TIME ON TEST
The Fixed Costs of 3-8 Standardized Testing in New York State
Resistance to Race to the Top reforms in New York State has manifest in widespread refusal by parents to allow their children to take state mandated exams associated with RTTT. Among the many objections to these reforms – and one that has become a symbolic and substantive rallying point – is the amount of time that is dedicated to standardized testing. Politicians and lawmakers have actively addressed concerns about testing. In its 2014 session, the New York State legislature enacted a law that capped annual instructional hours that could be dedicated to state-administered required assessments at 1 percent, with an additional 1 percent limit placed on other standardized assessments (Laws of the State of New York, 2014, Chapter 56, subparts E & F; see also Sokol, 2014). In the fall of 2015, Governor Andrew Cuomo announced the creation of the Common Core Task Force, which is charged with examining, among other issues, “how the State and local districts can reduce both the quantity and duration of student tests” (Cuomo, www.governor.ny.gov, 2015, Sept 28). At the national level, the Obama administration called on lawmakers to ensure that “no child would spend more than 2 percent of classroom instruction time taking tests,” (Zernike, 2015). Most recently, the New York State Education Department (NYSED) said it would reduce mandated 2016 3-8 ELA tests by one reading passage and one short essay and each math exam by about four questions, though “the maximum time available for students to complete the tests will remain the same as in past years” (Woodruff, 2015).

These actions demonstrate admirable responsiveness to one of the expressed concerns about New York State’s, and the nation’s, school reform movement (e.g. see nysape.org). However, these limits on testing focus only on the time that is devoted to the test itself. While this is important – clearly, lengthy tests take time away from instruction and may place undue stress on young students – it misses a key point.

Before a student picks up a pencil, and in the time after that pencil is put back down, there is a whole system of procedures that accompanies NYS standardized testing in grades 3-8 – all of which detracts from instructional time. This testing process includes a multitude of administrative tasks, such as setting up the classroom, ensuring certain students get their accommodations, counting and distributing the tests, and reading directions. These are the “fixed costs” of testing. They do not diminish even if the duration of each daily exam is reduced.

The federal government’s 2009 competitive grant program for elementary and secondary education, Race to the Top (RTTT), advanced common standards, statewide data systems, processes for improving low-performing schools, and performance-based evaluations for teachers and administrators. This initiative has led to sustained, intense multi-dimensional conflict over educational policy across the country.
This paper argues simply that the proper measurement of testing time must include both the “fixed costs” of standardized testing and time spent on the actual test itself. Using this common-sense standard, it becomes clear that the time (and resources) dedicated to testing are actually much greater than the tests’ duration. This is important because when students are engaged in this testing process, they are not engaged in learning. And when teachers are engaged in this testing process, they are not engaged in teaching.
New York State has a long history of standardized testing to measure progress toward standards and benchmarks. Regents exams were introduced in 1865 as an entrance exam to high school; beginning in 1878 they were administered at the end of the year and used to measure content knowledge. These tests became the basis of the Regents High School diploma, and the proud claim that, unique among the states, New York State had a consistent standard that guaranteed the quality of the education of its high school graduates.

The time for 3-8 testing in NYS, including the test itself and the fixed costs consume approximately 2 percent of the “required annual instructional hours.” This exceeds and is already double the 1 percent standard that was passed by the legislature.

Overall, within our analytic parameters, we find that the time given to mandated New York State 3-8 testing, including actual testing-taking and the associated fixed costs of this testing, consumes approximately 2 percent of the “minimum required annual instructional hours” (Sokol, 2014). This is more than the 1 percent limit that the NYS legislature placed on “State-administered required assessments” in its 2014 legislative session (Laws of the State of New York, 2014, Chapter 56, subparts E & F). We also demonstrate that the testing process forces a reallocation of resources for all students, regardless of whether or not they are in a testing grade. This is a displacement of resources from their intended and appropriate target in order to accommodate NYS tests.
The federal government entered in 1994. The Improving America’s Schools Act (IASA), which paralleled the Goals 2000: Educate America Act, required that states measure student attainment of state-developed performance standards. In New York, these exams began administration in 1999, in math and ELA for students in grades 4 and 8. When the Elementary and Secondary Schools Act was reauthorized in 2002, with legislation called the No Child Left Behind Act of 2001 (NCLB), accountability measures called for, among other things, annual testing of ELA and math in grades 3 through 8. These annual exams were first administered in NYS in 2006. In accord with New York State’s agreement under the Race to the Top Initiative, exams aligned to the Common Core Learning Standards in ELA and math for students in grades 3 through 8 were first administered in 2013.

