

Right-sizing Testing:

Useful information, efficient testing time

Brian Gong, Center for Assessment

Presentation at RILS, Boston, September 17, 2015

“Too Much Testing”?

Opt out

Publicity

Rationales

* Tests too long * Testing takes away from instruction *
Parents have right to refuse to participate in anything they wish
Harms kids * Doesn't benefit my child * High-stakes testing
destroying education * Teacher evaluation distorts good
instruction * Corporate reform profiteering from too much
testing * Teachers tests provide all the information needed *

Outcomes (by Legislatures,
Districts, etc.)

- Review testing
- Limit testing
- Redesign/Re-procure

Inventory Results

What Counts as Testing Time?

1. Time student is given to answer test
(e.g., “two 40-minute sessions”)
2. Time student is involved to complete
test (e.g., including registration, practice, breaks,
surveys, etc., e.g., “Allow 120 minutes from start
to finish”)
3. Time school/district takes to
administer test to all students, e.g.,
testing window with make-ups, getting students
onto computers
4. Time school/district personnel spend
on test administration (e.g., setting up CBT
software, registering IEPs, CBT profiles)
5. Time teachers/students spend on
preparing instructionally for test

*

Figure 1. Annual Hours of State and Federal Testing

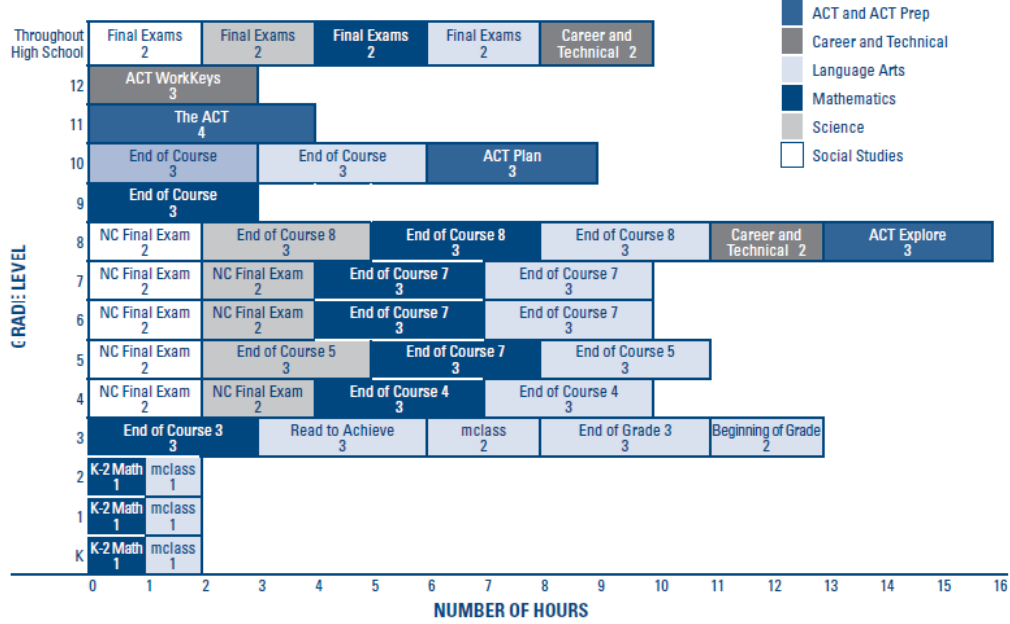


Figure 2. Annual State and Federal Hours of Assessment for an 8th Grader

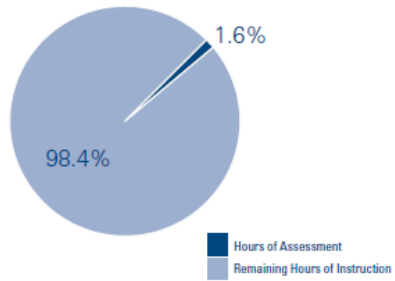
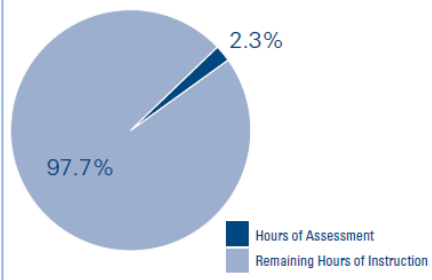
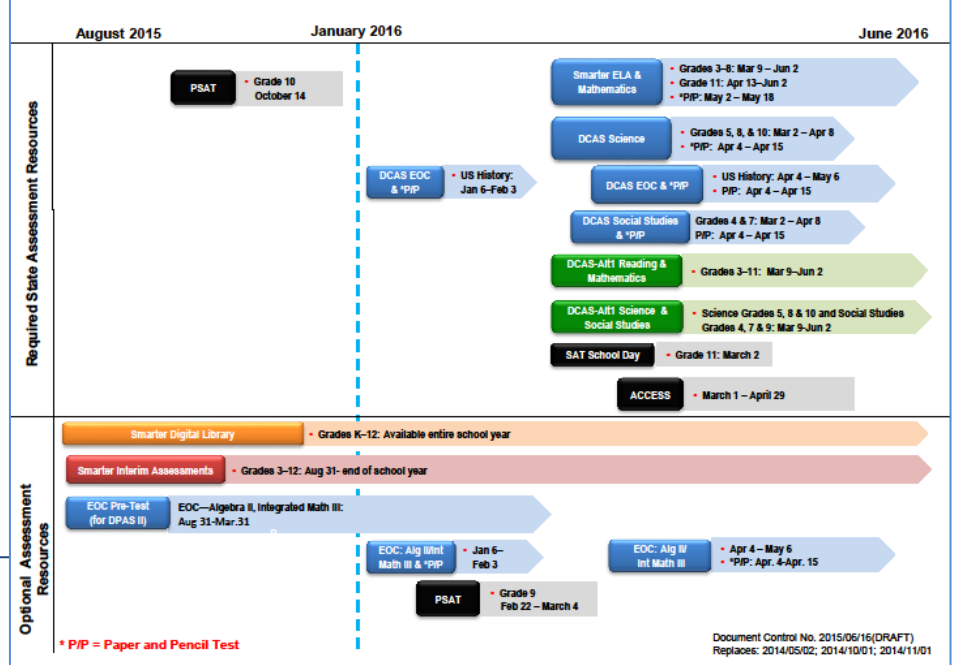


