
Intelligent Accountability: Re-Thinking the Concept of “Accountability” in the Popular Discourse of Education Policy

Scott Ellison

Pellissippi State Community College

I believe that the quality of our education system says as much about the long-term health of our economy as the stock market, the unemployment rate and the size of the gross domestic product. That’s because the quality of our work force and the intellectual breadth and depth of future leaders is directly related to the quality of education we provide today. So I begin...
by recognizing America’s common agenda to promote economic security through education. (U.S. Secretary of Education Arne Duncan)

In prepared remarks before the United States Chamber of Commerce’s “Education and Workforce Summit” in November of 2009, U.S. Education Secretary Arne Duncan expressed a sentiment now dominant in the popular discourse of education reform in the United States linking the economic fortunes of the U.S.A. in the global economy to the human capital produced by its education sector.¹ This perspective envisions economic success as emerging from an educational system that produces new generations of technological and scientific innovators and high-skill workers to drive an economic sector geared toward providing high value-added goods and services on a global marketplace.² For K-12 schooling, the challenge would appear to be to produce students possessing the foundational knowledge and skills of literacy, mathematics, and critical thinking required for university and technical training.

To that end, Duncan offers straightforward proposals for educational reform that require states to set rigorous, easily-measured academic standards, create assessments linked to those standards, use data generated by those assessments to raise academic achievement,

and, finally, to hold educational stakeholders [administrators, teachers, and students] accountable for failure. Most commonly subsumed under the conceptual rubric of ‘accountability,’ this model of education reform has become the default policy position of both political parties in the U.S. and has become the commonsensical position in popular debates over education reform. It is a policy position built on the assumption that it is possible to dramatically raise academic achievement in public education by re-organizing its incentive structures. However, critics of these reform policies charge that placing an undue emphasis on easily defined academic standards tied to standardized assessments and accountability regimes will lead to a narrowing of school curriculum to the easily assessed, the regimentation of schooling, the re-enforcement of teacher-centric pedagogies that provide little opportunity for student discovery and critical thinking, and an over-reliance on rote learning. It is from this indeterminate situation in debates over education policy that this inquiry finds its impetus.

The task of this article is to unpack the concept of accountability in order to clarify and critique the logic of this educational and political concept. To accomplish this task, I will employ a synthetic method of analysis that will, first, situate accountability within the larger framework of standards-based education reforms of which it is an integral element. From there, the second step is to examine the research literature in order to interrogate standards-based policy reforms at each point in its logical chain so as to unpack the unquestioned assumptions and problematics inherent to the concept of accountability that are often obscured by contemporary educational discourse. The results of this study suggest that critics of accountability policies are well justified in their concerns.

A Synthetic Method of Inquiry

The methodology for this inquiry is grounded in an ontological observation into the internal contradictions between the conceptual norms with which modern societies take their general orientations and the concrete realities generated by those conceptual understandings. This methodology is itself rooted in a long tradition of praxis philosophy tracing its lineage to the works of Hegel.³ In its most general sense, the synthetic mode of inquiry first outlined by Hegel and employed in this essay takes as its subject the concepts, theories, and ideas that represent the ‘spirit of the age,’ the conceptual knowledge through which individuals become conscious of their worlds and from which societies seek justification for concerted social action and governmental policy.

A synthetic method of inquiry can be characterized as being what Dewey called a “double movement” to and from meaning in which critique is located within the reflexive movement between the conceptual and concrete.⁴ Synthesis begins with a practical concept at work in popular discourse as it is most readily present as a discursive practice. From there the inquirer pulls apart the concept through an analysis of the concrete realities that is its subject. The product of this deductive movement establishes the foundation for the transformation of the concept under analysis, and hence the concrete realities appropriate to it, by providing the inquirer with the raw material for an inductive movement of re-conceptualization.

The synthetic mode of inquiry employed in this essay denotes a heuristic consisting of three “moments” of inquiry (definition, division, and synthesis) in which the inquirer works through the conceptual logic of the subject of inquiry in order to unlock its transformative potential, to put the subject of inquiry “to the test of reality, of contemporary reality, both to grasp the points where change is possible and desirable, and to determine the precise form this change should take.”⁵ The “double movement” of a synthetic analysis begins with an indeterminate situation in which cemented “common sense” ideas run aground the concrete realities of a contradictory, dynamic present-history.

