The Perfect Storm of Education Reform: High-Stakes Testing and Teacher Evaluation

Sheryl J. Croft, Mari Ann Roberts & Vera L. Stenhouse

No Child Left Behind (NCLB), Race to the Top (RT3), and now Common Core embody over a decade of federal and state education reform purportedly designed to address inequities for global majority and low-income students. However, these policies have in fact expanded inequities and exacerbated a discourse of failure regarding teachers, public schools, and teacher preparation programs. Consequently, public confidence in teachers, teacher preparation programs, and student performance is at an all-time low.

We contend that current reform initiatives (i.e., high-stakes testing and teacher evaluation from K-12 through higher education) are not, in fact, discrete singular efforts. Instead, they represent a confluence of systematic and orchestrated education reform efforts that are akin to storm fronts. These fronts comprise a perfect storm that is eroding the bedrock of public education in the United States through neoliberal policies. Neoliberal principles prescribe that market forces should determine the success or failure of any entity or organization; they support a reduction in public services; and they promote choice, competition, and accountability.

Using the state of Georgia as a case study, we present three interconnected fronts: political climate change, the testing industrial complex, and the resulting mesoscale evaluation system. We propose these fronts as a means to illuminate the gulf between the stated policy intentions of corporate reformers and the actual educational outcomes for public education and teacher education.

Following our analysis of the interconnected fronts, we challenge the assertion that the alignment of the reforms will lead to the claimed outcome—that is, an increase of academic achievement/success and global competitiveness for students, teachers, and the United States as a whole. Instead, we assert that the orchestrated...
alignment is actually being experienced as an assault on the intended beneficiaries. We conclude with responses by students, teachers, and professors to the elements of the perfect storm of education reform and our recommendations for K-12 and higher education practitioners to not just stem but turn the tides.

**Political Climate Change: Setting the Historical Context**

A perfect storm develops within the context of climate change. We posit that political climate change emerges as a series of orchestrated political and legislative efforts intended to drive policy and practice on the national, state, or local levels. The ongoing struggles between those who support equity in education and those who would lay the groundwork to destroy it have led to the juxtaposition of two political climates in Georgia, epitomized by the terms of Governors Barnes and Perdue, respectively. Georgia’s political climate has been gradually changing for over a decade from one of confidence and investment in public education to one of skepticism and funding deprivation. The implications of the political climate change illustrated within Georgia are being experienced nationwide (e.g., in California, Hawaii, Louisiana, New York, North Carolina, Tennessee, and Washington). During the terms of Governor Roy Barnes, from 1998 to 2002, and Governor Sonny Perdue, from 2003 to 2010, legislation in Georgia has exemplified two distinct types of educational reform: one inspired by equity and the other by inequity. Barnes’s A+ Reform package, influenced by the gubernatorial P-16 summit, pursued equity (Croft 2013, 83; Rochford 2007, 18), whereas the legislation passed by his successor, Perdue, seemed mostly influenced by private interests and corporations (Croft 2013).

**2000–2002: Ensuring Equity for All Students**

Through the introduction of his A+ Education Reform Act, during the second half of his tenure (2000–2002) Governor Barnes initiated a system of educational policies designed to insure accountability and to provide resources that supported systemic equity in public schools. Prior to and unlike the mandated and underfunded NCLB, Barnes’s measures coupled accountability with equity by legislating financial as well as structural supports to enhance student learning. For example, after decreasing class size, Barnes appropriated increased capital outlay to support the anticipated need for additional classrooms.

Structurally, his policies provided early intervention programs for all K-3 to K-12, extended the school day for middle school students to enhance instruction, and encouraged and financed dropout prevention initiatives for high school students. In terms of our climate change metaphor, Barnes’s tenure represented relatively calm seas for enhancing student achievement.

**2003–2010: Laying the Foundation for Educational Inequity**

Conversely, the two-term tenure of Sonny Perdue from 2003 to 2010 stunted the momentum achieved by his predecessor. His term represented a change from Barnes’s
political climate to an orchestration of inequalities, particularly for students of the
global majority (i.e., African American and low-income students). Augmented
by NCLB and corporations, Perdue’s tenure produced the initial conditions that
lead directly to a perfect education storm of financial deprivation and inequitable
educational access.

Examples of legislation detrimental to public education during Perdue’s tenure
are the establishment of charter schools, the allowance for charter school flexibility
(HB1190, /Act 449, 2004), and provision for tax credits and exemptions for private
schools (HB1133/Act 773, 2008; HB1219/Act 618). Most detrimental and central
to his educational plan were austerity cuts that from 2003 to 2010 stripped 4.5 bil-
lion dollars from public education (Croft 2013, 60; Henry and Pope 2010, A16). Whereas traditional public schools serving predominantly African American and
low-income students were exposed to a political onslaught of financial depriva-
tions that led to reductions in staff, resources, professional development funds,
and furloughed days, charter and private schools and the students they served were
sheltered by legislative maneuvers and financial appropriations (Croft 2013). These
legislated actions, coupled with the perceived school failure exposed by NCLB’s
accountability system, fostered the perception that public education was foundering
and that the only remedy lay in promoting accountability through high-stakes
testing and teacher evaluation.

