

Michael J. Strambler<sup>1</sup>, Clare W. Irwin<sup>2</sup>, Joanna L. Meyer<sup>1</sup>, and George A. Coleman<sup>3</sup>

<sup>1</sup>Yale School of Medicine

<sup>2</sup>Education Development Center, Inc.

<sup>3</sup>Cooperative Educational Services

## REPORT HIGHLIGHTS

### *Why this study?*

- The Connecticut Kindergarten Assessment Inventory (KEI) was developed to measure the skills children demonstrate at kindergarten entry across six domains (Language, Literacy, Numeracy, Physical/Motor, Creative/Aesthetic, and Personal/Social).
- PEER stakeholders wanted to know more about the performance of Connecticut children on the KEI and whether there were any group or district differences in performance.

### *Study Description*

- This study analyzed KEI data collected across four school years, from 2010/11 to 2013/14 (approximately 141,000 students).
- PEER compared KEI ratings among domains and over time.
- PEER also calculated each child's average KEI rating across all domains and compared average KEI scores across demographic groups and across district types.

### *Overall Trends*

- Teachers rated kindergarten students higher in creative/aesthetic and physical/motor skill domains than in other domains.
- Ratings of kindergarteners' skills were stable between 2010 and 2013.

### *Demographic differences*

- Overall, girls received higher ratings than boys. This gap was small, though somewhat larger than the national estimate.
- As a group, students who qualified for free or reduced-price lunch received lower skill ratings compared to those who did not qualify. The largest gap was for those qualifying for free lunch compared to those who were not eligible for free or reduced-price lunch. This gap was similar to national estimates of the gap between students from affluent and poor families.
- English learner students received lower skill ratings, on average, compared to those students who were considered proficient in English. The gap between English learners and English-proficient students was one of the largest demographics gaps for the timeframe of analysis. This gap was somewhat higher than national estimates.
- Compared to Whites, Black and Hispanic students, on average, showed larger gaps in skill ratings than did Asian students, and this gap was slightly higher than national estimates.

### *District Differences*

- On average, ratings of kindergarten entry skills were higher in smaller districts than larger districts.
- Kindergarteners were rated lower on entry skills in districts with high percentages of students who qualified for free lunch.

- Thirteen percent of the variability in students' skill ratings was accounted for by the grouping of students in school districts.

**Implications**

- Future analyses should explore the extent to which gaps represent actual differences in student skills versus inconsistencies or biases in teacher ratings.
- An examination of district-level characteristics that could be associated with KEI scores may yield valuable explanatory information. For example, the percentage of a district's students who are eligible for free lunch may be associated with the district's KEI scores.

**Background**

When the Partnership for Early Education Research (PEER) was formed in 2014, one of its main goals was to collaborate with early childhood stakeholders to develop a long-term [research agenda](#). Prior to establishing this agenda, PEER framed some [initial research questions](#) that would allow the Partnership to begin producing rigorous, actionable research that could inform policy and practice for children in the birth-to-eight age range. This brief focuses on one of these initial research aims, which was to understand the kindergarten readiness of children across Connecticut. PEER addressed this research aim by using data from a state-developed assessment tool, the Kindergarten Entrance Inventory (KEI).

**Goals of the study**

Using the KEI as an indicator of kindergarten readiness, this report is focused on addressing three main questions:

- (1) *What is the overall performance of children on the KEI?*
- (2) *Are there group differences in KEI scores?; and*
- (3) *Are there district differences in KEI scores?*

**The Connecticut Kindergarten Entrance Inventory (KEI)**

The KEI is a state-developed tool that was designed to provide teachers and other education stakeholders with information about what skills students demonstrate at entry into kindergarten. According to the Connecticut State Department of Education (CSDE), the “*Kindergarten Entrance Inventory was designed to provide a statewide snapshot of the skills students demonstrate, based on teachers’ observations, at the beginning of the kindergarten year.*”<sup>1</sup> Administration of the KEI involves teachers rating each of his/her students according to three performance levels. Table I indicates the skill domains and performance levels rated by teachers.

Table I. *KEI skill domains and performance levels*

<b>KEI Skill Domains</b>		
1. Language	2. Literacy	3. Numeracy
4. Physical/Motor	5. Creative/Aesthetic	6. Personal/Social
<b>KEI Performance Levels</b>		
<b>Level 1:</b> <i>Students at this level demonstrate emerging skills in the specified domain and require a large degree of instructional support.</i>		
<b>Level 2:</b> <i>Students at this level inconsistently demonstrate the skills in the specified domain and require some instructional support.</i>		
<b>Level 3:</b> <i>Students at this level consistently demonstrate the skills in the specified domain and require minimal instructional support.</i>		

The state required all public kindergarten teachers to complete the KEI each October between the school years of 2007/08 and 2013/14. Connecticut is currently in the process of preparing to launch a newly developed readiness assessment, the *Kindergarten Entrance Assessment*. The results from this brief are intended to provide a snapshot of kindergarten readiness in Connecticut, across time, using a well-established assessment. While direct comparisons to the future tool will not be possible, establishing a baseline understanding of the readiness of Connecticut's students and any gaps across groups of students may provide valuable information to the state as they roll out the new readiness assessment.



Photo by Scott Griffin, C.E.S.

### Strengths and Limitations of the KEI

Research has found evidence of the KEI's ability to predict students' kindergarten retention<sup>2</sup> as well as their academic performance in the third grade as measured by the Connecticut Mastery Test (CMT).<sup>3</sup> One potential limitation of the KEI relates to the consistency of the teachers' ratings of students. When using an observational measure like the KEI to compare readiness across classrooms and schools, teachers must assign ratings consistently. In other words, the teachers conducting the ratings must have a common understanding of the student skills and behaviors that the rating categories represent. One way to establish this consistency is through training raters to be consistent in how they assign ratings. In the case of the KEI, it is unclear the degree to which teachers were trained or if this training was consistent across schools or districts. Given that there is some evidence of validity and little evidence of reliability, it is important to avoid using the KEI as an accountability tool, since accountability procedures demand measures with strong evidence of multiple forms of validity and reliability. Despite its limitations, the evidence of the KEI's predictive validity, along with its broad use across the state for several years, make its results important to study.

### Caveats

When interpreting the findings of this report, there are two important cautions to note, which are addressed in more detail in Appendix C. First, most of the analyses in this brief treat the KEI as an interval measure rather than an ordered categorical measure. A key assumption of an interval measure is that each response category (level 1, level 2, and level 3) is equally spaced along a continuum. However, no psychometric analyses have been conducted to determine whether the measure is consistent with this assumption. Yet, our treatment of the data as interval is consistent with how CSDE and other researchers have summarized this scale in prior presentations<sup>4</sup> and peer-reviewed publications.<sup>2</sup>

Another important caveat relates to the use of national comparisons in this report. In this brief, national estimates of gap sizes are provided as a basis for approximate comparison, rather than as precise benchmarks. Although these national estimates are intended to show whether the gaps identified through these analyses are similar to the national-level gaps noted in other studies, the national estimates should be considered rough approximations rather than exact comparisons. No research has been conducted to verify whether or not the scores produced by the KEI are comparable to the scores produced by the measures used to calculate the national estimates of achievement gaps.