Moving from an indeterminate situation of conceptual failure and contradiction, inquiry begins with a definition of a concept in its most immediate form. From definition, inquiry moves through the logic of this conceptual framework in a deductive movement that articulates the complex determinants and processes conditioning its actualization in the concrete [division]. Moving inductively from this empirical base, the concept is then synthesized into a conceptual framework through which the indeterminate situation from which inquiry began is resolved and through which the epistemic validity of its normative claims are to be established.⁶

Thus, a synthetic mode of inquiry returns to its starting point as a rich totality of determinations built around a rigorous investigation of the empirical evidence relevant to the subject of inquiry and the reconciliation of the conceptual norms implicit to the subject of inquiry and the practices that seek justification in those normative ideals.

In this inquiry, I will begin by first defining the concept of accountability by situating it within the standards-based education reforms of which it is an integral element. To do so, I will examine popular texts produced by prominent advocates of standards-based reforms so as to flesh out the concept of accountability as it is employed in popular debates over education policy and reform. From there, I will peer beneath the

veneer of standards-based education reforms to examine the efficacy of those reforms in achieving the academic gains and functional rationality from which it seeks justification. I will accomplish this by examining peer-reviewed research literature so as to test the epistemic validity of accountability at each point in its logical chain, identifying moments of contradiction and epistemic failure. From there, I will then work back through the conceptual logic of accountability in order to establish the foundation for the reconciliation between the normative ideals animating the concept of accountability and the practices appropriate to those ideals. In the final section, I will conclude the article with a critical reflection that seeks to flesh out some of the larger lessons to be taken away from this synthetic analysis.

Standards, Assessment, and Accountability

Returning to Duncan's remarks, it would appear that the key to ensuring that students are taught the foundational knowledge and skills of literacy, mathematics and critical thinking required for success in a 21st century economy begins with a set of rigorous academic standards for those subjects. Kosar makes the case for a standards-based approach to reforming educational practices this way:

[T]here is a powerful [behaviorist] logic to standards-based reform: children will not learn to high levels unless they are taught challenging curricula... [T]o raise achievement, the level of skills and knowledge students are taught must be raised, and this can be done through establishing challenging education standards. Doing this will maximize the probability of good teaching or worthwhile content to all students. And the children will respond.⁷

The working assumption of this approach is that creating clearly defined, rigorous academic standards aligned to high-stakes assessments will push schools toward sound curricular and pedagogical practices that will then lead to higher student achievement. If the nation's system of public education is to meet the challenges of globalization, then federal and state policy makers should set high standards for achievement to ensure that students are taught the foundational knowledge they will need to compete in an information-based economy.

To ensure that students, administrators, and schools maintain this high level of achievement, a 21st century system of public education requires a rigorous regime of performance-based accountability. Such a regime begins with the necessary "mission, values and guiding principles" required to create a broad framework of academic and performance standards for students, administrators, and schools.⁸ These standards

would form the foundation for the design of instruction, curricula, and high-stakes assessments⁹ “that have real consequences for professionals in the schools.”¹⁰ Hess observes: “Performance information is meaningful only when it is backed by real incentives; when principles and superintendents use it to reward or penalize teachers based on how well they are serving their students.”¹¹ For teachers, accountability means that their performance evaluations would be linked to student performance on standardized assessments. Likewise, administrators and entire schools would be judged by student performance on standardized assessments with the explicit threat of dismissal of administration or the possibility of a complete re-structuring of schools either by quasi-privatization (e.g., charter schools) or re-constitution of school staff.

Beneath the Veneer: Research Literature

Beneath the veneer of these standards-based reforms lies an assumption of a functional reflexivity. Clear, measurable academic standards provide the foundation for the construction of rigorous curricula, sound instructional models, and assessments that will challenge students to achieve at high levels. Standardized assessments measure students’ achievement in reaching the prescribed academic standards, measure teacher and school effectiveness, and provide the necessary data for re-aligning curricular and pedagogical practices in order to address students’ academic needs. In turn, the tough accountability measures attached to performance on those assessments provide incentives for students, teachers, and administrators alike in ensuring student achievement. Put more simply, the assumption here is that standards-based reform will foster a rational organization of schooling that will increase student achievement. It is an assumption of reinforcing incentives where standards are the carrot accompanying the accountability stick and rational education decisions are made possible by the empirical data provided by objective assessments. However, a review of the research literature offers little to support these assumptions.

Sloan and Kelly identify two controversial issues involved with standards-based reforms and high-stakes assessments.¹² The first of these issues lie with the desirability [or even possibility] of developing clear, easily measured standards as the foundation of instruction, assessment, and accountability.