High Pressure Front: The Testing Industrial Complex
The testing industrial complex (TIC), an attempted system of education reform
catalyzed by standardized testing that emerged with NCLB, is a high-pressure front
that creates ideal storm/reform conditions for education at the state and national
levels. Over 10 years of NCLB policies yielded insignificant gains (if any) in
student achievement, and the federal government began to realize the mathematical
impossibility of expecting all children to reach a standardized proficiency level.
Yet, despite the colossal failure of the policy, in 2009 the Obama administration
attempted to salvage it with the creation of RT3, a program that created opportunities
for states to apply for NCLB waivers. The granting of a waiver was attached to
a state’s willingness to implement neoliberal policies such as establishment of
charter schools and increased teacher accountability through standardized testing.
Hence, RT3’s funding, waivers, and neoliberal policies have been integral to the
advancement and institutionalization of the TIC. Here we define the TIC and
describe the warm and cold fronts that bring it to bear.

The Testing Industrial Complex
The testing industrial complex directly relates to (and emulates) foundational
elements of the prison industrial complex, such as: (a) the use of surveillance and
unwarranted policing to feed punitive reform measures used to solve what are in reality
economic, social, and political problems; (b) a confluence of bureaucratic, political, economic, and racialized interests with the underlying purpose of diverting profit from public entities to private corporations; (c) increases in high-stakes outcomes and curricular coopting, even though neither has shown to improve outcomes; and (d) a perception that the complex is practically impossible to dismantle (see Roberts in press). Here, we will focus on three major interconnected factors that make up the TIC:

- Excessive high-stakes testing;
- False political narratives about improving education; and
- Transfer of curricular and financial governance from individual to local, local to state, and state to national/private entities.

**Excessive high-stakes testing.** Never before in the history of the United States have we based so many key education policy decisions on test score outcomes. Across the United States, high-stakes testing policies have caused a trickle-down effect in which politicians put pressure to increase standardized test scores on school boards and superintendents, superintendents put pressure on principals, principals on teachers, and teachers on students—all to little or no avail. Meanwhile, the psychological and academic stakes for failing these tests have become far too high. The examples are rampant: children’s loss of sleep and illnesses during test season, students’ academic disengagement, school closures in marginalized communities, and teacher/principal job losses are just a few of the outcomes of the current testing system (Ahlquist, Gorski, and Montaño 2011; Au and Tempel 2012; Farley 2009; Kohn 2000a, 2000b; Lipman 2004; Swope and Miner 2000).

Ironically, although over 12 years of evidence tell us that focusing exclusively on measurement, accountability, and standardized testing has produced outcomes contrary to those stated, we forge ever onward in the effort (Ravitch 2010; Sacks 2001). For example, the state of Georgia announced its intention to pull out of Common Core–regulated standardized testing; yet it has done so not because of the testing’s shortcomings, but only to save money. In the same announcement, the Georgia Department of Education stated its intent to continue the testing onslaught by creating its “own standardized assessments aligned to GA’s current academic standards” (GADOE 2013). The test, Georgia Milestones, is being administered to Georgia’s 1,702,750 students in 2014–2015.

**False political narratives about improving education.** Educational institutions—e.g., the US Department of Education (DOE), the Georgia Department of Education, the Georgia Professional Standards Commission, legislative bodies such as the American Legislative Exchange Council, and corporate entities such as the Bill & Melinda Gates Foundation, the Broad Foundation, and the National Council for Teacher Quality—insist that raising educational standards, improving school systems, and closing the education gap for students of the global majority are their only concerns.
However, the actual existence of such needs, the diagnosis of the reasons behinds these needs, and the ways these so-called needs should be addressed are all based on a neoliberal ideology promulgated through the TIC (Giroux 2012; Roberts in press). The aforementioned entities use print, digital, and social media to discredit teachers, teacher preparation programs, and schools in order to make the general populace believe that “something must be done about education.”

Even if we choose to believe that US public education is inadequate, there are areas in the country and the world that have had great success in their educational endeavors, usually by providing increased social services for students and their caretakers or rich professional development opportunities for teachers (Berliner, Glass, and Associates 2014; Ravitch 2010; Ripley 2013; Sahlberg 2011). Notwithstanding this evidence, participants in the TIC rely on false narratives of failure that provide the rationale for excessive high-stakes testing and neoliberal reforms (Berliner, Glass and Associates 2014; Fair Test 2014; Paarlberg 2012).

US Secretary of Education Arne Duncan often espouses a politically correct narrative in which he lauds teachers and teaching; yet, his true understanding of the source of the problem with public schooling is revealed by his support of policies that evaluate teachers through high-stakes testing of students, and in quotations such as the following. During an interview with Rolland Martin for the show Washington Watch, Duncan stated:

“I spent a lot of time in New Orleans, and this is a tough thing to say, but let me be really honest. I think the best thing that happened to the education system in New Orleans was Hurricane Katrina. That education system was a disaster, and it took Hurricane Katrina to wake up the community to say that “we have to do better.”” (Ravitch 2014a)

Duncan (2011) also showed his disdain for public educators and espoused a false narrative of failure when he described the transformation of Englewood High School in Chicago during the Teach for America 20th Anniversary Summit:

Same children, same community, same poverty, same violence. Actually went to school in the same building with different adults, different expectations, different sense of what’s possible. Guess what? That made all the difference in the world.