## What the study found

### Sample for this study

Analyses included all Connecticut kindergarteners from school years 2010/11 to 2013/14 for whom the CSDE provided KEI results and State Assigned Student Identifiers (SASIDs). Students without SASIDs were excluded from analysis to avoid potential duplicates. Students repeating kindergarten were included in the sample. The number of students without SASIDs (between 0 and 1.5 percent) and the number of repeaters (between 2.8 and 3.2 percent) were small. For more details on the sample, see Appendix B.

### Overall Trends

We begin by describing overall trends in readiness over time and across different domains. Table 2 displays the percentages of students at each performance level across the four years of analysis. To further examine readiness over time, we calculated average ratings across all six KEI domains. As figure 1 shows, there is little variability across the four years of data, with means ranging from 2.21 to 2.28. For more detailed results, see Appendix D. Values between Level 2 and Level 3 indicate that on average, students required a moderate level of instructional support.

Table 2. Distribution of KEI skill ratings in each domain

KEI Domain	Level 1	Level 2	Level 3
Language	22%	40%	38%
Literacy	23%	39%	38%
Numeracy	20%	41%	40%
Physical/Motor	11%	40%	50%
Creative/Aesthetic	11%	40%	50%
Personal/Social	16%	41%	43%

To examine readiness across domains, we calculated average ratings across time within each of the six KEI domains. As figure 2 shows, there was more variability in average ratings among the six domains than in average ratings across time, with mean scores ranging from 2.15 in literacy skills to 2.39 in creative/aesthetic skills. For more detailed results, see Appendix D. Teachers gave students somewhat higher ratings in the creative/aesthetic and physical/motor domains than they did in the pre-academic domains of language, literacy, and numeracy.

For the remaining analyses, we used average skill ratings across KEI domains as an indicator of global kindergarten readiness. Prior measurement studies support the approach of combining domains by suggesting that the KEI is best represented as a measure of a single domain of kindergarten readiness.<sup>5</sup>

Figure 1: Ratings of entry skills remain stable across years  
Mean Fall KEI Scores, 2010-2013

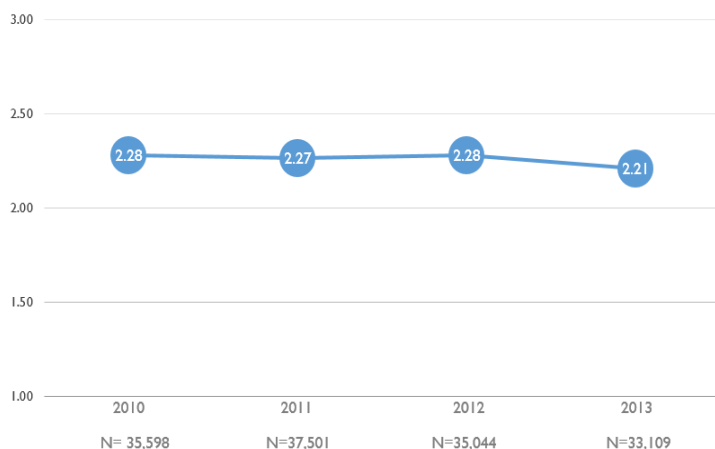
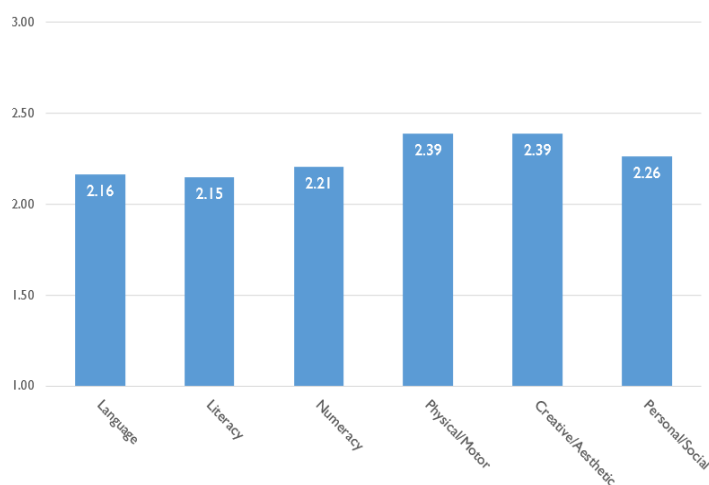


Figure 2: Children receive higher ratings on physical/motor and creative/aesthetic domains  
KEI Mean Scores, Combined 2010-2013



## Examining Group Differences in Readiness

Connecticut is often identified as a state with one of the largest achievement gaps in the country.<sup>6</sup> Research has shown that the achievement gap is often present when children are assessed at school entry<sup>7</sup>, but Connecticut does not administer a standardized achievement test at school entry. For this reason, we examined KEI data to learn about the size of any group differences in kindergarten entry skills. Specifically, we examined differences in KEI ratings by gender, socio-economic status, race/ethnicity, and English Learner status.



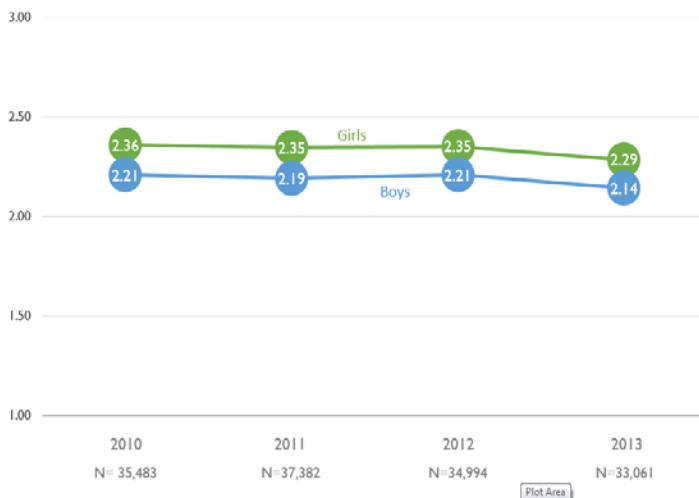
Photo by Scott Griffin, C.E.S.

Rather than representing the size of the difference between groups as a **mean difference**, we represent the size of the difference in terms of **standard deviation units**. Characterizing gaps in standard deviation units has a number of advantages over representing them as mean differences. The main advantage to looking at the gap in terms of standard deviation units is that it allows the gap to be directly compared to other gaps (or other kinds of effects), even with assessments that use different measures.

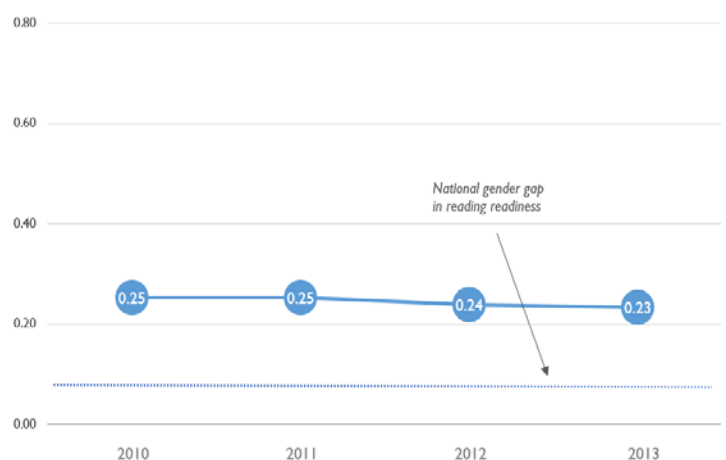
Such comparisons allow for a meaningful interpretation of the group differences. For example, it allows us to compare gaps in Connecticut KEI scores to gaps in scores from other states' kindergarten readiness assessments, assuming the assessments measure similar concepts. To provide this kind of context, we offer rough reference points on national group differences based on estimates obtained from other research on kindergarten readiness (see Appendix C for more information). These are not meant to be precise benchmarks, but rather approximate points for comparison.