Content mastery at some level is a cognitive event: the understanding of powerful, complex, and sometimes fuzzy ideas. For that reason, at least for challenging content, it may be difficult to write clear and simple standards, thereby making their operationalization for curriculum

development, test construction (of any genre, objective or authentic), and alignment between the two problematic.¹³

On the one hand, the logic of a standards-based framework requires challenging academic content as an incentive for high academic achievement; on the other hand, standardized assessments require clear, easily measured criteria by which to judge student success. Indeed, *prima facie*, it stands to reason that in the process of articulating challenging academic content into clear, easily-measured standards a good deal of the complexity, ambiguity, and contradictions that make content challenging to begin with would be lost. More importantly, even assuming the desirability of constructing clear, easily-measured academic standards, the types of assessment most often associated with standards-based reforms raise issues that call into question the assumption of reflexivity between these assessments and educational decision-making processes.

Sloan and Kelly point out that different types of assessments test different skills and serve different purposes. The multiple choice assessments associated with standards-based reforms in the U.S. are most often constructed within a psychometric framework designed to produce “economically tractable and defensible reliability indices” for ranking and norming purposes.¹⁴ Constructed along a Bell curve, it would appear that these types of assessments are far better suited to scaling students than providing useful information capable of improving student learning. Yet a key assumption of standards-based reform is that these “scientifically” constructed assessments must do exactly that, i.e., improve student learning.

At the district and state level, standardized assessments are being increasingly used to differentiate between effective and ineffective schools and individual teachers. While this may seem a rather straightforward and intuitive means of measuring effectiveness, it is far from rigorous. Measuring school proficiency/deficiency based on standardized assessments assumes a causal relationship between differences in student test performance and differences in the effectiveness of instruction. A school accountability framework based on standardized test scores alone fails to recognize other academic inputs, such as student-teacher characteristics, instructional practices, resources, curriculum, etc., involved in academic achievement and test performance. Robert Linn notes that, despite the rhetoric of many reformers, using standardized test scores to distinguish between effective and ineffective schools is decidedly unscientific.¹⁵ As a tool for judging the effectiveness of individual teachers, the National Academy of Sciences warns against the over-emphasis of standardized assessment scores as the primary mode of evaluation in Duncan’s *Race to the Top Initiative*, noting that, “while standardized tests

are helpful in measuring a reform's effects, evaluations should rely on multiple indicators of what students know and can do, not just a single test score."¹⁶ Further, in a more general sense, determining proficiency and deficiency using data from standardized assessments must *a priori* involve some sort of ranking process that requires a determination of "cut points" in the distribution of student scores that is, as Andrew Ho notes, necessarily "judgmental."¹⁷

Rupp et al. examine the common approach of standardized assessment designed to measure reading comprehension skills, which is to ask students to respond to a text passage with multiple-choice questions. Their findings suggest that students approach these questions as problem solving tasks and that these assessments actually test a rather limited range of reading skills. "[D]ifferent [multiple choice] questions do not merely tap but, indeed, create very particular comprehension and response processes. Therefore, a blanket statement such as '[multiple choice] questions assess reading comprehension' is nonsensical for any test."¹⁸

In order to evaluate the efficacy of standardized assessments as pedagogical tools Rupp and Lesaux compared measures of reading achievement produced by standardized assessments and the component skills of literacy as identified by a "diagnostic battery of standardized and experimental tasks."¹⁹ Their findings suggest that the proficiency classifications of standardized assessments mask a great deal of heterogeneity of reading skills at all levels. This raises two serious issues. First, it questions whether standardized assessments provide adequate measures of student reading comprehension skills, and, second, it calls into question whether standardized assessments provide actionable information for improving student achievement. "Given the reliance on standards-based assessments to guide educational decision-making... there is a need to seriously consider whether the properties of these tests support any interpretation at the level of the individual and similarly whether there is any instructional information to be gleaned from the results."²⁰

Looking at the component reading comprehension skills of students identified as failing by standardized assessments, Bully and Valencia also find that standardized assessments mask variations in the component skills of reading.²¹ By masking the considerable heterogeneity of student reading skills, standardized assessments appear not only to provide a limited picture of students' reading comprehension, but they also fail to provide the necessary data for tailoring instruction to students' specific needs and contribute to student achievement. Further, Bully and Valencia question the utility of policy decisions mandating specific instructional strategies and curricular programs for failing schools, such as Reading First, that are becoming increasingly common

in standards-based reforms. “Our data suggest that such policies aimed at specific, predetermined, instructional approaches cannot possibly fit the various needs of at-risk children.”²² In short, on both the student and school levels, it appears that standardized assessments are being put to tasks for which they are neither designed nor well-suited.

Turning now to the assumption that the incentive structure created by attaching high-stakes consequences to standardized assessments will lead to achievement gains, the research literature offers little supporting evidence. Examining state-level data, Nichols et al. investigate the relationship between high-stakes assessments and student achievement. They find that assessment pressures have a positive impact on 4th grade mathematics but no impact on 4th grade reading or 8th grade reading and math, suggesting that accountability pressure may only impact skills that can be effectively taught in a direct instructional approach.