Although there are multiple problems with Duncan’s statements, a large one is that they clearly overlook the causes of the poverty and violence he mentions and imply that those ills can be fixed by changing a school and its faculty. Further, the transformation of the school to which Duncan refers did not really include the same children, because the new charter school forced a number of them out (Rubenstein 2011). Duncan’s verbal jabs and use of false political narratives throughout his tenure demonstrate a belief that educators and public schools are inadequate at least, and usually disastrous.
In Georgia, the false narrative of failure recurs in the words of government officials who ignore structural inequities, such as the huge influence of systemic poverty, and instead blame teachers for student underachievement and then create legislation to “get rid of those bad teachers” (Arnold n.d.). Its most evident manifestation is often found in the front-page headlines of The Atlanta Journal-Constitution, a premier Georgia newspaper. Headlines have included: “Georgia scores low grade for content preparation of elementary teachers,” “Is anyone surprised at a critical review of how we train teachers? Are teachers?”, “Teachers to be graded on student test scores: Controversial new ratings,” “No four-star teacher prep programs in Georgia,” “Learning curve: Teachers too often MIA,” and “Middling marks for teacher training in Georgia. Why can’t we improve it?” (AJC.com).

Transfer of governance. Standardized testing demands a standardized curriculum to ensure the attainment of testing goals. The financial “encouragement” by the DOE toward reforms like the Common Core and value-added models (VAMs) of teacher evaluation through RT3 has resulted in a direct transfer of curricular and financial decision-making power from individual to local, local to state, and state to national/private entities. Although it denies it, through RT3 federal reforms and acquiescence from partner states such as Georgia the DOE has created a national curriculum (Strauss 2013).

High-pressure front. High-pressure fronts frequently manifest as a warm front (precipitation and fog) followed by a cold front (narrow) bands of thunderstorms and severe weather. The TIC warm front rains down neoliberal education polices under the guise of improving education while obscuring the free-market ideology of corporatization, standardization, and privatization as well as the reforms’ real intent—financial gain. For example, corporate CEOs have created educational foundations and brought forward unqualified educational spokespeople supported by corporate money. Espousing the intent to improve education for students of the global majority, these foundations promulgate a large amount of TIC-informed education policies and spew a fog of money that makes it hard for the average individual to see the true value of public education or the record amounts of financial profit generated by such policies (Karp and Sokolower 2014; Ravitch 2010).

After the warm front, a TIC cold front follows. This front manifests as the severe weather forcing local school systems to lay off teachers, close neighborhood schools, eliminate art and music programs, and dedicate more and more revenue to supporting standardized testing. Meanwhile, black, brown, and poor people are most grievously injured because the high-pressure front of the TIC weighs disproportionately on their backs and their communities (Fair Test 2010).

The Mesoscale Evaluation System

Riding on the crest of the TIC high-pressure front are national and local demands for accountability that provide justification for high-stakes evaluation tools and
for what we call a mesoscale evaluation system. In nature, a mesoscale storm is comprised of individual storms that combine to form a larger persistent/perfect storm. Similarly, a mesoscale evaluation system is a combination of individual evaluation efforts spanning kindergarten through higher education that are meant to serve as mechanisms of accountability for educators and educator preparation.

Just as a single drop of rain or gust of wind may not be inherently destructive, some tools of accountability proffered by education, legislative, or corporate entities may indeed be plausible and useful. In fact, such reform efforts claim to establish comprehensive standards aimed at professionalizing education and incentivizing the formation of career- and college-ready graduates who can better compete in a global market. Achieving such goals requires an interconnected system of high-stakes testing as the basis for determining the effectiveness and preparedness of teachers and teacher preparation programs.

Here we argue that prior to the introduction of RT3, education, legislative, or corporate entities might have individually attempted to reform education on local, state, or national levels. Currently, however, these entities, galvanized by the discourses on the failure of public education described earlier, have aligned in an unprecedented manner to aggressively advance a new era of reform. The apparently distinct but actually interrelated reform measures, when combined, comprise an overarching mesoscale evaluation system that brings destruction to teacher preparation programs and K-12 public education. In addition, the reforms mostly affect students and teachers already marginalized on the basis of race, ethnicity, language, economic status, and ability. Continuing to use Georgia as a paradigmatic case, we discuss the tools that comprise the mesoscale evaluation system at the national, state, and local levels. Table 1 highlights the various components of the mesoscale system and illustrates how they manifest in K-12 education and higher education, specifically teacher education.

National

Neoliberal ideas have recently reemerged in education in an effort to remedy the perceived loss in US global economic competitiveness and education prominence. RT3 has served as a catalyst to promulgate market-driven ideals of standardization, corporatization, and privatization designed to comprehensively alter the delivery of public education. Constitutionally, the US government cannot mandate state adoption of education reform initiatives. However, federal dollars are often a powerful decision-making factor in states’ educational policy, and their allocation may be contingent on the state’s implementation of education reform. As Georgia’s Governor Nathan Deal noted in an April 1, 2011 press release after the state won its RT3 bid, “In any year, this grant [RT3] provides a great opportunity to pursue new ideas for improving education, but in tough budget times such as these, this grant is truly extraordinary.”
Table 1: The Mesoscale Evaluation System

