications of these findings for educators and school psychologists. First, the finding that KSEP ratings measure the same readiness latent traits for boys and girls is particularly important, given that almost all schools and districts that would use the KSEP as a universal readiness screener will have male and female students. Although previous research has suggested that differences in readiness exist between boys and girls (Nylund-Gibson et al., 2014; Son, Lee, & Sung, 2013), the results of this study provide evidence that any differences found with the KSEP represent real differences in readiness rather than more complex differences on potentially different constructs (by gender), as might be found when measured by a tool without this form of evidence. Similarly, because of the well-documented association between school readiness and socioeconomic disadvantage (Kingston, Huang, Calzada, Dawson-McClure, & Brotman, 2013), the finding that the KSEP functions the same across parent education levels (a proxy for socioeconomic circumstances) provides critical validity evidence for schools and districts when interpreting readiness results across groups of children from various socioeconomic backgrounds.

The type of validity evidence provided by this study is particularly essential given the increased popularity of multitiered assessment procedures and the frequency with which schools, districts, and states are utilizing school readiness assessments to make important education decisions about incoming students. Absent measurement invariance evidence, it is impossible to discern whether differences in a measure's results across children from various subgroups are due to real differences on documentable latent traits or due to variations in how the measure functions across those subgroups (Milfont & Fischer, 2010). This is particularly critical in contexts where measures are used to make practically meaningful decisions about specific children, as is the case with school readiness screeners (e.g., placement, follow-up assessment, intervention, etc.).

Limitations

Although this study was conducted with an ethnically diverse sample of children, the demographics of our samples are not representative of those for classrooms across the United States. In particular, there was an overrepresentation of Latino/a students (45%) and an underrepresentation of White students (33%), with smaller disproportionalities for Asian (11%) and Black students

(6%). Although this might limit the generalizability of results based upon possible interactions between gender and parent education with ethnicity, the sample size across and within groups was large enough to provide robust estimates of factor structure, which were the central focus of this study. Additional research is needed to better understand how the KSEP functions with more nationally representative and regional samples.

Furthermore, we were only able to examine invariance across gender and parent education levels in the current study. Additional research is needed to better understand how the KSEP functions across other important subpopulation distinctions, such as English learners versus native English speakers, and ethnicity. For example, additional evidence is still needed to establish that the constructs measured at kindergarten are equivalently predictive of near- and short-term student outcomes for students from diverse backgrounds, which is what Millsap (2007) calls predictive invariance. Nonetheless, a first step in this direction was to establish measurement invariance for the fundamental variables of gender and parent education levels, as found in this study.

CONCLUSION

At this stage of the KSEP instrument's development and validation, the current study yielded essential findings that supplemented previous research evidence of its psychometric functioning and predictive validity. The replicated factor analyses in the present study provided additional independent evidence of the KSEP's construct validity and confirmed that it measures the same school readiness latent traits across genders and parental education levels. Evidence of invariance is particularly important, given that educators will likely use the KSEP to make important decisions regarding identification, further assessment, and possibly intervention. Previous research has also established that the KSEP ratings significantly positively predict students' later reading and math achievement (Quirk et al., 2013) and academic achievement trajectories through Grade 5 (Nylund-Gibson et al., 2014). Although this combination of findings indicates that KSEP ratings can be used as part of an assessment and decision-making process that provides children with early services and supports, it is likely that other factors affect children's academic trajectories. Hence, in practice there is a need to consider and use the KSEP within broader transactional ecological contexts (Pianta, Rimm-Kaufman, & Cox, 1999) that emphasize school readiness is based in the combinatorial effects of ready families, ready schools, and ready communities on each child's responsiveness to high-quality instruction in kindergarten (Rhode Island KID COUNTS, 2005; Sheridan, Marvin, Knoche, & Edwards, 2008).

ABOUT THE AUTHORS

Dr. Matthew Quirk and Dr. Michael Furlong are Professors in the Department of Counseling, Clinical, and School Psychology at the University of California, Santa Barbara, with research interests in the areas of assessment and student support. Ms. Ashley Mayworm and Ms. Jennica Rebelez are both doctoral students in the Department of Counseling, Clinical, and School Psychology at the University of California, Santa Barbara. Mr. Ryan Grimm is a doctoral student in the Department of Education at the University of California, Santa Barbara.