[O]ur findings (and lack of findings) lead us to the conclusion that high-stakes testing pressure might produce effects only at the simplest level of the school curriculum: Primary school arithmetic where achievement is most susceptible to being increased by drill and practice and teaching to the test.²³

Similarly, Merchant et al. find limited evidence that high-stakes assessment policies raise achievement only in the most basic skills but no evidence that tough accountability policies raise achievement in more advanced subject areas.²⁴ Lee and Wang find no evidence of achievement gains associated with high-stakes assessments nor a significant change in the distribution of achievement across student groupings.²⁵ Overall, results from my review of research literature examining state-level data do not support “any argument that high-stakes testing is necessary to raise student achievement.”²⁶

Interestingly, there is a growing number of studies that suggest that not only do high-stakes assessments have little impact on gains in student achievement overall but that implementation of these policies leads to strategic responses by educational players that generate significant social costs. Linking rigorous accountability regimes to standardized assessments creates powerful incentives to narrow instruction to only those curricular areas being assessed and to the wide-spread adoption of teacher-centric ‘drill and kill’ instruction.²⁷ More troubling, there is evidence that the incentives created by standards-based accountability reforms lead to perverse outcomes and gaming strategies. Through grade retention and inadequate reporting of graduation rates, high schools serving even the most challenging student bodies were able to raise average student achievement, but it is a success that appears to come

at the cost of the lowest performing students, who eventually disappear or drop-out.²⁸

In sum, the research literature fails to provide compelling evidence that the high-stakes tests and tough accountability measures associated with standards-based reform increase student achievement. Even assuming the desirability of constructing clear, easily measured standards, the assessments aligned to those standards fail to provide the rich data educators require to raise the achievement of struggling students. More troubling, the incentives created by the accountability measures associated with standards-based reforms appear to generate significant social costs in the form of higher retention rates and increases in the number of drop-outs. Thus, the functional reflexivity assumed by the standards-based reforms is fractured at each point in its logical chain. The concept of accountability appears to be internally contradictory and flawed. However, at each point of conceptual failure, the division of accountability demonstrates not only the failure of standards-based reforms to fulfill their own normative goals. It also provides us with a road map for constructing new policy ideas capable of making those goals an empirical reality.

Synthesis

If we take the mathematical-scientific knowledge, literacy, and creative-innovative thinking advocated by standards-based reformers as being the foundation for an intelligent system of accountability, then the preceding examination of the popular conceptualization of accountability opens up moments of possibility to break out of cemented conceptions and practices to develop new policy approaches capable of achieving the normative goals of preparing students for the intellectual demands of the 21st century. It is to this task that we now turn.

Rethinking Standards

As Sloan and Kelly point out, complex ideas and knowledge are necessarily fuzzy in that they denote “cognitive events” as opposed to being bits of data to be stored on organic hard-drives. Mathematical-scientific knowledge denotes more than having a grasp of the mechanics of mathematical computation and the ability to carry out formulaic “experiments” as one might follow a recipe. Instead, it denotes the ability to employ mathematical reasoning and the scientific method in such tasks as developing new technologies, answering scientific questions, and developing new materials. In short, mathematical-scientific knowledge is not an invisible cookbook to be carried around in a cognitive knap-

sack. It denotes an intellectual skill set that empowers one to utilize mathematical and scientific thinking in productive human activities of problem solving, knowledge construction, and creation.

Likewise, literacy surely denotes more than the simple decoding of information from text or the ability to solve vocabulary and analogy puzzles. Literacy denotes an intellectual skill set. It denotes the ability to interpret and communicate complex ideas through the use of symbolic systems that constitute the core of conscious thought, communication, and intellect. More specifically, literacy involves the decoding, processing, and interpretation of written text in a process of critical thinking in which the outcome is substantially more than the sum of its parts. Literacy involves the ability to organize one's thoughts in a logical framework; to symbolically represent those ideas so that others can decipher them; and to anticipate the needs, perspectives, and skills of an unknowable audience. In short, literacy denotes a mode of conscious thought and symbolic communication.