<table>
<thead>
<tr>
<th>Components of the Mesoscale System</th>
<th>In K-12 Education</th>
<th>In Higher Education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discourse of failure (used to justify need for increasing effectiveness and preparedness)</td>
<td>Effectiveness</td>
<td>Preparedness</td>
</tr>
<tr>
<td>Evaluation methods (used to determine educator and preparation program effectiveness based on quantitative and qualitative measures)</td>
<td>Teacher/Leader Keys Evaluation System (T/LKES); Student Learning Objectives (SLOs)</td>
<td>Teacher Preparation Program Effectiveness Measure (TPPEM)</td>
</tr>
<tr>
<td>Assessment tools (used to assess in-service teacher effectiveness and pre-service teacher preparedness)</td>
<td>Teacher Assessment of Performance (TAPS)</td>
<td>ed Teacher Performance Assessment (edTPA)</td>
</tr>
<tr>
<td>Standards (designed to encourage rigor, coherence, and consistency in curriculum and in educator preparation)</td>
<td>Common Core State Standards (CCSS)</td>
<td>The Interstate Teacher Assessment and Support Consortium (InTASC); Council for the Accreditation of Educator Preparation (CAEP)</td>
</tr>
<tr>
<td>Consortiums/corporations (hired via competitive bids to develop and disseminate evaluation tools as well as compile and analyze evaluation data)</td>
<td>Partnership for Assessment of Readiness for College and Careers (PARCC); SMARTER Balanced Assessment Consortium (SBAC)</td>
<td>Pearson Education</td>
</tr>
<tr>
<td>Catalyst (used to connect the components of evaluation via federal and state funding)</td>
<td>Race to the Top (RT3)</td>
<td>Race to the Top (RT3)</td>
</tr>
</tbody>
</table>

At the same time that the federal government started implementing RT3, higher education accreditation bodies were in the process of changing their national teacher preparation standards. Specifically, the Council for the Accreditation for
Educator Preparation (CAEP)—the merger of two former accrediting agencies, the National Council for the Accreditation of Teacher Education (NCATE) and the Teacher Education Accreditation Council (TEAC)—produced the 2013 CAEP Accreditation Standards, designed to use accreditation to “leverage change in teacher preparation and [help] ensure that our students are prepared to compete in today’s global economy.”

Even though RT3 and changes in higher education accreditation bodies occurred independently, they conveniently aligned with the aforementioned push for standardization, corporatization, and privatization. Jointly, NCATE and TEAC accredit programs in over 800 public and private institutions, which suggests that CAEP will have a comparable reach.

The CAEP standards promote criteria to assess teacher education programs aimed at increasing teacher quality, student academic success, and recruitment of teachers of color. The tool to achieve such aims is a standardization of the teaching and teacher preparation process according to RT3 expectations, as evidenced by CAEP’s stated goal of determining “P-12 student learning and development using multiple measures, e.g., value-added measures, student-growth percentiles, and student learning and development objectives required by the state.” The use of accreditation criteria and other measures of accountability and the standardization of education are likely to undermine the articulated aims of reform. The proposed logic assumes that raising and standardizing quantitative criteria equates to creating a better teacher; yet this does not mitigate the problem of current reform trends that narrow curricula or endemic issues such as inadequate teacher salaries, limited attention to social inequalities, and deprofessionalization, discouraging prospects for curricular innovation that might meet the needs of students and educators (Crocco and Costigan 2007; Dilworth and Coleman 2014; Milner 2013).

State

In exchange for a waiver from unmet NCLB requirements, states committed to altering the delivery of public education by implementing Common Core State Standards, and they revised teacher evaluation systems for K-12 teachers, state-wide assessment systems for determining student career and college readiness [i.e., Partnership for Assessment of Readiness for College and Careers (PARCC), SMARTER Balanced Assessment Consortium (SBAC)], and comprehensive evaluation systems for teachers and leaders. All these tools were anchored to achievement measures, value-added models, and students’ standardized test scores. However, researchers, education scholars, and practitioners have questioned the proposed measures’ validity, plausibility, and ability to yield the stated goals of accountability (Baker, Barton, Darling-Hammond, et al. 2010; Cody 2012; Milner 2013). In Georgia, RT3 has manifested itself as a deluge of evaluations: teacher effectiveness evaluation (TKES), teacher preparedness evaluation (edTPA), and teacher preparation program evaluation (TPPEM).
State-level K-12 teacher evaluation. The Teacher Keys Effectiveness System (TKES) categorizes K-12 teachers on four levels: ineffective, developing/needling improvement, proficient, or exemplary. A similar system, Leader Keys, is used for school leaders. Fifty percent of TKES is based on student growth and academic achievement, measured via growth percentiles/value-added models using students’ standardized test scores. Despite enduring concerns about this use of tests, the results of student growth percentiles determine a significant portion of a teacher’s effectiveness. The remaining portion includes results from administrators’ observations [Teacher Assessment on Performance (TAPS)], instructional artifacts, and student surveys on teachers’ instruction (grades 3–12).

As a means of establishing performance-based salary increases (such as those found in corporate models), Georgia is also in the process of confirming a tiered certification system that includes benchmarks such as passing scores on TKES and LKES. Tiered certification is presented as a way to provide educators and leaders with opportunities to grow in the profession; salary increases will no longer be based on higher education degrees but strictly on the results of the evaluation instruments. Not only have performance-based salary increases (i.e., merit pay) not been effective (Ravitch 2010), they also potentially minimize the significance of earning advanced degrees for education professionals. Instead of receiving quality salaries for their already high-stakes work, educators must rely on the acquisition of high student test scores as a narrow means for career advancement.

State-level higher education teacher evaluation. In addition to the evaluation of in-service teachers, a growing concern is high-stakes evaluation of pre-service teachers. To address concerns of teacher preparedness to enter the profession, pre-service teachers in teacher education preparation programs will produce a Teacher Performance Assessment (TPA) portfolio, such as the edTPA adopted in Georgia and 34 other states. In this evaluation, pre-service teachers submit a teaching video and written reflections on their lesson design and implementation. As a part of Georgia’s tiered certification, pre-service teachers will be required to have a passing score for initial licensure. edTPA portfolio scores, rather than the evaluation by the program’s faculty that is currently used, will determine whether pre-service teachers are eligible for certification and entry into the field.

