



## PA Keystone Exams: Analysis of 2013-14 Publicly-Available Data

June 2015

Kate Shaw ▾ Michael H. Norton ▾ John Sludden  
Daniel Long ▾ Marvin Barnes

### Overview

Pennsylvania's two-decade experience with standards-based education reform has been marked by an increasing reliance on test-based accountability. In the last five years, state assessments have become instrumental components in new teacher and principal evaluations and school building ratings. And beginning with the class of 2017, student performance on tests in three subjects—Algebra I, Biology, and Literature—will play a major role in whether students receive high school diplomas.

The state's move toward high-stakes exit exams, or Keystones, evolved over the last decade. Discussions began in 2006,<sup>1</sup> were codified in regulation in 2010, and have undergone multiple revisions since. Throughout this period, the policy has been supported by governors of both parties but heavily criticized by members of the legislature, education organizations, school districts, and civil rights groups. Today, nearly 10 years after the proposal was first outlined, opposition persists.<sup>2</sup>

Views on exit exams usually follow two tracks. Supporters believe exit exams will ensure greater equity in academic expectations statewide.<sup>3</sup> Opponents contend the exit exams will narrow curriculum, unduly burden both teachers and students, and negatively impact graduation rates—especially for disadvantaged students.<sup>4, 5</sup>

Test results for the 2013-14 administration of the Keystones were released by the state earlier this year; the results provide an initial indication of how students are performing on the exams. Research for Action's (RFA's) analysis of these data found:

- The percentage of students scoring proficient on their first attempts on the Keystone exams ranged from 52 (Algebra I) to 62 percent (Literature). For all subjects, fewer than a quarter of re-tests resulted in passing scores.
- Traditionally-disadvantaged students experienced particular difficulty: 31 percent of economically disadvantaged students were proficient on a first administration of the Algebra I Keystone; one in three first-time test-takers reached proficiency on the Biology Keystone.

<sup>1</sup> In prior roles, several RFA staff were involved in the development of the Rendell Administration's graduation requirements policy.

<sup>2</sup> Boccella, K. (Nov. 10, 2013). Many regional school leaders argue against Keystone exams: Philadelphia Inquirer. Retrieved from: [http://articles.philly.com/2013-11-10/news/43858373\\_1\\_keystone-tests-language-arts-test-supporters](http://articles.philly.com/2013-11-10/news/43858373_1_keystone-tests-language-arts-test-supporters).

<sup>3</sup> Warren, J.R., Jenkins, K.N. & Kulick, R.B. (2006). High School Exit Examinations and State-Level Completion and GED Rates, 1975 through 2002. *Educational Evaluation and Policy Analysis*, Vol. 28, No. 2. pp 131-152.

<sup>4</sup> Holme, J.J., Richards, M.P., Jimerson, J.B., & Cohen, R.W. (2010). Assessing the effects of high school exit examinations. *Review of Educational Research*. Vol. 80, No. 4, pp. 476–526. DOI: 10.3102/0034654310383147.

<sup>5</sup> Baker, O. & Lang, K. (June 2013). The effect of high school exit exams on graduation, employment, wages and incarceration. NBER Working Paper Series. Working Paper 19182.

- Statewide, more than a quarter-million re-test administrations of the Keystone exams resulted in another round of below-proficient scores.

While concerning, it is important to note the rollout of this policy is extraordinarily complex. Some challenges in the first years of implementation are to be expected, and may be addressed as the policy matures. Still, the assessment results bear careful watching. If early proficiency rates are accurate predictors of how students will perform on future keystones, Pennsylvania students will face serious obstacles in attaining high school diplomas when the policy goes into full effect in 2017.

The purpose of this Research Brief is to provide education stakeholders with accurate, comprehensive analysis to inform the growing debates about the keystones. To that end, we first provide background on exit exams nationwide, followed by details on Pennsylvania's graduation requirements and the keystones. We then present results of our analyses of publicly-available data on how students, districts, and charters fared on the tests during the 2013-14 school year.

## The Status of State-Level Exit Exam Policies

Nationwide, the movement toward state-level high school graduation exams began in the late 1970s and early 1980s when a number of states adopted exams to ensure that students attained a minimum level of competency before graduation. Over time, states transitioned to exams testing high school-level content.<sup>6</sup>

As of 2012, 25 states had test-based graduation requirements of some kind, with considerable variation in the policies including the number of exams taken, subjects tested, and provisions for students who fail to meet targets.<sup>7</sup> For example, New York requires students to pass five exams to earn a diploma, with performance determining the type of diploma offered: Students who earn a 90 or better receive a Regents with Honors; students who earn between 65 and 89 earn a Regents diploma, and local diplomas are offered for students who fail to reach the marks.<sup>8</sup> New Jersey students must pass one comprehensive, multi-subject exam, and students who fail are assessed in other ways.<sup>9</sup> Maryland students must pass tests in three subjects to graduate; students who repeatedly fail to reach the targets may complete projects administered by local districts.<sup>10</sup>

The effects of graduation requirements are difficult to discern due to the significant variation in policies across states. Early investigations (pre-2005) typically found that graduation exams do not lead to lower high school graduation rates.<sup>11</sup> More recently, a 2011 National Academy of Sciences review of three rigorous studies "suggests that high school exit exam programs, as currently implemented in the United States, decrease the rate of high school graduation without increasing achievement."<sup>12</sup> There are

<sup>6</sup> Holme, J.J., Richards, M.P., Jimerson, J.B., & Cohen, R.W. (2010). Assessing the effects of high school exit examinations. *Review of Educational Research*. Vol. 80, No. 4, pp. 476–526. DOI: 10.3102/0034654310383147.