As a curricular goal, creative-innovative thinking proves to be an even "fuzzier" concept. Denoting more of an intellectual disposition than curricular knowledge in the traditional sense, creative-innovative thinking would appear to demand that the previous two curricular standards, mathematical-scientific knowledge and literacy, be achieved in such a way so as to foster curiosity, a love of learning, problem-solving skills and creative expression. Or, to put it more generally, creative-innovative thinking denotes a mode of intellectual life. Perhaps a better model for this broad curricular goal can be found in Thayer-Bacon's conception of *constructive thinking* as an organizing principle that places equal emphasis on reason, creativity, intuition, collaboration, and care as necessary human tools of creation, innovation, and problem-solving.²⁹

Working through the logic of the curricular goals identified by standards-based reform advocates as being the necessary foundation for rigorous accountability through the lens of the research literature allows us to draw some broad conclusions about the nature of the academic standards that are appropriate to the task. Developing standards from the broad curricular goals of mathematical-scientific knowledge, literacy, and creative-innovative thinking would appear to require that those standards be broadly defined and de-coupled from a direct correlation with individual grade levels. If these broad curricular goals denote an intellectual skill set, a mode of conscious thought and symbolic communication, and a mode of intellectual life respectively, then it is clear that standards formulated from those goals must be articulated at a level of generality appropriate to the "cognitive events" they describe. Further, the uneven intellectual development and maturation of individual stu-

dents would seem to preclude the utility of linking individual standards to any one specific grade level or calendar year. Perhaps a more appropriate model would entail the use of benchmarks situated at transitional points in schooling, such as the transition from primary to secondary schooling. However, what is most important for the current study is to note that the broad curricular goals of creative-innovative thinking, mathematical-scientific knowledge, and literacy do not lend themselves to clearly defined academic standards that can be assessed on large-scale tests employing a multiple choice format. An accountability system constructed on the foundation of these curricular goals must employ an intelligent system of assessment that balances the needs of policy-makers and communities with the learning needs of individual students.

Re-Thinking Assessment

Thus, a constructive system of assessment would deploy a twofold strategy. At the classroom level, student assessment should primarily involve the use of authentic, formative assessments geared toward student learning. The most appropriate form of assessment for this task would be teacher-constructed performative assessments that provide rich data on student knowledge and skills as well as identify the educational needs to be addressed. In order to gear instruction to increased student learning, teachers require an accurate, in-depth perspective on student achievement and academic needs, and this perspective is best acquired through authentic assessments constructed by teachers for individual students according to their needs. These kinds of assessments would provide teachers not only with a rich perspective on student learning but would also provide an opportunity to assess more intangible skills and dispositions such as creativity, curiosity, and attitudes toward learning. For students, formative assessments provide immediate feedback on achievement and identify areas in which they need to focus their studies. Formative assessments provide students with the necessary feedback to develop meta-cognitive understandings of their own skills, knowledge, and interests, and it affords them the opportunity to take ownership of their learning.

At the district and state levels, the goal is to develop rich data on student achievement covering large populations in a manner that is both cost-efficient and practical. A reasonable way forward in achieving this goal is to adopt a balanced approach that employs more authentic assessments to increase the quality of the data available to communities and policy-makers in conjunction with random-sampling techniques that can gather large-scale data while addressing issues of cost and practicality. The quality of the data available to policy-makers and administrators

would benefit significantly from the use of more authentic assessments that incorporate elements of interactive, performative, and traditional assessment techniques and are capable of assessing the degree to which students can use knowledge as opposed to their ability to re-produce “dead” content. Utilizing random-sampling techniques will make it possible to carry out fewer assessments and produce high quality data as opposed to current trends toward large-scale assessments that yield relatively poor data. Further, valuable data on student and teacher success can be gleaned from broader measures such as tracking graduation rates, student success rates in the tertiary sector, and parental satisfaction.

Of course, carrying out these reforms would require structural shifts and at least an initial expenditure of resources. At the classroom level, a policy push toward encouraging the use of teacher-constructed performative assessments geared toward advancing student learning would require two actions by policy-makers. First, the wide scale use of authentic, performative assessments is predicated on significant investments in the professionalization of teaching through master-teacher and mentoring programs, professional development opportunities for practicing teachers, and an increased focus in teacher training on constructing and utilizing a variety of assessment techniques to increase student learning. Second, constructing authentic assessments tailor-made to individual students requires that teachers have an in-depth understanding of their individual students’ strengths, weaknesses, and academic needs. This would imply that policy-makers should strive to lower student-teacher ratios and experiment with grouping teachers and students together over several academic years, particularly in the primary years. At the district and state levels, implementing large-scale authentic assessments would require a coordinated effort to create a system of intelligent standards and appropriate assessments as well as a significant push to increase the number of educational specialists to carry out assessments and interpret the data. Finally, and perhaps most importantly, it is incumbent on policy-makers to provide the necessary resources to make these policy shifts at the classroom, district, and state levels a reality.