The time allotted for state testing has steadily increased over the years, as have the number of grade levels required to take the exams. The PEP tests assessed math and ELA in grades 3 and 6; in 1980, these tests were allotted 95 minutes for students in 3rd grade and 210 minutes for students in 6th grade.

Under NCLB, testing was expanded to include all students in grades 3 through 8; students in these grades sat for between 175 (3rd grade) and 325 (8th grade)
minutes over the course of 4-6 days, depending on grade level. Under Race to the Top, tests run for between 400 (3rd grade) and 520 (8th grade) minutes over 6 days. Judging by the number of grades tested and by the duration of those tests, we see that what began as standardized testing in a few grades to measure student achievement, identify students needing remediation, and assess progress toward NYS-specific benchmarks, has morphed into an expansive accountability process that tests all students in almost all grades.

**SURVEY FINDINGS**

Although the length of a school day in New York may be 6 hours or more, actual required instructional hours are fewer. Instructional hours do not include lunch or recess, although they do include “supervised study activities,” or study hall. In order for districts to qualify for state school aid, New York State requires 5 instructional hours for students in grades K-6 and 5.5 hours for students in grades 7-12.

The legislative promise of 1 percent, enacted in 2014, relies on “minimum required annual instructional hours” (5 hours per day for students in grades K-6 and 5.5 hours per day for students in grades 7-12) in its discussion of capping the time on “state-administered required assessments” (Laws of the State of New York, Chapter 56 of 2014, subpart E). Thus, we use these instructional hours to calculate the proportion of time that is dedicated to NYS 3-8 testing.

**METHODOLOGY**

The data analyzed in this paper were collected via a statewide web survey of teachers. A more detailed description of our methodology can be accessed on our website.

**Sampling**

We began with the 2014-15 School Directory and General Information database. This is a database of all school buildings in the state, publicly available on the NYSED website. We removed all schools that were not public schools and that did not include grades 3-8 and developed a process for randomly selecting one teacher from each building in the sampling frame.

**Data Collection**

Data were collected May 5, 2015 (right after the administration of the tests) through August 1, 2015 using Qualtrics web survey software. The total sample size collected was 143 completed interviews, garnering a simple response rate of 8.7 percent.

**Survey and Measures**

After validating that the teachers who responded proctored either the ELA or math 3-8 exams in the spring of 2015, the survey asked them to estimate the amount of time they spent, on an average test day, on the following tasks: Pre-test related tasks: room preparation; location changes; counting and distributing exams; and the delivery of test directions. Actual test taking: the number of minutes spent on actual testing. Post-test related tasks: collecting, counting, and securing exams; location changes; breaks and reorientation time to transition to non-testing related tasks. Day before prep: number of minutes spent preparing the room for testing conditions on the days preceding the exams. Non-testing activities: actual instruction, outside play, free time in class, classroom celebration, movies, lunch/recess. New content: likelihood (more/less/equally) to provide instruction on new content on testing days as opposed to non-testing days. Two open-ended questions at the end of the survey provided respondents the opportunity to provide additional comments. The full survey instrument can be accessed on our website.
The Testing Process

Teachers were asked to report on how much time was spent before, during, and after testing on testing related activities. Figure 2 presents the amount of time given to actual testing itself and the associated fixed costs, for ELA and math in grades 3-6 and 7-8, as well as an average of the two exams.

Our results demonstrate that on average, students in grades 3-6 sit filling out exams for 89 minutes, while students in grades 7-8 fill out exams for an average of 92 minutes. Taken together, we see that students spend, on average, about 90 minutes (as is expected in the Race to the Top parameters, see earlier discussion) taking tests with the administration of each ELA or math exam.