<sup>7</sup> McIntosh, S. (2012). State high school exit exams: A policy in transition. Center on Education Policy (2012). Retrieved from: <http://www.cep-dc.org/displayDocument.cfm?DocumentID=408>.

<sup>8</sup> New York State Department of Education. (2013). Diploma/credential requirements. Retrieved on June 1, 2015 from: <http://www.p12.nysed.gov/cjai/gradreq/diploma-credential-summary.pdf>.

<sup>9</sup> Center on Education Policy (2012). New Jersey: Profile of state high school exit exam policies. Retrieved from: <http://www.cep-dc.org/displayDocument.cfm?DocumentID=408>.

<sup>10</sup> Maryland State Department of Education. The Maryland testing requirement for graduation & the bridge plan for academic validation. Retrieved on June 1, 2015 from: [http://hsaexam.org/bridge\\_overview\\_09.html](http://hsaexam.org/bridge_overview_09.html).

<sup>11</sup> McDermott, K.A. (2009). The expansion of state policy research. *Handbook of Education Policy Research*. Eds. Sykes, G., Schneider, B., Plank, D.N. & Ford, T.G. New York: Routledge.

<sup>12</sup> National Research Council. (2011). *Incentives and test-based accountability in education*. Committee on Incentives and Test-Based Accountability in Public Education, M. Hout and S.W. Elliott, *Editors*. Board on Testing and Assessment, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

indications that the negative impact is strongest for traditionally-disadvantaged students, and that the impact increases as poverty rates increase.<sup>13</sup>

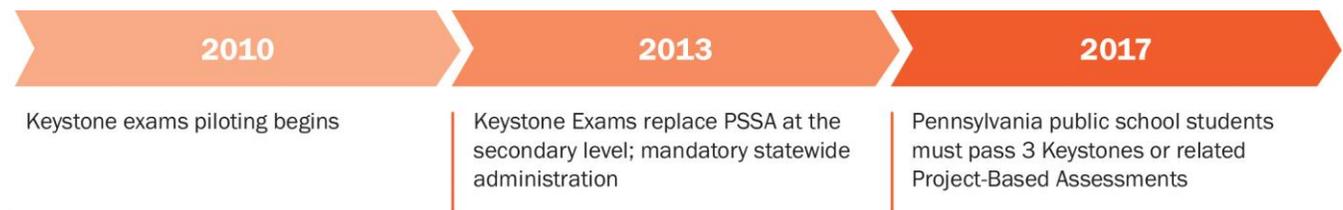
## Pennsylvania’s Graduation Requirements and Keystone Exams

Pennsylvania’s state graduation requirements center on a course-specific, final-exam structure, with students taking Keystones as they complete coursework in the three required subject areas. If a student fails to reach proficiency on the first attempt, he or she can re-take the exam until the requirement is met.<sup>14</sup>

As in many states, Pennsylvania also provides an alternative route for graduation. If a student fails to pass a Keystone on two separate attempts, he or she is eligible to take a Project-Based Assessment, or PBA, developed by the Pennsylvania Department of Education (PDE).<sup>15</sup> In addition, Pennsylvania regulations allow superintendents to waive Keystone and PBA requirements for students on “a case-by-case basis for good cause.” If a superintendent intends to grant waivers for more than 10 percent of the district’s graduating class, approval from the state must be obtained.<sup>16</sup>

There have been several revisions to the graduation requirements policy since its initial adoption in 2010. Field testing for the exams began that year, and the first full, statewide administration occurred in 2011. During the 2012-13 school year, the Keystones replaced the 11<sup>th</sup> grade PSSAs as the secondary level assessment required under *No Child Left Behind*. The regulation goes into full effect in 2017. See Table 1, below.

**Table 1. Timeline of Keystone Exam Implementation**



Source: Pennsylvania Department of Education, *ESEA Flexibility Request from Pennsylvania* – p. 23

## Data Sources and Methodological Note

On January 9, 2015, the *Philadelphia Public School Notebook* posted 2013-14 Keystone exam results for all Pennsylvania public middle and high schools using data provided by the PDE.<sup>17</sup> In this brief, we aggregate these school-level records to the local education agency (LEA) level—typically a school district or charter school—to present results as consistently as possible. Keystones are administered in the winter and spring; the *Notebook*-posted records include results from both test administrations for all three Keystone exams in 2013-14.

<sup>13</sup> Holme, J.J., Richards, M.P., Jimerson, J.B., & Cohen, R.W. (2010).

<sup>14</sup> The Pennsylvania Code. 22 Pa. Code §4.51b. (2014). Retrieved from: <http://www.pacode.com/secure/data/022/chapter4/s4.51b.html>.

<sup>15</sup> The Pennsylvania Code. 22 Pa. Code §4.51c. (2014). Retrieved from: <http://www.pacode.com/secure/data/022/chapter4/s4.51c.html>.

<sup>16</sup> The Pennsylvania Code. 22 Pa. Code §4.51d. (2014). Retrieved from: <http://www.pacode.com/secure/data/022/chapter4/s4.51d.html>.