**Gender Gaps.** The first graph below shows mean scores for boys and girls over the timeframe of the four years of this report. On average, girls received higher KEI ratings than did boys (see figure 3). When we frame these differences as effect sizes, small gaps are apparent, with standard deviation differences ranging from 0.18 to 0.25 (see figure 4). The largest gap of 0.25 indicates that the distribution of boys' and girls' KEI scores overlaps by about 90%. These small gaps are also in line with national estimates of differences between boys' and girls' kindergarten readiness. According to a report using a national dataset, girls score about 0.12 standard deviations higher than boys in reading (the difference in favor of girls for math is reported as "negligible").<sup>8</sup> Detailed results from these analyses are available in Appendix D.

**Figure 3: On average, girls receive higher KEI ratings than boys do**  
Mean KEI Scores for Boys and Girls, 2010-2013



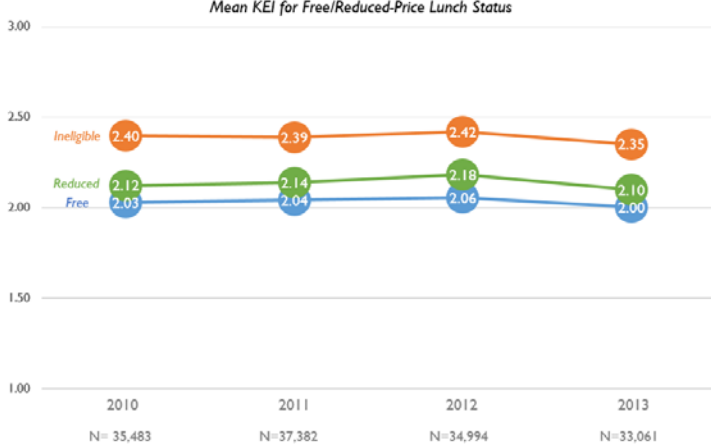
**Figure 4: KEI gender gap appears higher than national comparison**  
Gender Gap in KEI Scores in Standard Deviations, 2010-2013



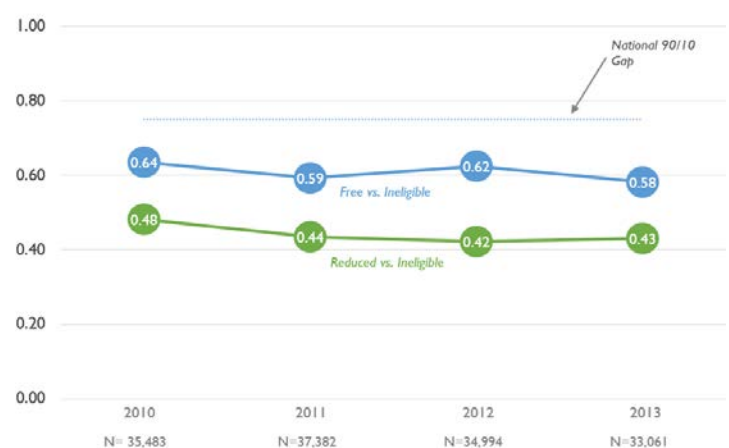
**Socio-Economic Gaps.** To compare students from different income backgrounds, we used a measure of whether a student was eligible for free lunch, reduced price lunch, or was ineligible for either. Districts use family income to determine eligibility for free or reduced-price lunch. This indicator is a commonly used measure of socio-economic status of students' families when more precise measures, such as family income, are unavailable.

Our analyses show that teachers gave higher ratings to students who did not qualify for free or reduced-price lunch, compared to those who qualified for reduced price or free lunch (figure 5). Additionally, they gave higher ratings to students who qualified for reduced price lunch, compared to those who qualified for free lunch. As shown in figure 6, the gap between free lunch students and ineligible students was largest and ranged from 0.54 to 0.64 standard deviations, whereas the gap between ineligible and reduced-price lunch students ranged from 0.38 to 0.44 standard deviations.

**Figure 5: Children with FRPL status are rated lower than those who are ineligible**



**Figure 6: KEI gap by FRPL status is comparable to national gap**



For a rough point of reference for interpreting these gaps, we used a study from Reardon and Portilla,<sup>9</sup> which combined three national datasets to assess socio-economic status and racial/ethnic readiness gaps. In that study, the national difference in scores on kindergarten readiness measures between students from poor and affluent backgrounds was 0.75 standard deviations, which is similar to the free-ineligible gap reported in this brief, but substantially larger than the reduced-ineligible gap. Again, detailed results from these analyses are available in Appendix D.

**English Learner Gap.** In this study, English learners were identified using an indicator for whether a student has been determined through a school assessment to have limited English proficiency. English learners received lower KEI ratings

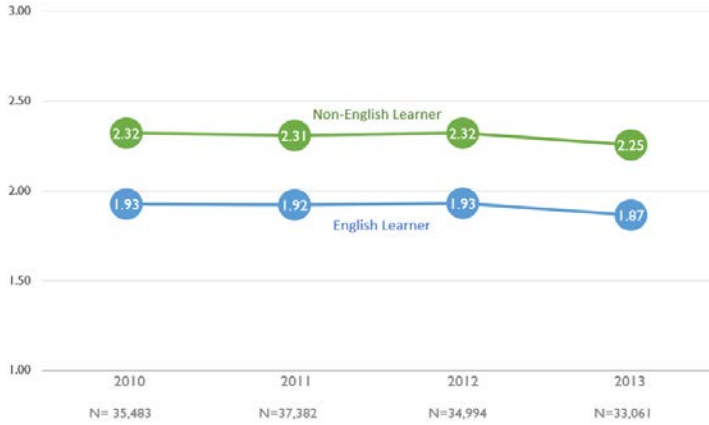


Photo by Scott Griffin. C.E.S.

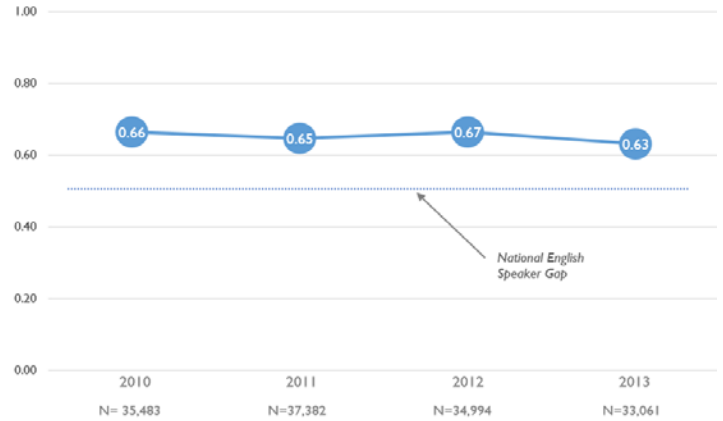
than their English-proficient counterparts (see figure 7), with sizeable gaps between the two groups (see figure 8). Specifically, the gap size ranges from 0.59 to 0.65 standard deviations. These gaps are somewhat larger than the national estimate of 0.50 standard deviation units for math and reading readiness estimates that compare students from English speaking home to those from non-English speaking homes.<sup>8</sup> Again, detailed results from these analyses are available in Appendix D.