Figure 2 also shows that the “fixed costs” more than double the time given to testing. For example, exams in grades 7-8 require nearly an hour (55 minutes) of pre-test activities: 20 minutes to prep the room, 14 to change locations for some students, 12 to count and distribute tests, and 9 to deliver instructions. And then after pencils are put down, 17 minutes to collect, count, and secure exams, 5 to change locations, and 20 to take a break and transition students to other activities. This amounts to 97 minutes of “fixed costs” associated with testing, on average, for 7-8 graders. The “fixed costs” amount to 96 minutes for students in grades 3-6. Again, these “fixed costs” of exam administration remain constant with the administration of each exam.

FIGURE 2

Actual Testing and Fixed Costs, ELA & Math with Averages

<table>
<thead>
<tr>
<th></th>
<th>Prep room</th>
<th>Change location</th>
<th>Counting, distributing</th>
<th>Deliver instructions</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-6 Math</td>
<td>14</td>
<td>11</td>
<td>12</td>
<td>87</td>
</tr>
<tr>
<td>3-6 ELA</td>
<td>15</td>
<td>12</td>
<td>13</td>
<td>91</td>
</tr>
<tr>
<td>3-6 Average</td>
<td>15</td>
<td>11</td>
<td>12</td>
<td>89</td>
</tr>
<tr>
<td>7-8 Math</td>
<td>19</td>
<td>14</td>
<td>13</td>
<td>8</td>
</tr>
<tr>
<td>7-8 ELA</td>
<td>21</td>
<td>14</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>7-8 Average</td>
<td>20</td>
<td>14</td>
<td>12</td>
<td>9</td>
</tr>
</tbody>
</table>

Total=180  Total=190  Total=185  Total=183  Total=194  Total=189
Combining the time of the test and the associated “fixed costs,” we see that on average, 185 minutes are dedicated to testing in grades 3-6 and 189 minutes in grades 7-8 during the administration of each daily exam. The entire testing process amounts to about three hours. This is more than half of the required 5 or 5.5 instructional hours in a school day (Figure 3). And, since these tests are administered on six separate days – to accommodate three ELA exams and three math exams – we can apply these counts to each day of testing; more than half of a school day on six separate days. Therefore the data show that students and teachers lose nearly four full days of instruction to the NYS state testing process in grades 3-6 and nearly 3.5 in grades 7-8.

The Rest of the Day

We also wanted to gain an understanding of the degree to which actual instruction occurs on testing days. Thus, we asked teachers to detail the activities that they and their students engage in, and how many minutes are dedicated to each, during the remainder of their time in school on a testing day (Figures 4 and 5).

Non-test time is distributed across multiple kinds of activities. In grades 3-6, 48 minutes are utilized for play, free time and class celebrations – giving young students time to unwind after a long test. This leaves approximately 98 minutes for actual instruction. In grades 7-8, less time
is consumed by non-test, non-instructional activities, and on average students in these grades receive approximately 118 minutes of actual instruction on testing days. This translates into just under ten hours of instruction for grades 3-6 over the course of six days of testing; as opposed to 30 hours of instruction that would occur over six non-testing school days. Students in grades 7-8 receive approximately 12 hours of instruction, as opposed to 33 hours, over course of the six testing days.

Finally, more than three-quarters of the teachers we surveyed stated that they are less likely to introduce new academic material to students on testing days. As one teacher stated, “state exams over the course of three days derail real instruction for the entire day a test is administered and tend to derail instruction for the whole week of a 3-day test administration period. This means that instruction is severely curtailed and stunted for the two weeks in the spring when exams are administered.” Clearly, according to our results, instruction is, indeed, sharply curtailed on testing days.