REFERENCES

- Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.
- Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist, 57*, 111–127. doi:10.1037/0003-066X.57.2.111
- Borsboom, D. (2006). When does measurement invariance matter? *Medical Care, 44*, 176–181.
- Brown, T. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford.
- Bulotsky-Shearer, R. J., & Fantuzzo, J. W. (2011). Preschool behavior problems in classroom learning situations and literacy outcomes in kindergarten and first grade. *Early Childhood Research Quarterly, 26*, 61–73. doi:10.1016/j.ecresq.2010.04.004
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233–255. doi:10.1207/S15328007SEM0902_5
- Chou, C., & Bentler, P. M. (1995). Estimates and tests in structural equation modeling. In R. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 37–55). Thousand Oaks, CA: Sage.
- Csapó, B., Molnár, G., & Nagy, J. (2014). Computer-based assessment of school readiness and early reasoning. *Journal of Educational Psychology, 106*, 639–650. doi:10.1037/a0035756
- 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. doi:10.1037/1082-989X.1.1.16
- Darling-Hammond, L. (2011, February). *The flat world and education: How America's commitment to equity will determine our future*. Keynote address at the meeting of the National Association of School Psychologists, San Francisco, CA.
- Denham, S. A., Warren-Khot, H., Bassett, H. H., Wyatt, T., & Perna, A. (2012). Factor structure of self-regulation in preschoolers: Testing models of a field-based assessment for predicting early school readiness. *Journal of Experimental Child Psychology, 111*, 386–404. doi:10.1016/j.jecp.2011.10.002
- DiPrete, T. A., & Jennings, J. L. (2012). Social and behavioral skills and the gender gap in early educational achievement. *Social Science Research, 41*, 1–15. doi:10.1016/j.ssresearch.2011.09.001
- Drouin, M., Horner, S. L., & Sondergeld, T. A. (2012). Alphabet knowledge in preschool: A Rasch model analysis. *Early Childhood Research Quarterly, 27*, 543–554. doi:10.1016/j.ecresq.2011.12.008
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., ... Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*, 1428–1446. doi:10.1037/0012-1649.43.6.1428
- Entwisle, D. R., & Hayduk, L. A. (1988). Lasting effects of elementary school. *Sociology of Education, 61*, 147–159. Retrieved from www.jstor.org/stable/2112624
- Foster, M. A., Lambert, R., Abbott-Shim, M., McCarty, F., & Franze, S. (2005). A model of home learning environment and social risk factors in relation to children's emergent literacy and social outcomes. *Early Childhood Research Quarterly, 20*, 13–36. doi:10.1016/j.ecresq.2005.01.006
- Fuhs, M. W., & Day, J. D. (2011). Verbal ability and executive functioning development in preschoolers at Head Start. *Developmental Psychology, 47*, 404–416. doi:10.1037/a0021065
- Furlong, M. J., & Quirk, M. (2011). The relative effects of chronological age on hispanic students' school readiness and grade 2 academic achievement. *Contemporary School Psychology, 15*, 81–92. Retrieved from <http://search.proquest.com/docview/888873786>
- Galindo, C., & Fuller, B. (2010). The social competence of Latino kindergartners and growth in mathematical understanding. *Developmental Psychology, 46*, 579–592. doi:10.1037/a0017821
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*, 1–55. doi:10.1080/10705519909540118
- Hulslander, J., Olson, R. K., Willcutt, E. G., & Wadsworth, S. J. (2010). Longitudinal stability of reading-related skills and their prediction of reading development. *Scientific Studies of Reading, 14*, 111–136. doi:10.1080/10888431003604058
- Kingston, S., Huang, K. Y., Calzada, E., Dawson-McClure, S., & Brotman, L. (2013). Parent involvement in education as a moderator of family and neighborhood socioeconomic context on school readiness among young children. *Journal of Community Psychology, 41*, 265–276. doi:10.1002/jcop.21528
- Lilles, E., Furlong, M. J., Quirk, M., Felix, E., Dominguez, K., & Anderson, M. (2009). Preliminary development of the Kindergarten Student Entrance Profile. *California School Psychologist, 14*, 71–80. doi:10.1007/BF03340952
- Matthews, J. S., Kizzie, K. T., Rowley, S. J., & Cortina, K. (2010). African Americans and boys: Understanding the literacy gap, tracing academic trajectories, and evaluating the role of learning-related skills. *Journal of Educational Psychology, 102*, 757–771. doi:10.1037/a0019616
- McWayne, C. M., Hahs-Vaughn, D. L., Cheung, K., & Wright, L. E. G. (2012). National profiles of school readiness skills for Head Start children: An investigation of stability and change. *Early Childhood Research Quarterly, 27*, 668–683. doi:10.1016/j.ecresq.2011.10.002