The Stanford Center for Assessment, Learning, and Equity (SCALE) is the lead developer for edTPA. Evaluation Systems, a unit of the Pearson Corporation, will administer the edTPA. Pre-service teachers are projected to pay $300 each for Pearson-trained scorers within the education profession; scorers are paid $75 per portfolio.

State-level teacher preparation program evaluation. Just as teacher evaluations (e.g., TKES) and teacher performance assessments (e.g., edTPA) allegedly measure the effectiveness of individual educators and leaders, the Teacher Preparation Performance Effectiveness Measure (TPPEM) in Georgia intends to evaluate the
effectiveness of teacher education programs (also a condition of securing RT3 funding). Fifty percent of a program’s effectiveness will be determined based on how well its graduates score on TKES. TPPEM, like TKES, is a high-stakes evaluation tool guided by value-added models. As noted above, numerous researchers and test makers report an overreliance on standardized tests and value-added models and warn against using tests in this manner (e.g., Baker, Barton, Darling-Hammond, et al. 2010; Berliner 2013; Cody 2012; Darling-Hammond 2012; Papay 2011; Smagorinsky 2014).

Additional state-level testing changes. Complicating matters, on June 4, 2014, Georgia announced that new tests were to replace the Criterion-Reference Competency Test (CRCT) for 4th–8th grade students and the End-of-Course Tests (EOCT) used to determine eligibility for graduation. The new system of tests, Georgia Milestones, has been recently developed and administered by CTB/McGraw Hill. The company has a five-year contract at $107.8 million beginning Fall 2014. The test will be aligned with Common Core Georgia State Standards (CCGSS). Revamping testing systems is not new for Georgia; however, the timing, in the midst of cumulating changes in current high-stakes reforms, seems ill timed, in part based on the Georgia’s State Superintendent’s acknowledgment that any new rollout brings the possibility of lower scores:

> The increased expectations for student learning reflected in Georgia Milestones may mean initially lower scores than the previous years’ CRCT or EOCT scores. That is to be expected and should bring Georgia’s tests in line with other indicators of how our students are performing.

This announcement introduces yet another variable in the education reform storm. Wide-scale distribution of these tests raises additional concerns about an evaluation system that estimates employability and effectiveness based on the results of tests that will necessarily need adjustments.

Over reported concerns regarding cost, Georgia withdrew from the use of Partnership for Assessment of Readiness for College and Careers (PARCC), subsequently altering RT3 implementation (GADOE 2013). Therefore, as Georgia continues to navigate the costs, drawbacks, and benefits of RT3, local school districts must negotiate the processes and expectations created by the reform.

Local

At the local level, the districts’ implementation of education reform has raised many questions and garnered limited answers. In Georgia, legislative, corporate, and education institutions answer those questions by reiterating the failures of the educators and the woeful academic performances of the students, as well as their confidence in the eventual success of reform initiatives. As a proof, they cite research (often limited) conducted by the very entities that are advancing or benefitting from
the reform initiatives (e.g., the Bill & Melinda Gates Foundation). Nevertheless, a rising tide of concern persists within the education community about the influence of these pervasive changes on local districts. In particular, we discuss here the implications of teacher evaluation system (TKES), Student Learning Objectives (SLOs), and Common Core State Standards (CCSS).

Local K-12 Teacher Evaluation (TKES). In Georgia, the implementation of TKES occurred after a truncated piloting period (five months) with 26 districts followed by summer of data analysis, after which, that same fall, the entire state became subject to mandated evaluation systems and Common Core curriculum expectations. Given the high-stakes evaluative nature of TKES, educators across the state have raised questions about the process, protocols, and key decision makers involved. For example, at a recent education reform conference, Georgia teachers raised questions such as:

- How will the student growth models affect the scores when the content areas compared are so different (e.g., US History in 9th grade vs. 10th-grade Economics)?
- Can teachers fight the results of students who may have received bad scores on a standardized test but have been performing well otherwise?
- How can teachers engage administrators in a discussion about their problems with the TKES evaluation system?

Reflecting not just a local but also a national concern, teachers across the country are asking the same types of questions about their own state-imposed value-added model evaluations. The answers they commonly receive are “We don’t know yet,” or statements of unquestioned confidence in the unproven reform mandates.

Local Student Learning Objectives (SLOs). For tested courses (as determined by the state), student growth is measured in percentiles based on state-level assessments (in Georgia, 4th–8th grade CRCT, based on Fall 2014 Georgia Milestones, and for high school EOCT). For non-tested subjects, local districts are held to developing Student Learning Objectives (SLOs) per each non-tested course, which are to be approved by the state’s Department of Education. Teachers administer pre- and post-assessments, the results of which are submitted to a district evaluator who will determine an end-of-year rating—exemplary, proficient, needs development, or ineffective—based on whether the SLO was met.

Local Common Core State Standards (CCSS). The role of CCSS (and the Common Core Georgia Performance Standards) continues to be a point of contention at the state and local levels in Georgia (Bluestein and Washington 2013; Gillooly 2014; Neely 2014; Smagorinsky 2013). Local concerns regarding the perceived imposition of state standards by the federal government have resulted in ongoing political debate. According to the Georgia Department of Education, “assessment is not supposed to drive curriculum” (at http://www.gadoe.org/); however, in order
to standardize the assessment process, state standards have formed the basis for the tests, benchmarking educational attainment goals and determining subject-matter emphases. For instance, CCSS’ main thrust, to advance literacy across the curriculum, has invited concerns that emphasizing literacy and mathematics (the two areas that have received the most attention) will significantly marginalize other significant subjects such as science, social studies, the arts, and physical education.