<sup>17</sup> Limm, D. (Jan. 9, 2015). Pa. releases some recent test results in an accessible format. The Philadelphia Public School Notebook. Retrieved from: <http://thenotebook.org/blog/158088/pennsylvania-keystone-pssa-test-results-in-spreadsheet-format>.

All findings represent the performance of students on individual Keystones. It is important to note that individual students may take multiple Keystones within a school year in one or more subject areas. For this reason, there are instances in which the number of Keystone exams administered in a single LEA exceeds the overall size of the LEA’s high school population.

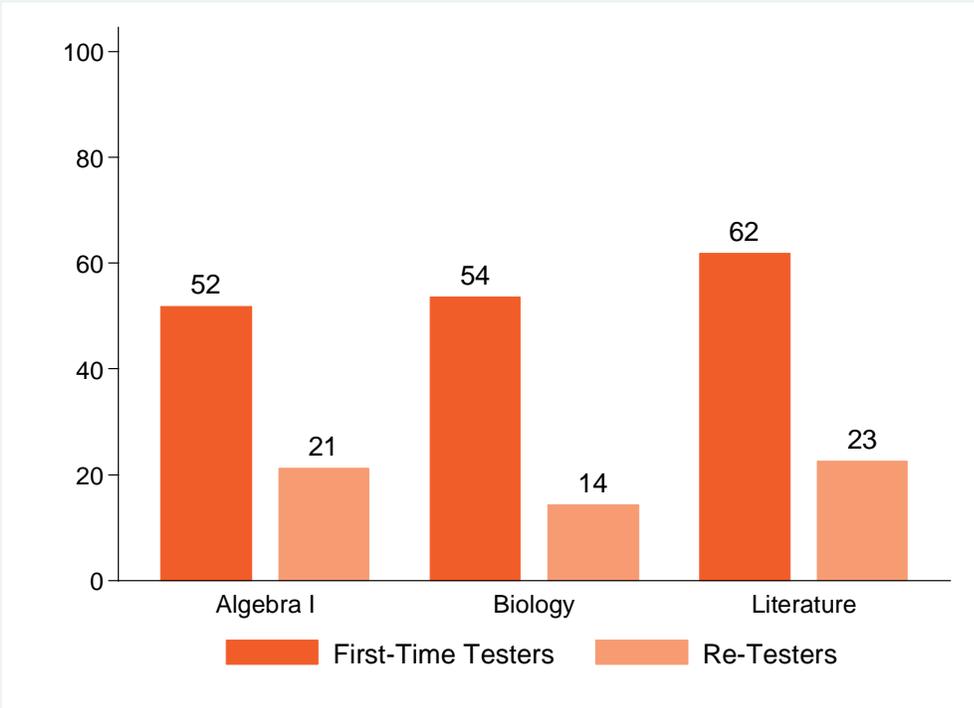
The results of our analyses are divided into three sections:

- 1. **Section I:** Overall performance in 2013-14. Results are presented for both first time test-takers and subsequent administrations to re-testers in all three subject areas.
- 2. **Section II:** Relationships between pass rates and the size of a district or charter’s traditionally-underserved populations. We present data for economically disadvantaged populations and non-white populations, for both first-time test-takers and re-testers.
- 3. **Section III:** Implications for Pennsylvania’s public schools.

### Section I: Keystone Exam Performance, School Year 2013-14

In 2013-14, a total of 705,965 Keystone exams were administered statewide; 43 percent, or 306,091, of these tests were classified as re-tests. Figure 1 presents the percentage of first-time test-takers and re-testers who scored proficient or above on each of the three Keystone exams in 2013-14.<sup>18</sup>

**Figure 1. Keystone Exam First-Time Test-Takers and Re-Testers, Percent of Students Proficient or Above, 2013-14**



*Sample Sizes: **Algebra I:** 139,895 First Time Tests; 155,616 Re-Tests; **Biology:** 132,916 First Time Tests; 89,781 Re-Tests; **Literature:** 127,023 First Time Tests; 60,694 Re-Tests*