Figure 8. Annual State, Federal, and Local Hours of Assessment for a Median 8th Grader



Delaware System of Student Assessments (DeSSA): 2015-2016 Calendar DRAFT



Limit Testing
See section on reducing testing time

Sample Tools & Reports

- Achieve Student Assessment Inventory for States/School Districts
- *Assessment Inventory Facilitation Process*, Illinois State Board of Ed.
- *Assessment Inventory Results, 2015*, Idaho Department of Education
- *Testing Report and Recommendations, 2015*, Ohio Department of Ed.
- *Districts receive grants to help reduce testing, increase instruction, 2015*, Connecticut State Department of Education
- *How Much Testing is Taking Place in North Carolina Schools at Grades K-12? An analysis of federal, state, and local required assessments. 2014.* North Carolina Department of Public Instruction



Right-size = Costs : Benefits

Right-size = Costs [including testing time] are acceptable for
the Benefits

Can reduce the Costs and/or

Can increase the Benefits

Reduce Testing Time

| Eliminate Whole Tests | Shorten Tests | Shorten Test Administration Time |
|--|---|--|
| <ul style="list-style-type: none"> • Redundant – do not replace • Low value – do not replace • Enhance to make multi-purpose/dual use <p><i>Salient Examples</i></p> <ul style="list-style-type: none"> • Eliminate state’s K-2 tests • Eliminate writing, social studies from certain grades • Eliminate district’s redundant or low value tests (cut tests by 50%) • Adopt ACT for state’s high school test • Use local assessments for federal purposes | <ul style="list-style-type: none"> • Move practice outside of tests; reduce instructions • Reduce number of content standards assessed • Reduce depth of content standards assessed • Reduce reliability of test scores with fewer items • CAT (variable length) <p><i>Salient Examples</i></p> <ul style="list-style-type: none"> • Provide practice tests/tools before testing • Eliminate essays • Design tests with fewer items | <ul style="list-style-type: none"> • Reduce start-up, elapsed time • Reduce permissible “extra” time for individual students • Reduce testing windows • Move towards 1:1 computing • Eliminate paper tests <p><i>Salient Examples</i></p> <ul style="list-style-type: none"> • Implement policy on maximum “additional testing time” • Implement policy with shorter testing windows, e.g., designated day(s) ala SAT • Promote group computer-based testing to reduce district overhead handling time |

Reduce Testing Pain

| Get below time pain threshold | Address test-use issues | Address non-test issues |
|---|---|--|
| <ul style="list-style-type: none"> • Do not have a single test much longer than alternatives • Do not have total testing contact time much longer than alternatives • Reduce test administration time that requires resources from schools/districts and/or disrupts instructional time <p><i>Salient Examples</i></p> <ul style="list-style-type: none"> • <i>Reduce test battery time from 7+ hours to around 4-6 hours maximum</i> • <i>Provide accommodations differently to free up staff</i> | <ul style="list-style-type: none"> • Make sure tests are not being used for purposes stakeholders object strongly to (different than do not support strongly) <p><i>Salient Examples</i></p> <ul style="list-style-type: none"> • <i>Address objections to teacher evaluation, for which testing is seen as an enabling tool/policy</i> | <ul style="list-style-type: none"> • Communicate and address issues where pain is with testing as a symptom or lightning rod, not the real cause of pain <p><i>Salient Examples</i></p> <ul style="list-style-type: none"> • <i>Clear up with objection is to the time spent on testing, or to their association with the Common Core State Standards</i> • <i>Clear up whether objection is to content of the CCSS or to “federal intrusion”</i> |

Improve Benefits

| Better Information | Better Use |
|---|---|
| <ul style="list-style-type: none">• Provide information clearly relevant to stakeholders and users<ul style="list-style-type: none">• Students• Parents• Teachers• Administrators• Policy makers• Public• Provide information when it is most useful/timely• Provide information in ways easy to understand and use <p><i>Salient Examples</i></p> <ul style="list-style-type: none">• <i>Content-based test interpretations</i>• <i>Action-relevant summaries and interpretations</i>• <i>User-flexible reporting</i> | <ul style="list-style-type: none">• Make sure intended uses are clear, e.g., claims• Make sure intended uses are supported by relevant stakeholders and users• Draw information together to provide context, greater credibility, applicability• Communicate intended interpretations and limitations on interpretations and uses in practical ways• Support assessment literacy in key user groups• Support feedback to improve systems <p><i>Salient Examples</i></p> <ul style="list-style-type: none">• <i>Link score reporting to supports for use, e.g., instructional supports, evaluation supports</i> |

Evaluate for Improved “Right-sizing”

- Be clear about values and purposes
- Encourage simultaneous consideration of Costs and Benefits
- Balance evaluations using technical criteria with educational criteria with political criteria with financial/resource criteria, etc.
- Allow sufficient resources to get information to inform deliberations (e.g., learn from others); support good deliberations
- Plan in appropriate feedback and revisiting/revision