The Drawbacks of Converging Reform Efforts

An issue worth of attention is the consequence of concurrently introducing multiple education reforms such as those discussed. Although designed to be complementary, the plethora of elements from RT3, state-level initiatives, and local implementations might obfuscate which variables are leading to the desired outcomes or, to the contrary, which measures have a negative effect on curriculum, evaluation, and student outcomes. The alignment of various evaluation efforts within the mesoscale evaluation system is very appealing, as it seems to minimize the time and energy educators need to expend on executing the provisions. Yet, it is of key importance to note that the emphasis undergirding the alignment is not supported by research. In fact, despite minimal research proclaiming the viability of reform aspects such as value-added models and standardized testing (Berliner 2014), a large amount of research by scholars within and external to the field of education points to the contrary (Berliner 2013, 2014; Milner 2013; see also fairtest.org). Hence, as Ravitch (2014b, 154) asserts, they are actually hoaxes, “a mandate, a legislative mandate, or a program that you must obey but has no evidence behind it,” which is fundamentally undermining rather than enhancing the educational experiences of students, families, and teachers.

The Perfect Storm: Alignment or Assault?

The perfect storm arrived in full in 2015 when many of the theorized and piloted efforts previously described became official and in many instances required by law. Proponents of recent education reform measures claim that their efforts are purposefully aligned to improve educational outcomes for public school students. Although we agree that the education reforms are aligned, we argue that the alignment to date has not and will not improve public education. Instead, the alignment amounts to a direct assault on the bedrock of public education that has been building over time and has accelerated under the guise of accountability.

If education outcomes are the determining factors as to whether educational reforms have been successful, then account after account tells us about the actual, rather than the proposed, results of educational reform since NCLB. Stories of school curricula narrowing (i.e., “teaching to the test”), inadequate funding and depleted human resources, and psychological costs to students and educators have been the telltale results from high-stakes testing and education reform. The perfect storm
has become tantamount to an assault on three major groups: K-12 public schools, public schools of higher education, and the educators and the students they serve.

**Narrowed Curricula**

The pressure of high-stakes testing influences school systems in general, but particularly those in racially, economically, and linguistically marginalized communities, which have attempted to raise test scores through measures such as curriculum narrowing, the elimination of enrichment courses, an increase in skill and drill instruction, and/or rampant cheating (Roberts 2010; Wellstone 2002/2003). Many “low-performing” schools allocate more than a quarter of the year’s instruction to test prep, often resulting in a narrowing of curricula (Crocco and Costigan 2007). This over-emphasis on testing has trickled down to the youngest students, causing some educators to replace much needed playtime with testing lessons. As the Alliance for Childhood reports, “time for play in most public kindergartens has dwindled to the vanishing point, replaced by lengthy lessons and standardized testing.”

The phenomenon of teaching to the test has been amply reported in K-12 to the detriment of students and educators, who are pressured to focus curriculum content on test preparation and further exposed to sentiments of de-professionalization. Milner (2013, 5) suggests that

> [w]hen news and other media report about the effectiveness and ineffectiveness of teachers and teaching based on the rise or fall of test scores and without other necessary information to make well-rounded judgments, the field of teaching is subject to unwarranted public criticism and consequently de-professionalization.

In higher education, the teacher education equivalent of teaching to the test in K-12 is the teacher performance assessment “test” (i.e., edTPA). Au (2013, 25) expresses a rising concern among teacher educators:

> The edTPA is shaping our program in some not-so-healthy ways. Instead of focusing on good teaching, our conversations are quickly turning to how to prepare our students for the edTPA. Our student teaching seminars increasingly emphasize the test’s logistics, choosing the right kind of video segment for the test, choosing the right kind of unit for the test, making sure everyone is using the same language as the test.

Education reform initiatives have caused K-12 schools and colleges of education to reevaluate curricular content, not toward expanding multidimensional learning opportunities, but rather toward adapting to the singular dimension of test expectations. However, school systems across the country are recognizing that less does not equate to more:
Milwaukee Public Schools is one of several school systems across the country—including Los Angeles, San Diego and Nashville, Tenn.—that are re-investing in subjects like art and physical education. The Milwaukee school district is hiring new specialty teachers with the hope of attracting more families and boosting academic achievement. (Toner 2014, para. 3)

A narrowed curriculum emphasizing test preparation has had an additional effect on survivors of NCLB who have been socialized to perform and seek the “right” answer. NCLB-affected curricula have produced students with lower capacities for creativity, problem-solving, effective communication, and critical thinking (all skills reportedly desired by corporations) (Wernert 2013).

Funding Priorities

As a result of high-stakes testing pressures, school systems have adjusted their funding priorities to support testing and testing materials rather than enrichment, recess, and resources for all students. Standardized testing seems to be funded carte blanche; yet student educational outcomes have been either inconclusive or unimproved (Berliner, Glass, and Associates 2014; Fair Test 2014). The outcomes of that failure have frequently been punitive, resulting in school closings, firing of teachers/administrators, and decreased school funding (see CREATE 2013; Fisher 2013). Furthermore, states like Georgia that claim a lack of available funding, and therefore furlough teachers, cut instructional days, and reduce instructional material, do not seem to hesitate to implement education reform requirements that demand additional personnel, time, resources, and development (likely paid via RT3 provisions that will no longer be available after 2015). Education in Georgia, like in many other states, is underfunded except when it comes to money to support neoliberal reforms (Strauss 2014; Suggs 2014).