**Teacher Time**

The testing process also consumed teacher time prior to test administration. Teachers reported that they spent, on average, 23 minutes creating testing conditions in their classroom before the first day of testing. These activities include covering posters and books that students might rely on during a test. According to one teacher, “I did have to cover anything that could be construed as ELA material. So, all the posters with the inspiring vocabulary words had to be covered. And as most of those were up high and

![Average Minutes of Instructional Time on Typical Exam Days Consumed by Testing Process, Actual Instructional, and Other Activities](chart)

**FIGURE 4**

**APPROXIMATELY ONE THIRD OF THE DAY IS LEFT FOR ACTUAL INSTRUCTION ON STATE TEST DAYS (30 PERCENT IN GRADES 3-6, 35 PERCENT IN GRADES 7-8).**
Teachers’ Likelihood to Present New Content on State Exam Days

77 PERCENT OF 7-8 TEACHERS AND 84 PERCENT OF 3-6 TEACHERS SAY THAT THEY ARE LESS LIKELY TO PRESENT NEW ACADEMIC MATERIAL ON TEST DAYS, COMPARED WITH NON-TEST DAYS.

“…state exams over the course of three days derail real instruction for the entire day a test is administered and tend to derail instruction for the whole week of a 3-day test administration period. This means that instruction is severely curtailed and stunted for the two weeks in the spring when exams are administered.”
Displacement of Time and Resources and Disruption of Instruction

Another important issue that surfaced from our research is the “displacement” of time and resources. These were instances in which teacher time or materials were diverted from their intended target to support testing. This was not an uncommon occurrence. For example, one teacher told of the use of the gym as a testing location for students with disabilities who require special testing accommodations; students in this school were not able to use the gym during testing time on the six days of testing. Another teacher tested her students with special needs in the library. Students in this school were unable to access the library during testing time; if a class was scheduled for library, the librarian would visit the classroom with a cart of books. In addition, this teacher was required to cover library materials so that her students could not use them to aid in their answers. These coverings – which, fortunately, did not extend to the bookshelves themselves – remained in place for the two weeks of testing; it would have taken too long to take them down, and put them back up, in between tests. Students who had library time outside of testing times had access to a covered, though still book-accessible, library.

Teachers reported that testing permeated the entire school during testing weeks. In many instances, rooms remained “prepped” for the duration of testing. This meant that students sat in sterilized classrooms, under testing conditions for two weeks, even when they were not being tested. According to one teacher in our sample, “prepping the classroom and covering up walls makes the class devoid of interest.” Another teacher told of subject areas, often specials like gym, music, and art, which had to adjust their curriculum in order to keep noise at a minimum, and students who had to be rerouted through the school building – they could not risk disrupting the testing classrooms as they walked by them on their way to specials. “When there is a test going on,” one teacher reported, “everything changes.”

And even after the exam period was over, displacement of time and resources endured, as teachers were required to proctor make up exams or score the tests. Teachers recounted, “I am pulled for make-up testing and [while I] only had to do an additional 90 minutes, other teachers did three or more makeup sessions and had to do many more hours of testing,” and “Teachers are out of the classroom after testing to get training to score tests and again to actually score tests. In fourth grade alone, there are eight instructional days lost to NYS test administration.”

The displacement of resources was also felt by students in non-testing grades; the reading teacher who was required to proctor and so could not provide services during testing time, or the speech teacher whose room was used as a testing site, and so had to find alternate spaces to provide services. One teacher spoke of the drain of make-up testing, when students missed an exam because of illness: “When I am assigned to do make-up testing, my services are cancelled in the ICT and resource room settings, as no substitute is brought in to cover for me. So students do not receive their mandated IEP services on the days I do make up testing.”

In the most extreme case, we learned of students in non-testing grades who were required to be out of their classroom during the administration of the tests because their classroom was being used as a testing site for students with special needs. Their creative teachers found ways to engage them during this time; a trip to a local bookstore, a trip to the park. On one day, the class spent the time learning in the cafeteria. Nevertheless, this amounts to considerable but not easily quantified hours of displacement and interrupted instruction. These examples show clearly how the testing process consumes the school community. Instructional time, and in some cases required services, are sacrificed in the mix.
DISCUSSION AND CONCLUSION

Testing takes more time than the duration of each exam. There is a whole process – the “fixed costs” of testing – that must be considered when accounting for time dedicated to testing. All of this time amounts to lost instruction for students in testing grades. Our research suggests that it also amounts to the displacement of resources – in some instances lost instructional time or lost services – that can affect students in all grades, irrespective of their own participation in exams.