- Meredith, W. (1993). Measurement invariance, factor analysis, and factorial invariance. *Psychometrika, 58*, 525–543. doi:10.1007/BF02294825
- Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research, 3*, 111–121.
- Millsap, R. E. (2007). Invariance in measurement and prediction revisited. *Psychometrika, 72*, 461–473. doi:10.1007/s11336-007-9039-7
- Muthén, L. K., & Muthén, B. O. (2008–2013). *Mplus user's guide: Statistical analysis with latent variables* (6th ed.). Los Angeles, CA: Muthén & Muthén.
- Nylund-Gibson, K., Grimm, R., Quirk, M., & Furlong, M. J. (2014). Latent transition mixture model using the three-step specification. *Structural Equation Modeling, 21*, 1–16. doi:10.1080/10705511.2014.915375
- Obama, B. (2013). *The White House, Office of the Press Secretary. Remarks by the president in the state of union address*. Washington, DC. Retrieved from www.whitehouse.gov/the-press-office/2013/02/12/remarks-president-state-union-address
- Obama, B. (2014). *The White House, Office of the Press Secretary. Remarks by the president in the state of union address*. Washington, DC. Retrieved from www.whitehouse.gov/the-press-office/2014/01/28/remarks-president-state-union-address
- 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. doi:10.1177/1529100610381908
- Pianta, R., Rimm-Kaufman, S., & Cox, M. (1999). An ecological approach to conceptualizing the transition to kindergarten. In R. Pianta & M. Cox (Eds.), *The transition to kindergarten* (pp. 3–10). Baltimore, MD: Brookes.
- Quirk, M., Furlong, M. J., Lilles, E., Felix, E., & Chin, J. (2011). Preliminary development of a kindergarten school readiness assessment for Latino students. *Journal of Applied School Psychology*, 27, 77–102. doi:10.1080/15377903.2010.540518
- Quirk, M., Nylund-Gibson, K., & Furlong, M. (2013). Exploring patterns of Latino/a children's school readiness at kindergarten entry and their relations with Grade 2 achievement. *Early Childhood Research Quarterly*, 28, 437–449. doi:10.1016/j.ecresq.2012.11.002
- Quirk, M., Rebelez, J., & Furlong, M. J. (2014). Exploring the dimensionality of a brief school readiness screener for use with Latino/a children. *Journal of Psychoeducational Assessment*, 32, 259–264. doi:10.1177/0734282913505994
- Rhode Island KID COUNTS. (2005). *Getting ready: Findings from the National School Readiness Indicators Initiative: A 17 state partnership*. Retrieved from www.gettingready.org/matriarch/d.asp?PageID=303&PageName2=pdfhold&p=&PageName=Getting+Ready+%2D+Full+Report%2Epdf
- Rimm-Kaufman, S. E., & Pianta, R. C. (2000). An ecological perspective on the transition to kindergarten: A theoretical framework to guide empirical research. *Journal of Applied Developmental Psychology*, 21, 491–511. doi:10.1016/S0193-3973(00)00051-4
- Romano, E., Babchishin, L., Pagani, L. S., & Kohen, D. (2010). School readiness and later achievement: Replication and extension using a nationwide Canadian survey. *Developmental Psychology*, 46, 995–1007. doi:10.1037/a0018880
- Sabol, T. J., & Pianta, R. C. (2012). Patterns of school readiness forecast achievement and socioemotional development at the end of elementary school. *Child Development*, 83, 282–299. doi:10.1111/j.1467-8624.2011.01678.x
- Santa Maria–Bonita School District, First 5 of Santa Barbara County, & University of California Santa Barbara. (2005). *Kindergarten Student Entrance Profile*. Santa Barbara, CA: University of California Santa Barbara.
- Sheridan, S. M., Marvin, C., Knoche, L., & Edwards, C. P. (2008). Getting ready: Promoting school readiness through a relationship-based partnership model. *Early Childhood Services*, 2, 149–172. Retrieved from <http://digitalcommons.unl.edu/psychfacpub/566>
- Snow, K. (2006). Measuring school readiness: Conceptual and practical considerations. *Early Education and Development*, 17, 7–41. doi:10.1207/s15566935eed1701_2
- Son, S. H., Lee, K., & Sung, M. (2013). Links between preschoolers' behavioral regulation and school readiness skills: The role of child gender. *Early Education & Development*, 24, 468–490. doi:10.1080/10409289.2012.675548
- Stipek, D. J., & Ryan, R. H. (1997). Economically disadvantaged preschoolers: Ready to learn but further to go. *Developmental Psychology*, 33, 711–723. doi:10.1037/0012-1649.33.4.711
- Torgesen, J. K., & Burgess, S. R. (1998). Consistency of reading-related phonological processes throughout early childhood: Evidence from longitudinal-correlational and instructional studies. In J. Metsala & L. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 161–188). Mahwah, NJ: Erlbaum.
- Tsaousis, I., & Kazi, S. (2013). Factorial invariance and latent mean differences of scores on trait emotional intelligence across gender and age. *Personality and Individual Differences*, 54, 169–173. doi:10.1016/j.paid.2012.08.016