Re-Thinking Accountability

Of course, the critical role public education plays in modern society demands that educational actors and institutions be held accountable for fulfilling their social roles and societal mandates. It is both reasonable and desirable that society articulate a uniform set of curricular standards that do justice to the complexity of the cognitive acts they describe; to construct an intelligent assessment regime to evaluate the

degree to which individual teachers, schools, and districts are helping students achieve those standards, to reward educational actors for success; and, finally, to hold educational actors accountable for failure. That said, the devil is in the details.

In light of the previous analysis, the questions raised about the efficacy and desirability of accountability policies now dominant in the U.S. requires that researchers and policy-makers remain open to and seek out alternative approaches to education reform that achieve similar goals and ideals as those animating this current era of reform. One alternative approach can be located in one of the most successful education systems in the world in which, interestingly, the concept of accountability is not even a topic of discussion among policy-makers: Finland.³⁰ The Finnish model of education reform is now widely regarded as having fostered a highly effective, equitable public education system that produces consistent results at a reasonable cost.³¹ It is also unique in that, contrary to current international trends toward top-down external accountability measures, the locus of accountability pressures in the Finnish model is centered around the ballot box.

In the midst of an economic crisis in the early 1990's, Finland embarked on an aggressive re-structuring of its educational system from a highly centralized top-down approach toward a system built around a commitment to mutual accountability, professional responsibility, formative assessment, and trust. Finland followed a path of de-centralization coupled with loose national standards that provided teachers and administrators with a high degree of autonomy in achieving those standards and relied on investments in the professionalization of teaching to foster school communities engaged in "research-based teaching" to raise student achievement.³² In order to balance the needs of policy-makers to provide oversight and teachers to utilize classroom assessment as an instructional tool, Finland embraced teacher-made formative assessments as the dominant method for assessing student achievement and instituted authentic assessments utilizing random-sampling techniques to provide policy-makers with rich data to gauge system and school performance without wide-scale disruption of classroom learning.

Sahlberg notes that the de-centralization of the educational sector and increased autonomy for teachers and school administrators has distributed accountability pressures throughout the system by strengthening a linkage often severed by the test-based accountability pressures now dominating education policy in the United States: professional autonomy and professional accountability.

Educational reform principles since [the] early 1990s—when much of public sector administration went through decentralization—have relied

on building professional responsibilities within schools and encouraging lateral capacity building among teachers and schools, rather than applying external accountability structures. Therefore, sample-based testing, thematic assessments, reflective self-evaluations and putting learning first have established [a] culture of mutual responsibilities and trust.³³

Despite the fact that the concept of accountability as we think about it in the U.S. is not a topic of discussion among its policy-makers, the Finnish model of education reform offers us a new perspective on how to re-think accountability. To that end, Sahlberg offers three broad recommendations for building what he terms the “intelligent accountability” exemplified by the Finnish model.

- Build trust and shared responsibility by giving school administrators and teachers greater autonomy in developing strategies for achieving academic standards and constructing internal systems of accountability.
- Employ a balanced approach to assessment that provides teachers with the freedom to construct formative assessments as an instructional tool while utilizing authentic assessments [quantitative and qualitative in nature] to provide policy-makers, communities, and parents with rich data to gauge achievement.
- Build lateral capacity by creating spaces within schools for teacher collaboration and research and networking schools to spread ideas throughout the system.

Interestingly, Sahlberg’s recommendations for building trust and shared responsibility, a balanced approach to assessment, and lateral capacity also corresponds with the themes that emerged from Darling-Hammond’s analysis of successful accountability reforms in the U.S., and it bears a marked resemblance to the conclusions that have emerged from this analysis.³⁴ While it is impossible to adopt the Finnish reform model wholesale, Finland’s educational successes can help educational actors on this side of the Atlantic ocean expand the horizon of the possible in our on-going pursuit of education reform.

In contrast to a punitive system of accountability that relegates school knowledge to the easily assessed, intelligent accountability must work to build the learning capacity of public schooling. If the goal of policy reforms built around the organizing principle of accountability is to ultimately make a positive contribution to student learning and global competitiveness then it would appear necessary to develop an intelligent system of accountability that seeks to build and nurture the institutional capacity of schools. Such a system of intelligent accountability would involve a balanced approach to constructing micro- and

macro-accountability structures that provide necessary supports to help teachers and administrators build their professional skills and create productive learning environments for their students. Constructing an intelligent system of accountability would appear to require policy-makers and communities to, at a minimum, carry out the following tasks:

- Coordinate efforts to articulate academic standards from the broad curricular goals of mathematical-scientific knowledge, literacy, and innovative-creative thinking that do justice to the intellectual complexity implicit to each;
- De-centralize the decision-making processes of education policy by providing teachers and school administrators greater professional autonomy in reaching academic standards and constructing internal accountability policies at the school level;
- Make significant investments in teacher training and building lateral capacity across schools, districts, and states;
- Coordinate efforts to construct and implement an intelligent assessment regime that employs a balanced approach of authentic, performative and traditional assessments using random-sampling techniques.