The NYS legislature has asserted that time for testing “cannot exceed, in the aggregate, 1 percent of the minimum required annual instructional hours” for each grade (Laws of the State of New York, Chapter 56 of 2014, subpart E). As stated, in NYS, grades 3-6 are required to have at least 5 hours of instruction, and 7-8 are expected to have at least 5.5 hours. At the mandated 180 days of school per year this amounts to 900 hours for grades 3-6 and 990 hours for grades 7-8; 54,000 minutes and 59,400 minutes, respectively. Within this context of required annual instructional hours, our results reveal that:

- in the 2014-15 school year, on average, 1,110 minutes were dedicated to the NYS standardized testing process for students in grades 3-6 and 1,134 minutes were dedicated to this process for students in grades 7-8, and
- this represents 2 percent of “required annual instructional hours” for grades 3-6 and 1.9 percent for grades 7-8, and exceeds, and almost doubles, the standard set by the legislature.

We know, however, that this percentage of time is an underestimation. First, we do not include science testing, which is mandated for students in 4th and 8th grade. In addition, we use the 180 day school year as a basis for establishing instructional hours in a year, even though previous research demonstrates that students in New York State do not receive the “180 required days of instruction” (NYSED, 2014) as typically assumed (Jacobowitz, 2015). Finally, we did not measure all tasks associated with 3-8 NYS standardized testing, such as practice tests or test prep. Thus, the 2% that our exploratory study evidences is a floor for the amount of time devoted to 3-8 NYS testing.

Elected leaders and appointed policy makers have been talking about shortening the exams (Harris, 2015, Sept 16; NYSSBA, 2015). We agree it is a good idea to reduce the time given to these tests. But as our research demonstrates, reducing the duration of each test will have minimal effect; less time will be spent on each exam, but the fixed costs remain and will still detract from instruction on those days. Moreover, apportioning the tests over six days means that these “fixed costs” – and the time given to them – are replicated with each administration. The only way to eliminate these fixed costs is to reduce the number of exam days.
A final contextualizing comment: the “1 or 2 percent for tests” paradigm is arbitrary. It is grounded less in science and more in rhetoric. We believe that where time on testing is concerned, our students would be better served by thoughtful, deliberative testing policies that account for time as it relates to the loss of instruction as well as the capacity of our children to sit for an exam. And while we agree a reasonable amount of testing might be necessary to assess our efficacy in educating our children, the cost to instruction, and student emotional health, must be considered.

We, in New York State, give a lot more time to testing than we think we do. We need to be concerned about the amount of instructional time that is lost, and resources diverted, in this process.

ENDNOTES


2. This is an overview of the history of assessment in NYS; it is not exhaustive. The NYS testing program was subject to continuous change over the years. For example, preliminary competency tests in reading and writing were given in grades 8 & 9 for students who scored below a certain level on the PEP. Additional exams were added to the PEP program in 1983 (5th grade writing). Program Evaluation Tests, to assess the content of NYS curriculum, were administered in social studies in 6th grade (1987), and 8th grade (1989) and in science in 4th grade (1989). It is unclear how many years these tests were administered. And today, students in grades 4 and 8 are tested in science. Information came from the following sources. www.p12.nysed.gov/assessment/timeline-historev.pdf, www.nysl.nysed.gov/edocs/education/edhist.htm#exam, and www.archives.nysed.gov/common/archives/files/ed_background_overview_essay.pdf. We also consulted numerous school Administrator’s Manuals and Teachers Directions distributed through the NYSED from 1966 through 1999. Our reference list cites each of these documents. See also Grant, 2000; Layton, 1986; New York State Education Department, 1984; University of the State of New York, 1991.

3. By 1987, PEP tests were administered in May.

4. Tests from 2006-2009 were administered in January (ELA) and March (math). Spring administration of exams began in 2010.

5. University of the State of New York, 1980. Tests, and the time given to them, likely changed over the many years of PEP implementation. These data are for 1980, though they are likely applicable to other years as well.


8. The minimum length for school day for ½ day kindergarten is 2.5 hours. See https://stateaid.nysed.gov/attendance/attendance memo.htm

9. There was a lack of school websites with teacher emails in New York City. Overall, 56 percent of schools outside of NYC had accessible email addresses for teachers; in New York City only 18 percent provided teacher emails. Thus, our sample was skewed geographically to districts outside of NYC. We weighted the sample to adjust for this as well as for grade organization.