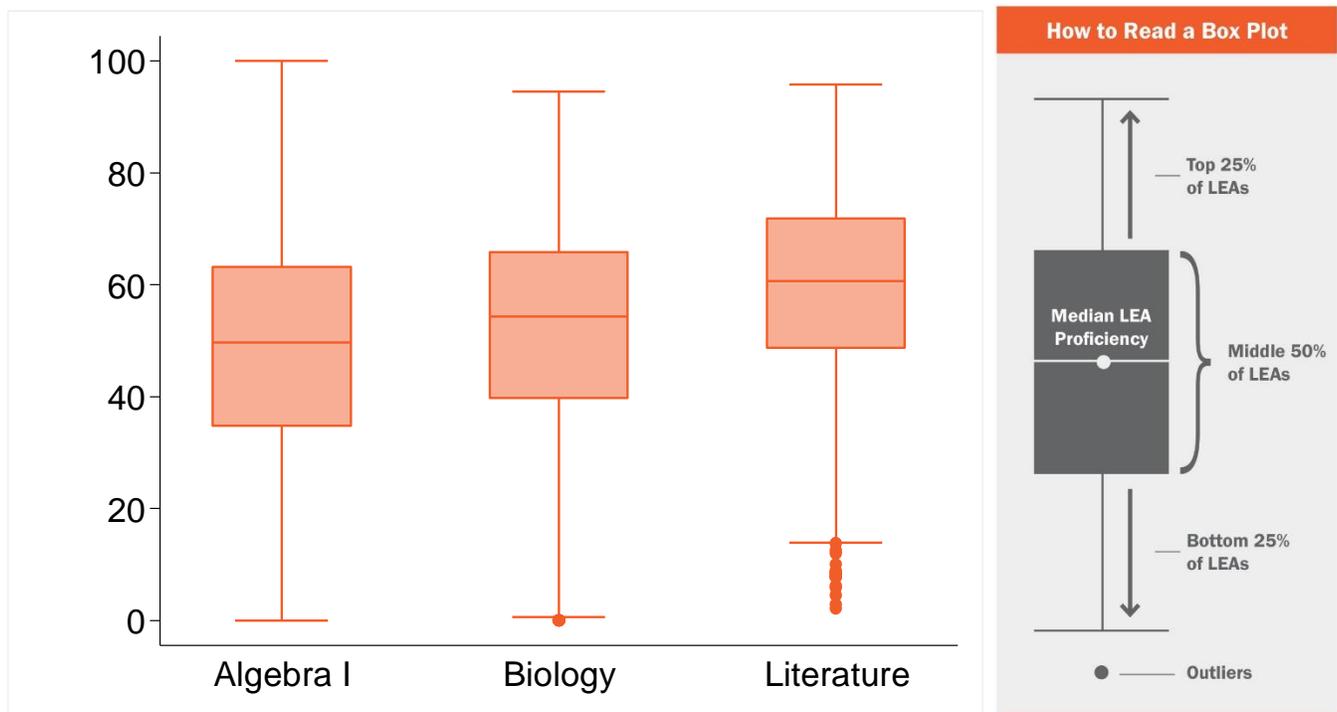
<sup>18</sup> First time test-takers within each subject represent unique students. It is possible that an individual student could take multiple re-tests on the same Keystone exam within the same academic year.

For all subjects, the percentage of first-time test-takers reaching proficiency ranged between 52 and 62 percent. For all subjects, proficiency levels on re-tests were far lower, ranging from 14 percent in Biology to 23 percent in Literature.

### Variation in First-Time Pass Rates by LEA

At the district and charter school level, there was considerable variation in the percentage of first-time test-takers achieving proficiency on each Keystone exam. Figure 2 presents the state-wide distribution of proficiency levels for first-time test-takers at the district and charter level for each Keystone. Each of the orange boxes in Figure 2 represents the proficiency levels (percentage of students scoring proficient or above) for the middle 50 percent of districts and charters; the line separating the boxes represents the median pass rates among the districts and charters. The lines above and below the boxes extend to the highest and lowest proficiency levels for each Keystone.

**Figure 2. Distribution of the Percentage of First-Time Test-Takers Scoring Proficient or Above, by LEA**



*Sample Sizes: Algebra I: 578 LEAs; Biology: 571 LEAs; Literature: 561 LEAs. The whiskers in each box plot extend to either the highest or lowest LEA proficiency rate on each Keystone exam, or to a value that is twice the size of the inter-quartile range (the distance between the 25th percentile and the 75th percentile) above or below the median value—the outlier LEAs had proficiency rates below this level in Biology and Literature.*

- Across all Keystone exams, more than half of all districts and charters had first-time proficiency rates below 60%.
- In roughly three-quarters of districts and charters, at least 25 percent of all first-time test-takers were not proficient on at least one Keystone exam.

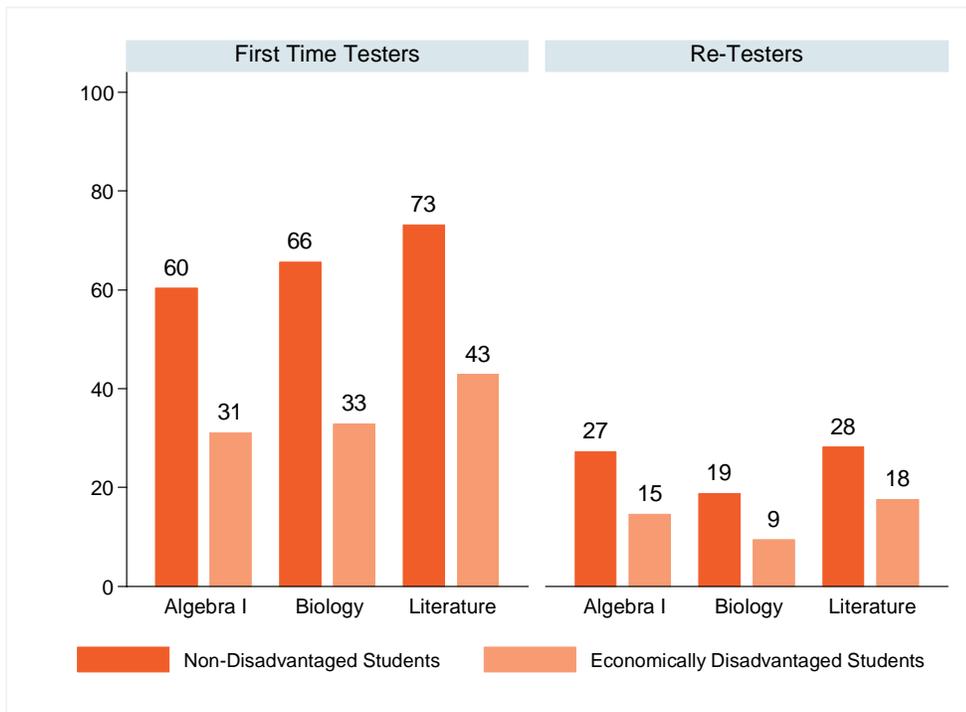
## Section II. Keystone Proficiency Rate Patterns by Income and Race/Ethnicity

Proficiency levels on the keystones were particularly low among two groups of traditionally-disadvantaged students: students living in poverty and non-white students. The results of our analyses for both groups are presented below.