**Racial/Ethnic Gaps.** We used family-reported race/ethnicity to report on overall readiness for these groups, and calculated the size of gaps for Hispanic, Black, and Asian students, compared to White students. The graph in figure 9 shows the average KEI scores across all four years for each racial/ethnic

**Figure 7: English Learners receive lower KEI ratings**  
Mean KEI Scores for English Learner Status



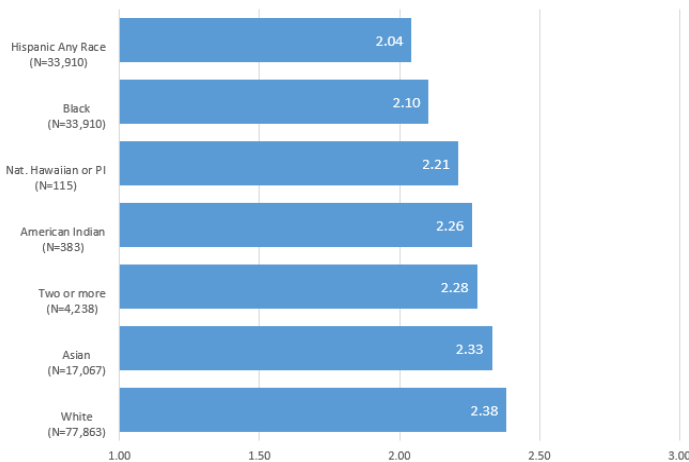
**Figure 8: Gap between English speakers and non-native English speakers is comparable to national gap**  
Gaps in KEI Scores in Standard Deviation Units



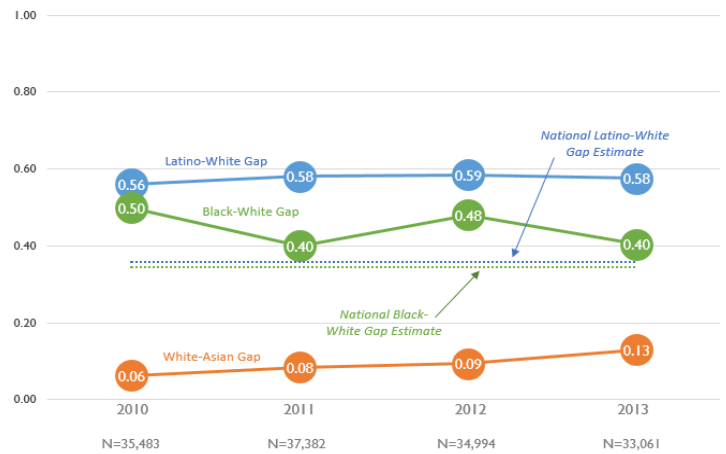
group. As shown in figure 9, Hispanic students received the lowest KEI scores with a mean of 2.04 (standard deviation = 0.62), and Whites received the highest scores with a mean of 2.38 (standard deviation = 0.57). The size of the Hispanic-White, Black-White, and Asian-White gaps are presented in figure 10 in standard deviation units. This figure shows that the gap was largest for Hispanic and Black students, with the Hispanic-White gap remaining stable over time and the Black-White gap slightly decreasing over time. The Asian-White gap was small in overall size and differed substantially from the gaps involving the other groups.

To compare the size of the racial/ethnic gaps in KEI scores to national estimates, we used the previously noted Reardon and Portilla study, which examines three national datasets.<sup>9</sup> The readiness gap from that study, averaged across domains, is 0.36 for the Hispanic-White gap and 0.35 for the Black-White gap. When using these comparisons, the data shows that the Connecticut racial/ethnic gaps were moderately larger than the national gaps. Detailed results from these analyses are available in Appendix D.

**Figure 9: Average KEI ratings vary across racial/ethnic groups**  
Mean KEI scores by group, combined 2010-2013



**Figure 10: KEI race/ethnicity gaps are largest for Latinos & Blacks**  
Race/Ethnicity Gaps in KEI Scores Measured in Standard Deviations



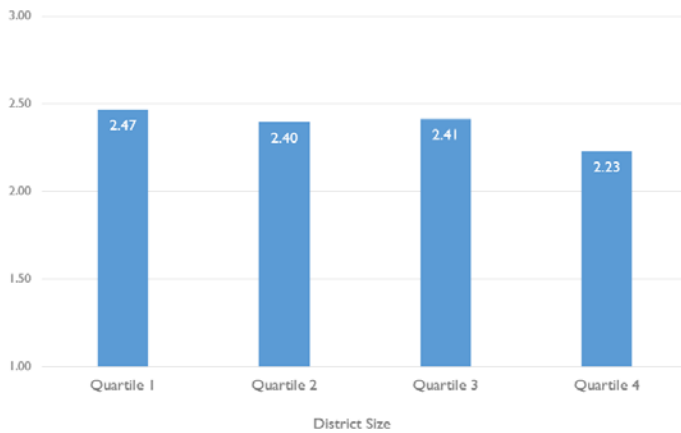
## District Characteristics and Readiness

To examine how district factors were associated with readiness scores, we first explored how KEI scores vary by the size of districts' kindergarten population. The first graph (figure 11) shows that teachers in the largest districts (quartile 4) rated students lower on the KEI than did teachers in smaller districts. The readiness gap between districts in the smallest and largest quartile was 0.41. See Appendix D for detailed results.

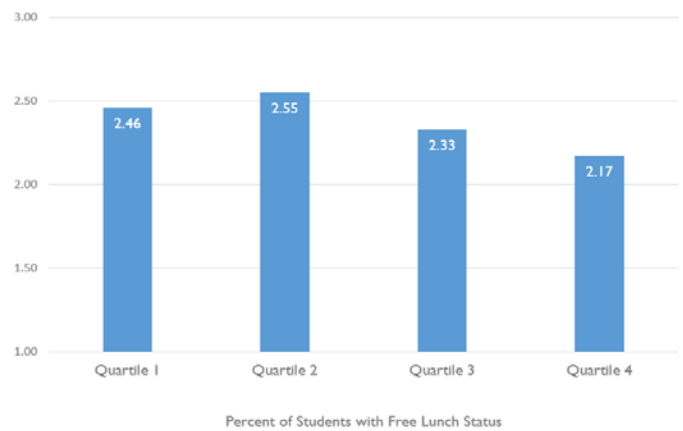
We also explored how KEI scores varied by the percentage of the kindergarten population that was eligible for free lunch. In districts with the highest percentage of kindergarten students who were eligible for free lunch (quartile 4), KEI ratings were lowest. The readiness gap between districts in the wealthiest and poorest quartile was 0.51.

We also assessed how much of the variability in KEI scores existed among individual students versus among groups of students defined by school district of attendance. To make this assessment, we used a multilevel modeling technique to produce a statistic called the **intraclass correlation coefficient (ICC)**<sup>10</sup>. This value was .13, indicating that 13 percent of the variability in students’ KEI scores can be accounted for by the grouping of students within districts. It is also important to note that other forms of grouping not assessed in this report, such as the grouping of students in classrooms, schools, and neighborhoods, likely contribute to variability in KEI scores as well. Given the meaningful variation in KEI scores across districts (see note in Appendix C)<sup>a</sup>, it may be worth exploring district-level characteristics<sup>b</sup> that could explain this variability. For example, it may be valuable to consider the accessibility, type, and quality of early childhood education available to each district’s children before kindergarten<sup>c</sup>.