Psychological Costs

Most discouraging is the reality that education reform has led to negative physical and emotional consequences for students and educators. According to researcher Gregory J. Cizek (2001), anecdotes abound “illustrating how testing ... produces gripping anxiety in even the brightest students, and makes young children vomit or cry, or both.” On March 14, 2002, the Sacramento Bee reported that “test-related jitters, especially among young students, are so common that the Stanford-9 exam comes with instructions on what to do with a test booklet in case a student vomits on it.” A three-year study completed in October 2010 by the Gesell Institute of Human Development showed that increased emphasis on testing makes “children feel like failures now as early as PreK” (at http://www.gesellinstitute.org/). Georgia parent Stephanie Jones (2014) states:
I am well aware of many Georgia families being sick and tired of the hyper-focus on the tests, recess being taken away, Saturday school being mandatory, after-school being mandatory, and summer school being mandatory all in the name of passing some test. Kids are stressed out and anxious, and learning that school is a place where anxiety is normal, and that the only real reason to “learn” something in school is so that you can pass a test at the end of the year.

Student academic engagement and academic outcomes have also experienced serious damage. As pointed out by Senator Wellstone (2002/2003), “the effects of high-stakes testing have [had] a deadening effect on learning.” Wellstone’s words are illustrated by the skyrocketing numbers of students who have given up and dropped out of school because of the inability to pass a gateway test or a feeling of disengagement (Tyler and Loftstrom 2009).

In conjunction with students’ stress and disengagement, K-12 teachers and teacher educators have expressed sentiments of profound demoralization (Santoro 2011). Educators experiencing high anxiety, frustration, and hopelessness have published open letters of resignation (e.g., see Downey 2012), brought lawsuits against the state (e.g., Florida; see http://feaweb.org/teachers-file-federal-736-lawsuit), and have been fired for expressing dissent (Hayes and Sokolower 2012–2013; Madeloni and Gorlewski 2013). Sarah Wiles, a science teacher in Charlotte, North Carolina, as cited by Megan (2014), clearly illustrates teacher demoralization when she says:

I am so tired of being lied to about how important I am and how valuable I am…. I am also sick and tired of politicians making my profession the center of attention and paying it lip-service by visiting a school, kneeling next to a child, shaking my hand and thanking me, telling the nightly news that I deserve a raise, and then proceeding to speak through the budget that I am not worth it. If you aren’t going to do anything, and you know nothing will change, just leave me alone. I would rather be ignored than disrespected.

In Georgia, teachers and teacher educators have written multiple editorials and left numerous comments in local media blogs in which they “speak of the tremendous pain that they feel in being part of a profession that is continually battered by [inaccurate] public commentary from education officials, taxpayers, and other stakeholders from outside the system” (Smagorinsky 2011).

**Conclusion and Implications**

In the face of the perfect storm we have described, our clarion call is not to endure or weather the storm, as educators have done with education reforms of the past. Previous survival techniques of battening down the hatches and waiting for the waves of reform to pass are insufficient to withstand this convergence of storm
elements. Meteorologically, perfect storms are almost impossible to avoid; however, the repercussions may be so severe that, if we simply wait for this storm to pass, when we finally emerge from our hiding places we will find only remnants and fragments of our public schools.

Neoliberal policy making (i.e., privatization, corporatization, and standardization) has dictated current iterations of education policy in the hope that “this time, things will be different.” Yet, historically, we have seen that no matter how idyllic current education reform initiatives appear on paper, they are most likely to leave educators and students adrift, feeling consumed, overwhelmed, and subjected to political finger-pointing, disappointment, disengagement, and shame. In order to secure our best chances that indeed things will be different, we advocate that K-12 educators and teacher educators: (a) escalate actions to stop the eventual and present negative consequences of current education activities, laws, and reforms; and (b) demand the provision of the financial, physical, emotional, and psychological infrastructure that must accompany education reform to achieve authentic, healthy, and sustainable success.

Indeed, just as reforms have been growing nationwide, so have national, state, and local resistance efforts by educators, students, and parents/caretakers. Of many, we share selected exemplars of resistance to indicate how a rising number of individuals and collectives are striving to turn the tides (see Table 2; see also Strauss 2012). At stake is the education of all children: not just the ones who deserve it, not just the ones who do well on tests, not just ours, and not just the ones we like.

Table 2: Exemplars of Resistance to Current Education Reforms

<table>
<thead>
<tr>
<th>National</th>
<th>Georgia</th>
</tr>
</thead>
</table>
| National Center for Fair & Open Testing  
www.fairtest.org | Georgia-United Opt Out  
| National Resolution on High Stakes Testing  
http://timeoutfromtesting.org/nationalresolution/ | GREATER (Georgia Researchers, Educators, and Advocates for Teacher Evaluation Reform)  
http://greater2012.blogspot.com/ |
| Network for Public Education  
www.networkforpubliceducation.org | |
Corporately privatized winds of change have gathered their forces in ways that are deeply disturbing and unprecedented. Yet, there is still time to deflect total destruction. To do so, however, those who are authentic stakeholders must answer enduring questions about education and education reform: What kind of education do we want and need? For whom, for what aims, and at whose expense? Whether the edifice of public education is completely destroyed, rebuilt in the image of corporations in the United States, or saved will be determined by how hard we fight to salvage what is left.

NOTES

1. The term “global majority” is used to represent many populations variously characterized in the United States as minority, at risk, underserved, non-white, of color, urban, of low socioeconomic status, and poor—all terms that are used to mask the hegemony of European American populations and the numeric and political reality of black, brown, and lower-income people worldwide. We use the term “global majority” to reflect a more affirming and accurate sense of the vast diversity of individuals represented in the United States.