### Keystone Proficiency & Poverty

Figure 3 compares the total percentage of first-time test-takers and re-testers who scored proficient or above on each Keystone exam in 2013-14 for two groups of students: economically disadvantaged students v. their non-economically disadvantaged peers.<sup>19</sup>

**Figure 3. Keystone Exam Test-Takers: Non-Disadvantaged Students v. Economically Disadvantaged Students: % Proficient or Above, 2013-14**



- Across the state, proficiency levels for economically disadvantaged students were substantially lower than those for other students.
- This pattern was consistent across all three Keystone exams, for both first-time test-takers and re-testers.

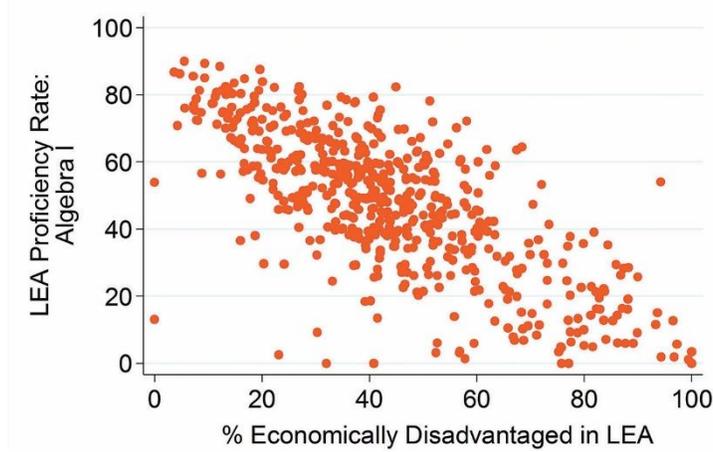
<sup>19</sup> Economically disadvantaged populations are submitted to PDE by individual districts and charters. Poverty data sources may include: Temporary Assistance for Needy Families cases; U.S. Census Bureau; Medicaid, children living in facilities for neglected and delinquent youth; those supported in foster homes; or free/reduced price lunch eligibility.

## Keystone Proficiency Levels by Size of Economically Disadvantaged Population in LEAs

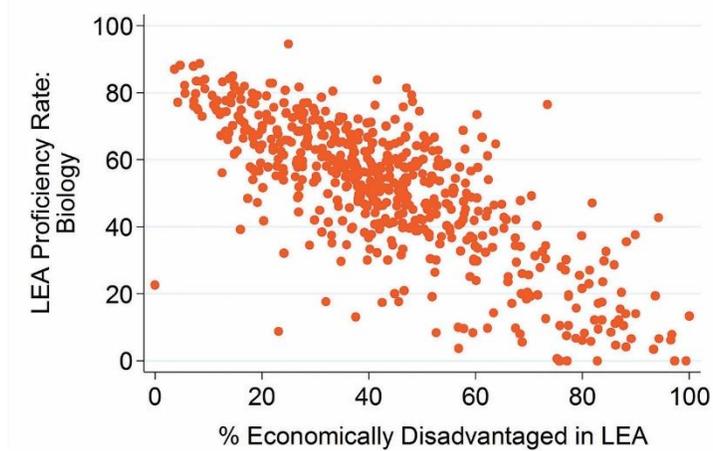
We also observed strong, negative relationships between district or charter proficiency levels on Keystones and the size of their economically disadvantaged populations. That is, those districts and charters with sizable economically disadvantaged populations had lower proficiency levels on Keystones than did districts and charters with fewer economically disadvantaged students. This pattern was consistent for all three exams, as seen in Figures 4, 5, and 6. In these figures, each dot represents a single district or charter.

Research note: The **r-value** reported for each Keystone represents the direction and strength of the relationship between district or charter proficiency levels, and the size of their economically disadvantaged population. r values range from -1 to 1; a value of 0 indicates no relationship; a value of 1 is perfect positive association; and -1 a perfect negative association.

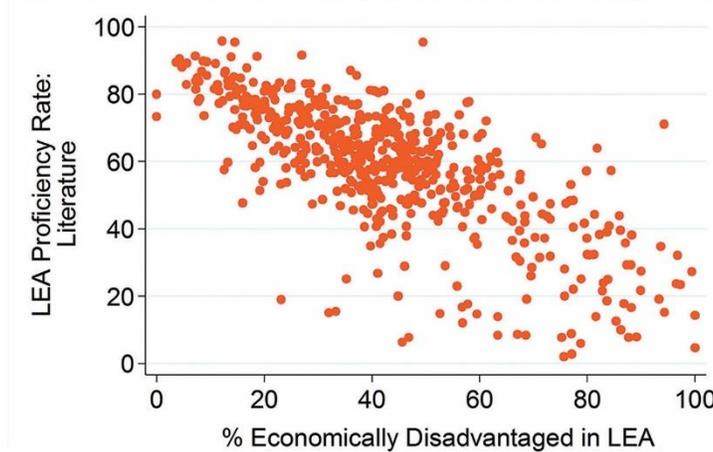
**Figure 4. LEA Proficiency Rates and Economically Disadvantaged Populations: First-Time Algebra I Test-Takers**