**Figure 11: KEI ratings are lowest in the largest districts**  
District Mean KEI by District Kindergarten Size (Quartiles)



**Figure 12: KEI ratings are lower in poorer communities**  
District Mean KEI by % of Kindergarten Students with Free Lunch (Quartiles)



## Conclusion

This brief explored overall trends in students’ kindergarten entry skills as measured by the state-developed KEI, which involves teachers’ ratings of students’ skills at the start of kindergarten. Regarding general trends, teachers rated students

<sup>a</sup> There is some debate about the threshold at which an ICC value is considered meaningful. PEER believes that an ICC value of 0.13 is meaningful, although it may be smaller than similar values calculated in past research.

<sup>b</sup> As stated above, the term “district-level predictors” includes factors that are within and outside of the districts’ control, including factors that influence child performance prior to school entry. We refer to “district-level” predictors rather than “municipality-level” or “feeder town-level” predictors because the ICCs were calculated at the school district level (using a district identifier).

<sup>c</sup> Other possible sources of homogeneity that contribute to an ICC include the subjectivity of teacher ratings, the rating system itself, the items on the instrument, as well as shared characteristics of students grouped together within the same district.



higher in creative/aesthetic and physical/motor skill domains than preacademic domains, and overall skill levels were relatively stable over time.

This brief also explored differences in mean KEI scores among student demographic groups. Gender gaps were small and slightly above a national estimate. Gaps involving free and reduced lunch status were large, with the gap between students eligible for free lunch and students ineligible for free or reduced-price lunch being similar to a national estimate of poor versus affluent families. The gaps between English learner and English-proficient students was one of the largest demographic gaps in this study and was somewhat higher than a national estimate. For racial/ethnic differences, Blacks and Hispanics showed the largest gap in KEI ratings when compared to Whites, and this gap was slightly higher than national estimates.

Finally, the brief explored the association between district size and district socio-economic status with KEI scores. After separating districts into quartiles based on size and percent of students eligible for free and reduced-price lunch, teachers in small districts and districts with fewer students eligible for free and reduced-price lunch gave their students higher ratings than teachers did in large school districts and districts with more students eligible for free and reduced-price lunch. Given that 13 percent of the variance in entrance skills is accounted for by grouping at the district level, studying how district-level factors contribute to students' skills at kindergarten entry could be a fruitful next step.

An important consideration for the future examination of gaps in Connecticut entry skills is the extent to which gaps in teachers' ratings represented actual differences in students' abilities. It is possible that the ability to measure gaps with precision was hampered by inconsistencies across teachers' ratings of student behavior and/or biased perceptions about certain student groups. If training of raters is provided in the future for the current or new kindergarten entry assessment, this would be the ideal context in which to study these questions.

## Terminology

**Reliability** describes the accuracy of a measurement, in terms of its consistency or repeatability.

**Validity** describes the degree to which inferences based on scores are appropriate.

**Predictive validity** describes how strongly a measurement is related to expected outcomes.

**Inter-rater reliability** describes the degree to which different raters or observers agree on their measurements of the same phenomena.

**Mean difference** describes the difference between the mean (average) values for two or more groups.

**Standard deviation (SD)** describes the variability of a measure by expressing how much scores vary from the mean. One SD on either side of the mean encompasses 68% of scores in a normal distribution. Two SDs represents 95% of scores.

**Effect size in SD units** describes a standardized way of indicating how much groups differ on a measure by expressing this difference as a percentage of a standard deviation. This is accomplished through taking the mean difference between two groups and dividing it by the pooled standard deviation of the two groups. An effect size of .10 means that the two groups differ by .10 standard deviations. For more background, see Lakens (2013).

**Intraclass correlation coefficient (ICC)**, in the context of this report, describes how much variability in KEI scores is accounted for by the grouping of students in districts. This is expressed as a proportion of the variability at the district-level divided by the variability at the district level plus the variability at the individual level. For more background, see Peugh (2011) and McCoach & Adelson (2010).

## Acknowledgements

This work was supported by grant R305HI40142 from the U.S. Department of Education's Institute of Education Sciences (IES) under the Researcher-Practitioner Partnerships in Education Research program. The content of this brief is solely the responsibility of the authors and does not necessarily represent the official views of the Institute of Education Sciences (IES), the U.S. Department of Education, the Connecticut State Department of Education (CSDE), or the Connecticut Office of Early Childhood (OEC).

PEER would like to thank the CSDE representatives who made this study possible, including Ajit Golpalakrishnan, Alison Zhou, Kendra Shakir, Gilbert Andrada, and Charlene Russell-Tucker. PEER would also like to thank the OEC representatives who provided input on this study, including Linda Goodman, Harriet Feldlaufer, and Michelle Levy. Finally, PEER would like to thank Isiah Cruz, Student Research Assistant, Yale University, for his assistance in preparing this brief for publication and Scott Griffin, Public Relations and Marketing Specialist, Cooperative Educational Services, for the images in this brief.

This report is in the public domain. Permission to reproduce is not necessary. The report is available at: <http://PEER.yale.edu/publications> and should be cited as:

Strambler, M. J., Irwin, C. W., Meyer, J. L. & Coleman, G. A. *Assessing Kindergarten Entry Skills in Connecticut: The Kindergarten Entrance Inventory, 2010-2013*. New Haven, CT: Partnership for Early Education Research (PEER). April 2018.

## References

1. CT State Department of Education. *Fall Kindergarten Entrance Inventory*. 2017 June 1, 2016]; Available from: <http://www.csde.state.ct.us/public/csde/cedar/assessment/kindergarten/fall.htm>.
2. Goldstein, J., M. Eastwood, and P. Behuniak, *Can teacher ratings of students' skills at kindergarten entry predict kindergarten retention?* The Journal of Educational Research, 2014. **107**(3): p. 217-229.
3. Goldstein, J. and D.B. McCoach, *The predictive validity of kindergarten readiness judgments: Lessons from one state*. Journal of educational research, 2017. **110**(1): p. 50-60.
4. Adesso, K. and J. Goldstein. *Connecticut's Fall Kindergarten Inventory: Policy and Technical Issues*. 2010 March 2017]; Available from: <http://assessment.education.uconn.edu/assessment/assets/Conferences/2010-CAF/CAF%20Kindergarten%20Inventory%20Addesso%20Goldein.pdf>.
5. Goldstein, J. and D.B. McCoach, *The Starting Line: Developing a Structure for Teacher Ratings of Students' Skills at Kindergarten Entry*. Early Childhood Research & Practice, 2011. **13**(2): p. n2.
6. Connecticut Commission on Education Achievement. *Every child should have chance to be exceptional. Without exception: A plan to help close Connecticut's achievement gap*. 2010; Available from: [http://ctedreform.org/wp-content/uploads/2012/11/CCER\\_2012Report.pdf](http://ctedreform.org/wp-content/uploads/2012/11/CCER_2012Report.pdf).
7. Burchinal, M., et al., *Examining the Black-White achievement gap among low-income children using the NICHD study of early child care and youth development*. Child development, 2011. **82**(5): p. 1404-1420.
8. Nores, M. and W.S. Barnett, *Access to high quality early care and education: Readiness and opportunity gaps in america (ceelo policy report)*. 2014, Center on Enhancing Early Learning Outcomes: New Brunswick, NJ.
9. Reardon, S.F. and X.A. Portilla, *Recent trends in income, racial, and ethnic school readiness gaps at kindergarten entry*. AERA Open, 2016. **2**(3): p. 2332858416657343.
10. Peugh, J.L., *A practical guide to multilevel modeling*. Journal of school psychology, 2010. **48**(1): p. 85-112.
11. Lakens, D., *Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs*. Frontiers in psychology, 2013. **4**.
12. Hedges, L.V. and E.C. Hedberg, *Intraclass correlation values for planning group-randomized trials in education*. Educational Evaluation and Policy Analysis, 2007. **29**(1): p. 60-87.