2. Although presenting the two governors as political binaries may be a simplification, their tenures differed in significant ways. Barnes’s educational package mirrored the growing trend toward accountability that had been mounting since the 1983 report “A Nation at Risk,” and he served just prior to the full implementation of NCLB. Perhaps indicative of his political acumen, he was able to push through a largely Republican educational agenda, but with equity built in through funding. In fact, one source indicated that many of the ideas were Republican in origin (Brooks 2011). Notwithstanding a heavily swayed Republican agenda, Barnes’s educational platform built in safety nets from K through 12, all of which he funded (Cumming 1999, H8). In essence, even though his educational program reflected a growing concern for accountability, he counterbalanced this demand with built-in equity through comprehensive funding. Ironically, however, despite his equitable distribution of funding, provision of teacher raises, establishment of a career ladder for teachers, and increase in pay for National Board Certified teachers, he is more often remembered as the governor who was opposed to teachers because of his revocation of teacher tenure. In a personal interview, Governor Barnes lamented that he never opposed teachers; he just wanted to make it easier to replace ineffective teachers (R. Barnes personal interview, November 15, 2012).

In contrast, Governor Perdue ended his tenure just prior to the full implementation of RT3. In fact, the US Department of Education awarded Georgia the Race to the Top Grant in August 2010 (http://gosa.georgia.gov/race-top), even though implementation did not begin until 2011 under Governor Deal’s tenure (http://gov.georgia.gov/00/press_print/). As much as Barnes’s educational package represented bipartisan ideology (R. Barnes personal interview, November 15, 2012), Perdue’s educational platform was dominated and supported by a majority Republican legislature. Hence, from 2003 through 2010, a series of measures (including austerity cuts and tax credits) were passed that undermined public schools rather than supporting them. Also, although one of Perdue’s first acts was to overturn the revocation of teacher tenure, during his term classes were increased, school days were shortened, funding for national board certification was reduced, and furloughed days were instituted.

In retrospect, both governors worked in a climate of testing and accountability. One used legislation to support public education; the other used legislation to undermine it. Whereas one governor supported teachers in all but seemingly the most significant way, the other undermined teachers in more subtle ways. One left office shrouded in a legacy of disdain for teachers; the other left office with educators realizing that public education had been significantly weakened. In conclusion, and notwithstanding
Barnes’s stance on teacher tenure, his efforts still reflected a desire to level the playing field (Croft 2013; Cumming 1999, H8).

6. For details, see https://sites.google.com/a/pearson.com/score-edtpa/.

REFERENCES

Ahlquist, Roberta, Paul Gorski, and Theresa Montaño

Arnold, Jim

Au, Wayne
2013 “What’s a Nice Test Like You Doing in a Place Like This? The edTPA and Corporate Education ‘Reform.’” Rethinking Schools 27(4): 22–27.

Au, Wayne and Melissa Bollow Tempel


Berliner, David C.
2014 “Exogenous Variables and Value-Added Assessments: A Fatal Flaw.” Teachers College Record 116(1); at www.tcrecord.org, ID Number: 17293.

Berliner, David C., Gene V. Glass, and Associates
2014 50 Myths and Lies That Threaten America’s Public Schools: The Real Crisis in Education. New York, NY: Teachers College Press.

Bluestein, Greg and Wayne Washington

Brooks, F. Erik

Cizek, G.J.

CReATE (Chicagoland Researchers and Advocates for Transformative Education)
Cody, Anthony

Crocco, Margaret S. and Arthur T. Costigan

Croft, Sheryl J.

Cumming, D.
1999 “Raising the Grade: A look at Improving Schools across the state of Georgia. Governor Touts US Standards Barnes Says State Should Measure up with Rest of Nations.” Atlanta Journal and Atlanta Constitution, June 21: H8

Darling-Hammond, Linda

Dilworth, Mary E. and Marcus J. Coleman

Downey, Maureen

Duncan, A.

Farley, Todd

Fisher, Alexandria

GADOE

Gillooly, Jon

Giroux, Henry A.

Hayes, Nini, and Sokolower, Jody
Henry, J. and J. Pope  

Jones, Stephanie  

Karp, Stan and Jody Sokolower  

Kohn, A.  

Lipman, Pauline  

Madeloni, Barbara and J. Gorlewski  

Megan, Kathleen  

Milner, H. Richard  

National Center for Fair and Open Testing (Fair Test)  
2014 “Stagnant Grade 12 NAEP Results Underscore Failure of Test-Driven Public Education; ‘No Child Left Behind’ and State High-Stakes Exams Did Not Lead to Improved Performance.” At www.fairtest.org/stagnant-grade-12-naep-results-underscore-failure.


Neely, Clay  

Papay, John  

Ravitch, Diane  


The Perfect Storm of Education Reform

Ripley, Amanda

Roberts, Mari A.


Rochford, J.A.

Rubenstein, Gary
2011 “Same Kids, Same Building, Same Lies.” At http://garyrubinstein.wordpress.com/2011/03/06/same-kids-same-building-same-lies/.

Paarlberg, Michael

Sacks, P.

Sahlberg, Pasi

Santoro, Doris A.

Smagorinsky, Peter


Strauss, Valerie


Suggs, C.  

Swope, Kathy and Barbara Miner  

Toner, Erin  

Tyler, J.H. and M. Lofstrom  

Wellstone, P.  

Wernert, Sean  

Acts

Act 394, 2008 Ga. L. 82, § 1/HB 1209  