10. A few other questions related to opt-outs were included, and will be reported on in a separate report.

11. While lunch and recess are part of the total school day, Figures 3 and 4 do not include time for lunch and recess as these activities are not considered to be “instructional hours,” as defined by NYS.

12. These totals compute to more than the required instructional hours because here, we are measuring the school day, which as we state previously is often longer than the required minimum instructional hours (usually 6 hours or more), as reported by teachers in our sample, even though we exclude lunch and recess from the analysis.
AUTHOR BIOS

Robin Jacobowitz, Ph.D., is the director of education projects at The Benjamin Center at SUNY New Paltz. Previously, Robin worked with Janice Hirota and Associates on an evaluation of school reform initiatives in New Orleans, Washington DC, New York City, and Dallas. She also worked at New York University’s Institute for Education and Social Policy, where her research centered on the growth and development of charter schools in New York State, the organizational structures that facilitate teaching and learning in New York City small high schools, and leadership transitions in new schools in New York City. She worked with the University of Chicago’s Chapin Hall Center for Children, where her research focused on the relationship between constituency building and policy work in effecting systemic school reform in New York State. Prior to beginning her career in research, Robin worked with the Public Education Network in Washington DC, where she provided technical assistance to local education funds around the country. Robin holds a MEd in education policy from the Harvard University Graduate School of Education, and a Ph.D. from the Robert F. Wagner Graduate School of Public Service at New York University. She is currently a trustee on the Kingston City School District Board of Education and serves on the executive committee of the Ulster County School Boards Association.

KT Tobin, M.S., is the associate director of The Benjamin Center at SUNY New Paltz where she is responsible for designing, conducting, managing, and producing studies on regional issues and concerns. Her recent publications include: Mid-Hudson Municipal Use of Technology & Participation; Agriculture Supporting Communities in the Mid-Hudson Region; Communicating Student Research at SUNY New Paltz to State and Local Elected Officials; Poughkeepsie Plenty: A Community Food Assessment; and the Mid-Hudson Arts and Culture: the Economic Impact. KT is also editor of the CRREO Discussion Brief Series. KT is the former Vice President of the New Paltz School Board and a former member of their Legislative Action committee. She has served on several regional and community committees including the Mid-Hudson Regional Sustainability Planning Consortium, the Mid-Hudson Regional Economic Development Council Metrics Committee, and in New Paltz: the Government Efficiency & Effectiveness Project, GreenWorks, Flood Aid, the Regatta, and the Wallkill River Watershed Alliance. KT is also an adjunct lecturer in Sociology, teaching Intro to Sociology, Social Inequality, and Research Methods.

ACKNOWLEDGEMENTS

Many people contributed to the development of this report. Benjamin Center Cetrino Scholar Emily Vanderpool (Economics ’16) and Micaela Kayser (Women’s Gender, and Sexuality Studies ’16) assembled our email sample. Several teachers gave their time and suggestions to refining the survey used in this research. We are grateful to the many teachers who took the time to complete our survey. Their thoughtfulness and insights drove the content of this report. Thanks are due to the librarians at SUNY New Paltz. Gerald Benjamin and Janis Benincasa provided much appreciated substantive and editorial guidance.

For a complete list of works cited, detailed tables, and full methodology for this paper please visit www.newpaltz.edu/benjamincenter

Citation

Questions or Comments?
Contact Robin Jacobowitz at jacobowr@newpaltz.edu
THE BENJAMIN CENTER for Public Policy Initiatives

Independently and in collaboration with local governments, businesses, and not-for-profits in the Hudson Valley, The Benjamin Center (formerly CRREO):

• conducts studies on topics of regional and statewide importance;
• brings visibility and focus to these matters;
• fosters communities working together to better serve our citizenry;
• and advances the public interest in our region.

The Benjamin Center connects our region with the expertise of SUNY New Paltz faculty. We assist in all aspects of applied research, evaluation, and policy analysis. We provide agencies and businesses with the opportunity to obtain competitive grants, achieve efficiencies, and identify implementable areas for success.

www.newpaltz.edu/benjamincenter