If the overarching goal of public schooling should be the empowerment of students to successfully compete in a new global reality then it is vital that policy-makers and educational leaders make the strategic investments needed to make that a reality. The Finnish model demonstrates that these kind of reforms are both practical and cost effective, and it demonstrates that a de-centralized education system relying on democratic processes to hold education actors accountable for educational outcomes can produce high levels of achievement and equity. However, carrying out such a reform agenda would require an initial expenditure of resources. This again returns us to the role communities and governments should play in an intelligent system of accountability, and it demands critical reflection.

Critical Reflection

In contemporary educational discourse, it is now quite common for policy-makers and educational experts to demand that we hold teachers and school administrators accountable for educational failure. However, what is notably absent from those debates is the voicing of a similar need to hold policy-makers and educational experts accountable for providing the necessary resources and know-how to foster academic achievement. There is a great deal of discussion in policy circles and in popular media of how to best hold students and teachers accountable for success and

failure, no doubt a tribute to American individualism, but very little is said about holding policy-makers and community leaders accountable for fulfilling their duties in the grand *contrat social* that is public education. In effect, policy-makers and educational leaders have succeeded in pushing accountability “downward” while absolving themselves of being held accountable for fulfilling their vital role in fostering academic achievement in public schools. However, it would be a mistake to view this simply as being a bug in the system. It is not. It is a feature. It is an essential element of the current system.

Education policies which absolve policy-makers and educational experts of accountability for student learning actually work to undermine their own normative justifications. This is not accidental nor is it the inevitable outcome of a structural formation, although that is certainly an element of the equation. The disjunctions and contradictions of current education policies that seek justification in the concept of accountability are a function of a particular articulation of power composed of corporate lobbies, venture philanthropists, think-tanks, conservative political organizations, and advocacy groups that wield educational concepts such as accountability as political tools with frightening efficiency.³⁵ Indeed, it is the prevalence of these private actors in the social construction of public knowledge and education policy that led me to employ popular texts written by authors connected to national think-tanks currently driving education policy in my initial treatment of the concept of accountability. This articulation of power is a manifestation of a specific form of politics that seeks to subject the public domain to private interests in the pursuit of profit, whether that be through direct per-child payments for schooling “services” or by grabbing a “4 or 5 percent slice of a couple trillion dollar public education real estate market” for speculative investment.³⁶ It is a politics that seeks to institutionalize failure in public schooling through the implementation of increasingly impossible accountability regimes [e.g., No Child Left Behind] in order to justify its colonization by private actors. It is a politics that seeks to carve out politicized markets from a “failing” public institution, markets that represent a direct conduit for private interests to tap public treasuries.³⁷

The political characteristics of this articulation of power clearly emerged in the preceding inquiry. It is a politics that above all rejects ambiguity. The constellation of social forces, institutions, cultural practices and economic formations always already present in public schooling and individual student achievement is distilled down by this political bloc into that most transcendent of Enlightenment fetishes: numbers.³⁸ Indeed, it is a politics that demonstrates an instrumentality of reason borne of business plans and MBA curricula the world over, and it renders school

learning down to the simple act of production in the form of test scores that are easily plotted on charts and presented in digital slideshows. That business leaders would formulate a model of schooling built on clearly defined goals, quantitative metrics, and punitive accountability measures should come as no surprise to anyone.³⁹ Yet, like all politics, it is a politics that is necessarily mediated through normative concepts at work in popular discourse and dominant ideological formations, concepts that are themselves battlegrounds in a ‘war of position’ among political actors vying for hegemonic dominance.⁴⁰ It is a politics that is not only contested but is a politics that creates the logical framework for its own critique and transformation. It opens itself up to its own colonization. Thus, we find in the logical framework of domination a logic of *potential* liberation.

As we have seen, the concept of accountability is being employed by this articulation of power to make specific epistemic claims that establish criteria for judging the validity of those claims, an epistemic test that current accountability policies fail. However, the critical moment consists of more than simply identifying epistemic failure or “contradiction.” Critique must also work through the logic of the concepts employed by this articulation of power in order to both bring contradictions to the surface of public discourse and to open up a political space for the transformative negation of those concepts, or to establish the foundation for a positive reconstruction of those concepts through an articulation of the necessary pre-conditions for achieving their own normative ideals.