## Appendix A: Fall Kindergarten Entrance Inventory

The following Performance Level (PL) Literals describe the characteristics of a typical student at each performance level. These will be used to rate each student on each of the six domains.

**Performance Level 1:** *Students at this level demonstrate emerging skills in the specified domain and require a large degree of instructional support.*

**Performance Level 2:** *Students at this level inconsistently demonstrate the skills in the specified domain and require some instructional support.*

**Performance Level 3:** *Students at this level consistently demonstrate the skills in the specified domain and require minimal instructional support.*

---

**Directions:** The indicators listed below each domain are examples of the skills a student should be able to demonstrate at the beginning of the kindergarten year; however, these are not the only skills to be considered. Rate each student in your class on each of the six domains. Use the Performance Levels (PL) above and all available and pertinent information when rating a student.

---

### Language Skills

PL Rating

At what level does the student:

- Participate in conversations
- Retell information from a story read to him/her
- Follow simple two-step verbal directions
- Speak using sentences of at least 5 words
- Communicate feelings and needs
- Listen attentively to a speaker

---

### Literacy Skills

PL Rating

At what level does the student:

- Hold a book and turn pages from the front to the back
- Understand that print conveys meaning
- Explore books independently
- Recognize printed letters, especially in their name and familiar printed words
- Match/connect letters and sounds
- Identify some initial sounds
- Demonstrate emergent writing

### Numeracy Skills

PL Rating

At what level does the student:

- Count to 10
- Demonstrate one-to-one correspondence while counting (e.g., touches objects as he/she counts)
- Measure objects using a variety of everyday items
- Identify simple shapes such as circles, squares, rectangles, and triangles

- Identify patterns
- Sort and group objects by size, shape, function (use), or other attributes
- Understand sequence of events (e.g., before, after, yesterday, today, or tomorrow)

---

### Physical/Motor Skills

PL Rating

At what level does the student:

- Run, jump, or balance
- Kick or throw a ball, climb stairs or dance
- Write or draw using writing instruments (e.g., markers, chalk, pencils, etc.)
- Perform tasks, such as completing puzzles, stringing beads, or cutting with scissors

---

### Creative/Aesthetic Skills

PL Rating

At what level does the student:

- Draw, paint, sculpt, or build to represent experiences
- Participate in pretend play
- Enjoy or participate in musical experiences (e.g., singing, clapping, drumming, or dancing)

---

### Personal/Social Skills

PL Rating

At what level does the student:

- Engage in self-selected activities
- Interact with peers to play or work cooperatively
- Use words to express own feelings or to identify conflicts
- Seek peer or adult help to resolve a conflict
- Follow classroom routines

## Appendix B: Analysis Sample for Connecticut Kindergarteners 2010-2013

Following the development and approval of a data sharing agreement between Yale University and the Connecticut State Department of Education (CSDE), the PEER team acquired four years of data from CSDE. These data came from two data management systems, one that contains KEI scores and one that contains student and school characteristics. We merged these datasets across the four years to produce the dataset used for the analyses in this report.

**State-assigned student identifiers.** We focused our analyses on students who had state-assigned student identifiers (SASIDs) to avoid potential duplicates and to allow us to assess kindergarten repeaters (see below).

**Repeaters.** The sample for this study includes some students who repeated kindergarten. We used each student's SASID to identify whether he or she appeared in kindergarten across two different time points. Specifically, we examined whether the student appeared in school years 2010/11 and 2011/12, 2011/12 and 2012/13, or 2012/13 and 2013/14. For the sake of methodological consistency, we included repeating kindergartners in the sample because we could not remove them for the first year of data (to do so would require having data from the prior year, or limiting the analysis to three years, instead of four years). See Table B.1 for the aforementioned values.

Table B.1. *Students repeating kindergarten*

Time	Status	Frequency	Percent
2010 & 2011	Repeaters	1,137	3.0
	Non-repeaters	36,364	97.0
	Total	37,501	100.0
2011 & 2012	Repeaters	1,112	3.2
	Non-repeaters	33,932	96.8
	Total	35,044	100.0
2012 & 2013	Repeaters	1,078	2.8
	Non-repeaters	32,031	97.2
	Total	33,109	100.0

Table B.2. *Students with missing State Assigned Student Identifiers (SASID)*

Year	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
2010	39099	98.5%	583	1.5%	39682	100.0%
2011	37501	100.00%	0	0%	37501	100.0%
2012	38422	98.9%	418	1.1%	38840	100.0%
2013	38255	98.9%	420	1.1%	38675	100.0%

## Appendix C: Study Methodology

### Readiness gaps

In analyzing KEI results, the main aim was to assess the size of any differences in ratings of kindergarten entry skills among groups by gender, socio-economic status, race/ethnicity, and English learner status. We assessed the size of such differences in terms of standard deviation units, a type of measure commonly referred to as an *effect size*. Characterizing gaps in this way has a number of advantages over representing gaps as mean differences. The main advantage of looking at the gap as a standard deviation is that it allows the gap to be directly compared to other gaps (or other kinds of effects), even with assessments that use different measures. Such comparisons allow for a meaningful interpretation of the group differences. For example, it allows one to compare gaps in Connecticut KEI scores to gaps in scores from other states' kindergarten readiness assessments, assuming the assessments measure similar concepts.

This report uses the Cohen's  $d_s$  measure (sometimes referred to as Cohen's  $g$ ) to assess KEI gaps between groups. This measure is derived by taking the mean difference between the groups and dividing it by the pooled standard deviation of the two groups. Specifically, the following equation is used:

$$d_s = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{(n_1 - 1)SD_1^2 + (n_2 - 1)SD_2^2}{n_1 + n_2 - 2}}}$$

In this equation,  $X_1$  and  $X_2$  represent the means (averages) of each group;  $n_1$  and  $n_2$  represent the sample sizes of each group, and  $SD_1$  and  $SD_2$  represent the standard deviations of each group. See Lakens (2013)<sup>11</sup> for more detail on the Cohen  $d_s$  measure and other measures of effect sizes.

### National Estimates

To provide context to which gaps in Connecticut's KEI scores can be compared, this report offers rough reference points on national group differences based on estimates obtained from other research on kindergarten readiness. As we noted in the report but cannot stress enough, these values are not meant to be precise benchmarks, but rather reasonable points for comparison.