**Figure 5. LEA Proficiency Rates and Economically Disadvantaged Populations: First-Time Biology Test-Takers**



**Figure 6. LEA Proficiency Rates and Economically Disadvantaged Populations: First-Time Literature Test-Takers**



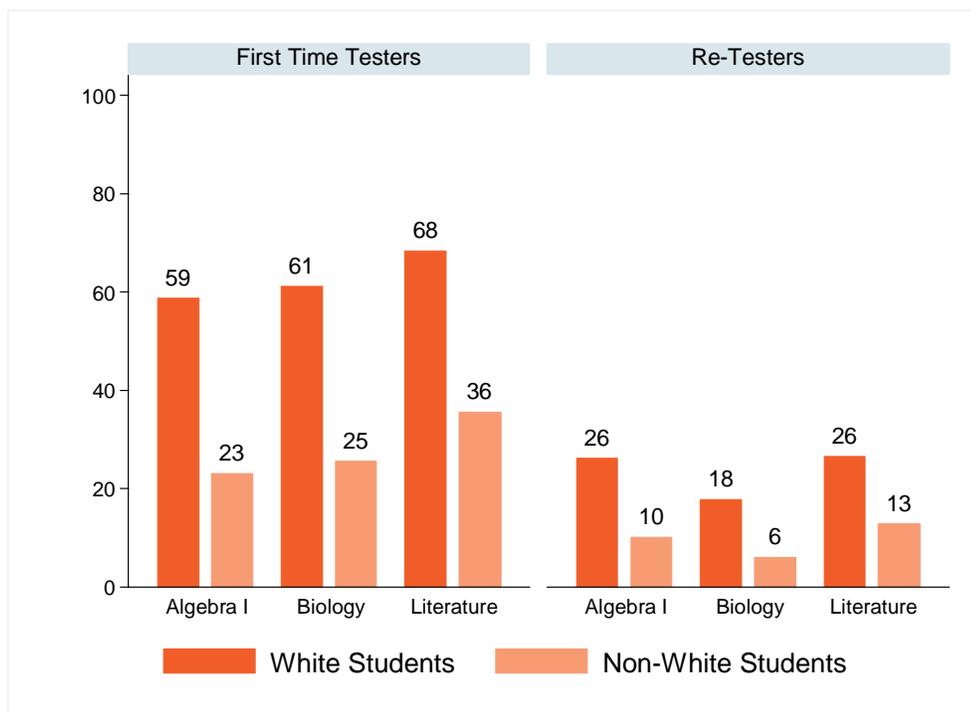
For all three keystones, there were strong, negative associations between LEA proficiency levels and the size of a district or charters' economically disadvantaged population:

- Algebra I:  $r = -.73$ ;
- Biology:  $r = -.77$ ;
- Literature:  $r = -.73$

## Keystone Proficiency and Race/Ethnicity

A similar pattern emerged with respect to race and ethnicity. Across the state, the proficiency levels of non-white students<sup>20</sup> were substantially lower than those of white students. These patterns were consistent across all Keystones, for both first-time test-takers and re-testers. Figure 7 presents the percentages of both white and non-white students who scored proficient or above on each of the Keystones in 2013-14.

**Figure 7. Keystone Exam Test-Takers: White v. Non-White Students Percent Proficient or Above, 2013-14**



Particularly notable are the following patterns:

- For all but first-time test-takers in Literature, the pass rates for white students were at least double that of non-white students.
- Across all three tests, the non-white student first-time pass rates do not rise above 36 percent.

A moderately strong, negative relationship was observed between the size of a district or charter's non-white population and Keystone performance. Figures 8, 9 and 10 present the associations between LEA proficiency levels for first-time test-takers and the size of a district or charter's non-white population across each of the three Keystone exams. Each of the points in Figures 8 through 10 represent a single district or charter.

<sup>20</sup> Non-white students include African American, Latino/Hispanic, Asian, Native American, Pacific Islanders, and all other, non-White, racial and ethnic groups.

Figure 8. Association between LEA Proficiency Levels and LEA Non-White Populations: First-Time Algebra I Test-Takers

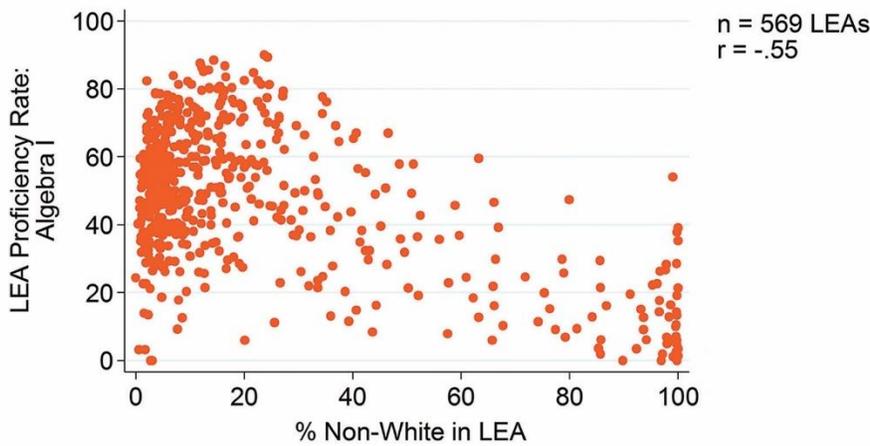


Figure 9. Association between LEA Proficiency Levels and LEA Non-White Populations: First-Time Biology Test-Takers