For racial/ethnic and socio-economic gaps, we use estimates from Reardon & Portilla (2016). Table C.1 reports the national estimates from the Reardon & Portilla study. It is important to note that the estimates reported in Reardon and Portilla range (sometimes widely) depending on the domain of assessment (for example, math vs. approaches to learning) and the assessment respondent (student vs. teacher). For this report, we averaged across these different assessments to approximate our approach of averaging across the six KEI domains. While we acknowledge the flaws in this approach due to the measurement variability in the Reardon & Portilla estimates mentioned above, we believe this comparison to provide more meaningful interpretation than other common approaches such as interpreting effects by way of general "small, medium, large" conventions that lack specific context.

For DLL and gender gaps, we use estimates from Nores & Barnett (2014), which analyzes data from the Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 (ECLS-K). The authors report the gap in 2010 math and English test scores for students from non-English speaking homes compared to those from English speaking homes.

Table C.1. National estimates of group differences reported in Reardon & Portilla (2016)

Domain	Assessment	Black-White Gap	White-Hispanic Gap	Income 90/10 Gap
Math	Direct measure of students' math skill	0.547	0.672	1.172
Reading	Direct measure of students' reading skill	0.319	0.559	1.056
Self-control	Teacher-reported measure of students' ability to control behavior and emotions	0.32	0.09	0.527
Approaches to learning	Teacher-reported measure of behaviors that promote learning	0.269	0.11	0.58
Externalizing	Teacher-reported measure of students' acting-out behaviors	0.288	-0.027	0.412
PEER calculated total mean		0.349	0.358	0.749

### District Analyses

Since the size of districts' kindergarten classes varied across the four years of analysis, we used the largest size over this time period as an indicator of district size. On page 8, we note that the ICC of .13, which indicates how much of the variability of KEI scores is attributable to the grouping of student in districts, is meaningful. We made this judgement based on estimates from the research literature indicating an average national average ICC of .23 for reading achievement among kindergarten students.<sup>12</sup> While this national average value is larger than the value reported in this brief, we judged the ICC of .13 to be meaningful in size.

## Appendix D: Supplementary Tables

### Descriptive statistics

Table D.1: Descriptive statistics for 2010-2013 KEI data

	Time	Language	Literacy	Numeracy	Physical/ Motor	Creative/ Aesthetic	Personal/ Social
Fall 2010	Mean	2.18	2.17	2.23	2.40	2.41	2.29
	N	35598	35598	35598	35598	35598	35598
	Std. Deviation	0.758	0.763	0.738	0.672	0.668	0.717
	Std. Error of Mean	0.004	0.004	0.004	0.004	0.004	0.004
Fall 2011	Mean	2.18	2.15	2.21	2.40	2.39	2.28
	N	37501	37501	37501	37501	37501	37501
	Std. Deviation	0.757	0.766	0.742	0.666	0.668	0.716
	Std. Error of Mean	0.004	0.004	0.004	0.003	0.003	0.004
Fall 2012	Mean	2.18	2.17	2.22	2.41	2.41	2.28
	N	35044	35044	35044	35044	35044	35044
	Std. Deviation	0.753	0.760	0.738	0.660	0.660	0.711
	Std. Error of Mean	0.004	0.004	0.004	0.004	0.004	0.004
Fall 2013	Mean	2.11	2.10	2.16	2.34	2.34	2.21
	N	33109	33109	33109	33109	33108	33109
	Std. Deviation	0.766	0.773	0.754	0.688	0.690	0.732
	Std. Error of Mean	0.004	0.004	0.004	0.004	0.004	0.004
Total	Mean	2.16	2.15	2.21	2.39	2.39	2.26
	N	141252	141252	141252	141252	141251	141252
	Std. Deviation	0.759	0.766	0.743	0.672	0.672	0.719
	Std. Error of Mean	0.002	0.002	0.002	0.002	0.002	0.002



## KEI performance levels over time by domain (counts and frequencies)

Table D.2a: Distribution of KEI ratings across Language performance levels, over time

			Time				Total
			Fall 2010	Fall 2011	Fall 2012	Fall 2013	
Language	1	Count	7602	7983	7393	8038	31016
		% within Time	21.4%	21.3%	21.1%	24.3%	22.0%
	2	Count	13964	14832	14107	13242	56145
		% within Time	39.2%	39.6%	40.3%	40.0%	39.7%
	3	Count	14032	14686	13544	11829	54091
		% within Time	39.4%	39.2%	38.6%	35.7%	38.3%
Total	Count	35598	37501	35044	33109	141252	
	% within Time	100.0%	100.0%	100.0%	100.0%	100.0%	

Table D.2b: Literacy levels over time

			Time				Total
			Fall 2010	Fall 2011	Fall 2012	Fall 2013	
Literacy	1	Count	7816	8625	7597	8402	32440
		% within Time	22.0%	23.0%	21.7%	25.4%	23.0%
	2	Count	13840	14703	13733	12993	55269
		% within Time	38.9%	39.2%	39.2%	39.2%	39.1%
	3	Count	13942	14173	13714	11714	53543
		% within Time	39.2%	37.8%	39.1%	35.4%	37.9%
Total	Count	35598	37501	35044	33109	141252	
	% within Time	100.0%	100.0%	100.0%	100.0%	100.0%	

Table D.2c: Numeracy levels over time

			Time				Total
			Fall 2010	Fall 2011	Fall 2012	Fall 2013	
Numeracy	1	Count	6506	7243	6517	7232	27498
		% within Time	18.3%	19.3%	18.6%	21.8%	19.5%
	2	Count	14260	15266	14237	13448	57211
		% within Time	40.1%	40.7%	40.6%	40.6%	40.5%
	3	Count	14832	14992	14290	12429	56543
		% within Time	41.7%	40.0%	40.8%	37.5%	40.0%
Total	Count	35598	37501	35044	33109	141252	
	% within Time	100.0%	100.0%	100.0%	100.0%	100.0%	

Table D.2d: Physical/Motor levels over time

			Time				Total
			Fall 2010	Fall 2011	Fall 2012	Fall 2013	
Physical	1	Count	3751	3830	3392	4102	15075
		% within Time	10.5%	10.2%	9.7%	12.4%	10.7%
	2	Count	13749	14994	13845	13489	56077
		% within Time	38.6%	40.0%	39.5%	40.7%	39.7%
	3	Count	18098	18677	17807	15518	70100
		% within Time	50.8%	49.8%	50.8%	46.9%	49.6%
Total	Count	35598	37501	35044	33109	141252	
	% within Time	100.0%	100.0%	100.0%	100.0%	100.0%	

Table D.2e: Creative/Aesthetic levels over time

			Time				Total
			Fall 2010	Fall 2011	Fall 2012	Fall 2013	
Creative Aesthetic	1	Count	3639	3906	3384	4163	15092
		% within Time	10.2%	10.4%	9.7%	12.6%	10.7%
	2	Count	13625	15016	13742	13457	55840
		% within Time	38.3%	40.0%	39.2%	40.6%	39.5%
	3	Count	18334	18579	17918	15488	70319
		% within Time	51.5%	49.5%	51.1%	46.8%	49.8%
Total	Count	35598	37501	35044	33108	141251	
	% within Time	100.0%	100.0%	100.0%	100.0%	100.0%	

Table D.2f: Personal/Social levels over time

			Time				Total
			Fall 2010	Fall 2011	Fall 2012	Fall 2013	
Personal	1	Count	5488	5853	5356	6139	22836
		% within Time	15.4%	15.6%	15.3%	18.5%	16.2%
	2	Count	14299	15393	14677	13954	58323
		% within Time	40.2%	41.0%	41.9%	42.1%	41.3%
	3	Count	15811	16255	15011	13016	60093
		% within Time	44.4%	43.3%	42.8%	39.3%	42.5%
Total	Count	35598	37501	35044	33109	141252	
	% within Time	100.0%	100.0%	100.0%	100.0%	100.0%	