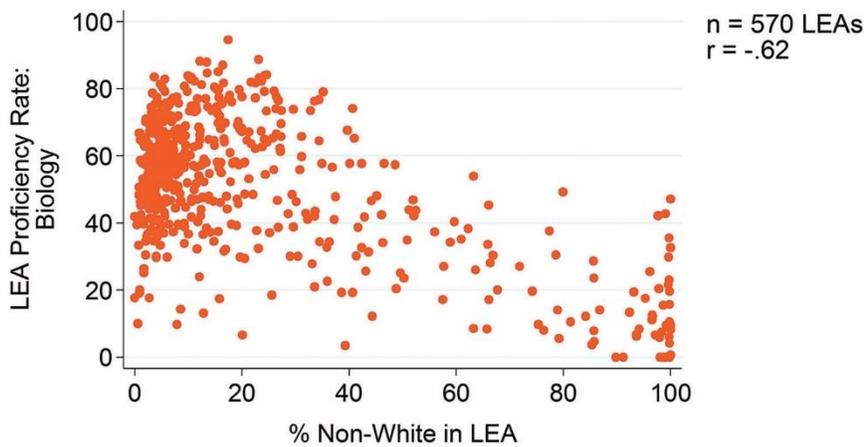
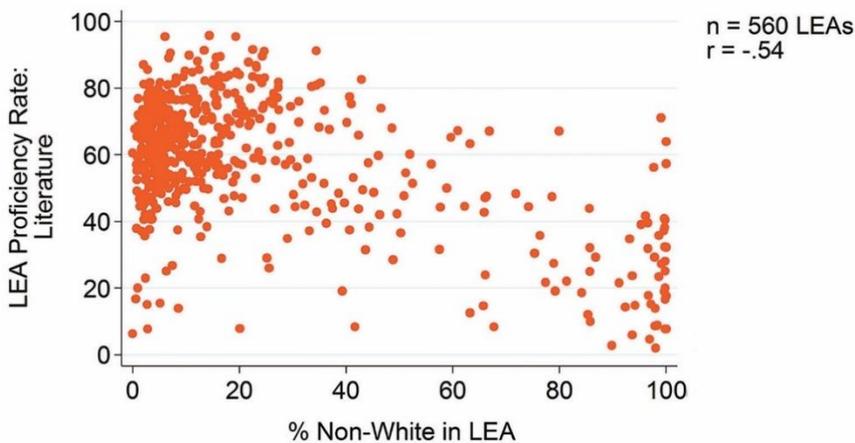


Figure 10. Association between LEA Proficiency Levels and LEA Non-White Populations: First-Time Literature Test-Takers



For all three Keystones, there was a moderately strong, negative correlation between LEA proficiency levels and a district or charters' non-white population:

- Algebra I:  $r = -.55$ ;
- Biology:  $r = -.62$ ;
- Literature:  $r = -.54$

**Notably, the relationships between proficiency levels for first-time test-taker and race/ethnicity were weaker than those between first-time proficiency levels and poverty.**

### Section III. Implications for PA Districts & Schools

The results presented in this brief clearly illustrate that students, districts, and charter schools across the state have faced considerable challenges in passing keystones. In this section, we highlight two significant, related concerns if these patterns persist:

1. If the low proficiency rates seen in 2013-14 continue, keystones will create significant strains on the project-based alternative (PBA) route to a diploma when the policy goes into full effect (graduating class of 2017); and
2. School districts serving the highest numbers of traditionally-disadvantaged students will be faced with disproportionate burdens in administering the PBAs.

#### Project-Based Alternative Routes to a Diploma: The Scale of the Challenge

##### A. Total Number of Project-Based Assessments

State regulation provides that students who fail to pass a Keystone exam after two administrations are eligible to attempt to graduate via PBAs. **In the 2013-14 school year, nearly a quarter-million Keystone tests were failed upon a second try.** The breakdown for each subject is as follows:

- 122,426 in Algebra I
- 76,313 in Biology
- 46,734 in Literature

As noted above, state regulations provide an exception to the graduation requirements under specific conditions: District superintendents can waive the graduation requirements for up to 10 percent of a graduating class for students who have met all expectations except the PBA, and participated satisfactorily in the required supplemental instruction in the specific subject area.<sup>21</sup> Even if superintendents were to take full advantage of this provision, the 2013-14 Keystone results would require students to take, and pass, roughly 220,000 alternative assessments presuming the regulation was in full effect.<sup>22</sup> For perspective, this figure is roughly equivalent to the entire K-12 enrollment of the School District of Philadelphia *plus* the entire K-12 enrollments of the next five largest school districts statewide.

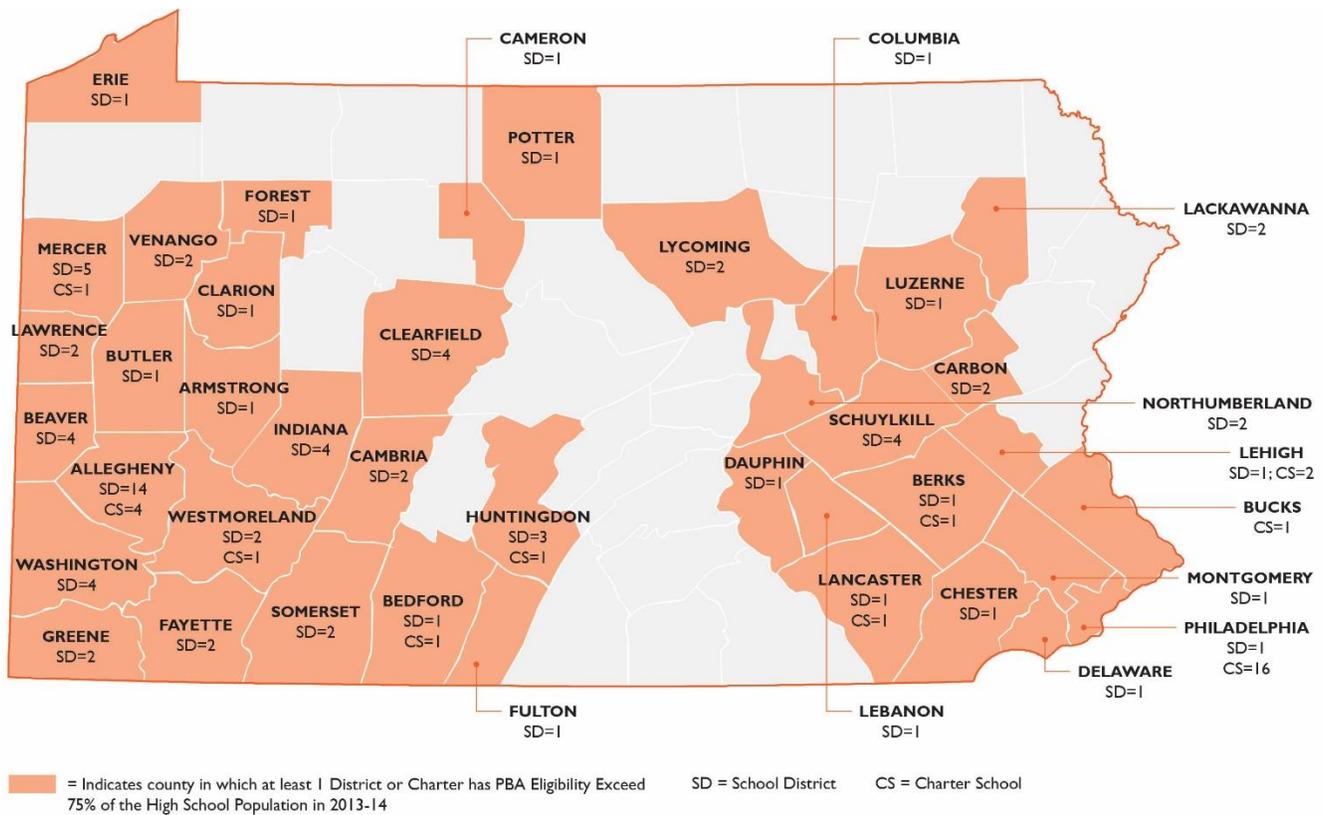
To get a sense of the magnitude of the potential PBA burden in specific local contexts, we calculated the total number of PBAs that would need to be administered by district or charter to allow all students to satisfy state requirements; we then compared these totals to high school enrollments. The results of these analyses are presented in Figure 11 by county.

---

<sup>21</sup> The Pennsylvania Code. 22 Pa. Code §4.51d. (2014).

<sup>22</sup> It is likely the case that a subset of the failed re-tests represented in these results were taken by students graduating before 2017; however, it was not possible to identify the graduating cohorts for test takers in the publicly released data.

**Figure 11. Counties where PBA Eligibility Exceeded 75% of the High School Population in at least one LEA, 2013-14**



We found:

- In more than 100 districts and charters, the number of Keystone re-tests that have been administered with below-proficient results exceeds three-quarters of the total high school enrollment for the district or charter.
- Further, in 10 percent of all districts and charters (n=59), the number of Keystone re-tests that have been administered and failed actually exceeds the entire high school enrollment for the district or charter.
- High rates of Keystone re-take failures are present in districts and charters of every type across the state.

## B. Capacity Challenges for Pennsylvania’s Schools Administering PBAs

The scope of the PBA challenge is compounded by its complexity. Guidance from the Department of Education on the PBAs requires the following of each school administering PBAs: Identification of a lead staff person to administer the assessment, a provision for remediation in related content for students as they complete portions of the PBA, and technical requirements for computer-based testing.<sup>23</sup> These provisions represent new burdens and costs for districts and charters.

<sup>23</sup> The Pennsylvania Code. 22 Pa. Code §4.51b(f). (2014). Retrieved from: <http://www.pacode.com/secure/data/022/chapter4/s4.51b.html>.

## Conclusion

Our analyses of the most recent data available on Keystone exams suggest that Pennsylvania faces significant challenges as the 2017 graduation requirement deadline approaches. These challenges are likely to be felt in nearly every part of the state, with the prospect of especially severe impacts in school districts and charters that enroll large traditionally-underserved student populations. These entities already face systematic barriers from accountability systems that rest on standardized test results; they have been disproportionately impacted by state education cuts; and they are continuing to reduce tutoring, expanded learning, and other vital programs for their students.

The analyses presented in this brief should inform policymakers as they assess the efficacy of the Keystones and consider additional supports to schools or modifications to the graduation policy. While one year of test results does not provide a basis for definitive analysis of the policy, initial results are sobering.