## KEI Group Differences

Table D.3a: Descriptive statistics for KEI ratings over time by gender

Time		Mean	N	Std. Deviation	Std. Error of Mean
Fall 2010	Female	2.361	17255	0.586	0.004
	Male	2.209	18228	0.613	0.005
	*	2.054	115	0.610	0.057
	Total	2.282	35598	0.605	0.003
Fall 2011	Female	2.346	17977	0.588	0.004
	Male	2.194	19405	0.615	0.004
	*	2.060	119	0.619	0.057
	Total	2.266	37501	0.607	0.003
Fall 2012	Female	2.353	16937	0.587	0.005
	Male	2.209	18057	0.614	0.005
	*	2.143	50	0.617	0.087
	Total	2.279	35044	0.605	0.003
Fall 2013	Female	2.287	15907	0.613	0.005
	Male	2.142	17154	0.631	0.005
	*	1.851	48	0.653	0.094
	Total	2.211	33109	0.627	0.003
Total	Female	2.338	68076	0.594	0.002
	Male	2.189	72844	0.619	0.002
	*	2.040	332	0.624	0.034
	Total	2.260	141252	0.612	0.002

\* Gender not provided

Table D.3b: Descriptive statistics for KEI ratings over time by FRPL Eligibility

Time		Mean	N	Std. Deviation	Std. Error of Mean
Fall 2010	Free	2.029	10094	0.616	0.006
	Ineligible	2.397	24143	0.564	0.004
	Reduced	2.124	1246	0.615	0.017
	*	2.054	115	0.610	0.057
	Total	2.282	35598	0.605	0.003
Fall 2011	Free	2.041	12126	0.613	0.006
	Ineligible	2.388	23949	0.569	0.004
	Reduced	2.140	1307	0.606	0.017
	*	2.060	119	0.619	0.057
	Total	2.266	37501	0.607	0.003
Fall 2012	Free	2.058	12642	0.610	0.005
	Ineligible	2.419	20968	0.559	0.004
	Reduced	2.182	1384	0.609	0.016
	*	2.143	50	0.617	0.087
	Total	2.279	35044	0.605	0.003

Time		Mean	N	Std. Deviation	Std. Error of Mean
Fall 2013	Free	2.002	12395	0.626	0.006
	Ineligible	2.353	19415	0.587	0.004
	Reduced	2.099	1251	0.630	0.018
	*	1.851	48	0.653	0.094
	Total	2.211	33109	0.627	0.003
Total	F	2.033	47257	0.617	0.003
	N	2.390	88475	0.570	0.002
	Reduced	2.138	5188	0.615	0.009
	*	2.040	332	0.624	0.034
	Total	2.260	141252	0.612	0.002

\* FRPL Eligibility not provided

Table D.3c: Descriptive statistics for KEI ratings over time by ELL status

Time		Mean	N	Std. Deviation	Std. Error of Mean
Fall 2010	NonELL	2.320	32108	0.594	0.003
	ELL	1.926	3375	0.593	0.010
	*	2.054	115	0.610	0.057
	Total	2.282	35598	0.605	0.003
Fall 2011	NonELL	2.307	33512	0.595	0.003
	ELL	1.921	3870	0.598	0.010
	*	2.060	119	0.619	0.057
	Total	2.266	37501	0.607	0.003
Fall 2012	NonELL	2.323	31053	0.592	0.003
	ELL	1.929	3941	0.592	0.009
	*	2.143	50	0.617	0.087
	Total	2.279	35044	0.605	0.003
Fall 2013	NonELL	2.254	29420	0.615	0.004
	ELL	1.866	3641	0.615	0.010
	*	1.851	48	0.653	0.094
	Total	2.211	33109	0.627	0.003
Total	NonELL	2.302	126093	0.599	0.002
	ELL	1.911	14827	0.600	0.005
	*	2.040	332	0.624	0.034
	Total	2.260	141252	0.612	0.002

\* ELL status not provided

Table D.3d: Descriptive statistics for KEI ratings over time by race/ethnicity

Time		Mean	N	Std. Deviation	Std. Error of Mean
Fall 2010	White	2.39	21236	0.57	0.00
	Latino/a of any race	2.06	7371	0.62	0.01
	Black	2.10	3991	0.62	0.01
	Asian	2.35	1927	0.60	0.01
	American Indian or Alaska Native	2.32	100	0.57	0.06
	Native Hawaiian or Other Pacific Islander	2.29	34	0.60	0.10
	Two or More Races	2.31	824	0.60	0.02
	Total	2.28	35483	0.60	0.00
Fall 2011	White	2.39	20701	0.57	0.00
	Latino/a of any race	2.05	8946	0.61	0.01
	Black	2.09	4613	0.61	0.01
	Asian	2.34	1851	0.62	0.01
	American Indian or Alaska Native	2.25	108	0.64	0.06
	Native Hawaiian or Other Pacific Islander	2.22	23	0.66	0.14
	Two or More Races	2.26	1140	0.60	0.02
	Total	2.27	37382	0.61	0.00
Fall 2012	White	2.41	18614	0.57	0.00
	Latino/a of any race	2.07	8952	0.61	0.01
	Black	2.13	4240	0.61	0.01
	Asian	2.35	1899	0.59	0.01
	American Indian or Alaska Native	2.23	99	0.60	0.06
	Native Hawaiian or Other Pacific Islander	2.15	30	0.63	0.12

Time		Mean	N	Std. Deviation	Std. Error of Mean
	Two or More Races	2.30	1160	0.60	0.02
	Total	2.28	34994	0.61	0.00
Fall 2013	White	2.34	17312	0.59	0.00
	Latino/a of any race	1.99	8641	0.63	0.01
	Black	2.10	4223	0.62	0.01
	Asian	2.26	1667	0.64	0.02
	American Indian or Alaska Native	2.24	76	0.61	0.07
	Native Hawaiian or Other Pacific Islander	2.17	28	0.51	0.10
	Two or More Races	2.24	1114	0.63	0.02
	Total	2.21	33061	0.63	0.00
Total	White	2.38	77863	0.57	0.00
	Latino/a of any race	2.04	33910	0.62	0.00
	Black	2.10	17067	0.62	0.00
	Asian	2.33	7344	0.61	0.01
	American Indian or Alaska Native	2.26	383	0.61	0.03
	Native Hawaiian or Other Pacific Islander	2.21	115	0.60	0.06
	Two or More Races	2.28	4238	0.61	0.01
	Total	2.26	140920	0.61	0.00

## District Differences

Table D.4a: Mean KEI ratings by district size in quartiles

	Mean	N	Std. Deviation
Quartile 1	2.47	44	0.561
Quartile 2	2.40	43	0.536
Quartile 3	2.41	44	0.570
Quartile 4	2.23	43	0.583
Total	2.37	174	0.565

Table D.4b: Mean KEI ratings by % free lunch in quartiles

	Mean	N	Std. Deviation
Quartile 1	2.46	43	0.522
Quartile 2	2.55	44	0.491
Quartile 3	2.33	44	0.596
Quartile 4	2.17	43	0.589
Total	2.37	174	0.565