The results of this study suggest that the necessary pre-conditions for achieving the normative ideals built into the conceptual logic of accountability would look dramatically different from the accountability policies associated with standards-based reforms. The foundation for achieving the goals of preparing students for the intellectual rigors of a 21st century economy begins with a re-commitment to public schooling as a public institution in which all share responsibility and in which accountability pressures are centered around the ballot box. Upon this foundation, the first step toward achieving the normative ideals implicit within the conceptual framework of accountability is a coordinated effort to formulate academic standards that embrace the ambiguity of the complex knowledge we seek to impart to our students. The second step requires a coordinated effort to re-orient assessment and accountability away from punitive measures that create perverse incentive structures that punish the weakest in society. Instead, what is needed is a movement toward formative assessments that both do justice to the complex knowledge we seek to impart to our students and are geared toward increasing student achievement and learning. And, finally, it

demands an institutional structure of shared accountability among all educational actors. It demands that society re-affirm its commitment to schooling as a public institution.

Notes

¹“Economic Security and a 21st Century Education: Secretary Arne Duncan’s Remarks at the U.S. Chamber of Commerce’s Education and Workforce Summit,” *Speeches and Testimony*, <http://www.ed.gov/news/speeches/2009/11/11092009.html> (Accessed: November 10, 2009).

² Of course, this economic perspective is not unique to the Obama administration. Popular debates around the relationship between global economic competition and public education can be traced back to the publication of *A Nation at Risk* during the Reagan administration.

³ Scott Ellison, “In the Shadow of Hegel: Toward a Methodology Appropriate to the Sociological Consciousness of Philosophic Inquiry,” *Education and Culture*, 26(1), (2010): 44-66.

⁴ John Dewey, *How We Think* (New York: D.C. Heath & Co., 1910), 79.

⁵ Michel Foucault, *The Politics of Truth* (2nd ed.). (Los Angeles, CA: Semiotext(e), 2007), 114

⁶ Ellison, 61.

⁷ Kevin Kosar, *Failing Grades: The Federal Politics of Education Standards* (Boulder, CO: Lynne Rienner Publishers, 2005), 47-48.

⁸ David T. Kearns & James Harvey, *A Legacy of Learning: Your Stakes in Standards and New Kinds of Public Schools* (Washington, DC: Brookings Institution Press, 2000), 139.

⁹ Kosar, 68.

¹⁰ Herbert J. Walberg, “Standards, Testing, and Accountability,” in John E. Chubb (Ed.), *Within Our Reach: How America Can Educate Every Child* (Lanham, MD: Rowman & Littlefield Publishers, 2005), 55.

¹¹ Frederick M. Hess, *Common Sense School Reform* (New York: Palgrave MacMillan, 2004), 51.

¹² Finbarr C. Sloan & Anthony E. Kelly, “Issues in High-Stakes Testing Programs,” *Theory Into Practice*, 42(1), (2003): 12-17.

¹³ Sloan & Kelly, 15.

¹⁴ Sloan & Kelly, 13.

¹⁵ Robert L. Linn, “Validity of Inferences from Test-Based Educational Accountability Systems,” *Journal of Personnel Evaluation in Education*, 19(1-2), (2006): 14.

¹⁶ “Race to the Top Education Reforms Should Be Evaluated,” National Academy of Sciences: The Office of News and Public Information of the National Academies, <http://www8.nationalacademies.org/onpinews/newsitem.aspx?RecordID=12780#> (Accessed: October 7, 2009).

¹⁷ Andrew Dean Ho, “The Problem With ‘Proficiency’: Limitations of Statistics and Policy Under No Child Left Behind,” *Educational Researcher*, 37(6). (2008): 351-360.

¹⁸ Andre A. Rupp, Tracy Ferne, & Hyeran Choi, "How Assessing Reading Comprehension with Multiple Choice Questions Shapes the Construct: A Cognitive Processing Perspective," *Language Testing*, 23(4), (2006): 470.

¹⁹ Andre A. Rupp & Nonie K. Lesaux, "Meeting Expectations? An Empirical Investigation of a Standards-Based Assessment of Reading Comprehension," *Educational Evaluation and Policy Analysis*, 28(4), (2006): 319.

²⁰ Rupp & Lesaux, 331.

²¹ Marsha Riddle Bully & Sheila W. Valencia, "Below the Bar: Profiles of Students Who Fail State Reading Assessments," *Educational Evaluation and Policy Analysis*, 24(3), (2002): 232.

²² Bully & Valencia, 233.

²³ Sharon L. Nichols, Gene V. Glass, & David C. Berliner, "High-Stakes Testing and Student Achievement: Does Accountability Pressure Increase Student Learning?" *Educational Policy Analysis Archives*, 14(1), (2006): 51.

²⁴ Gregory J. Merchant, Sharon E. Paulson, & Adam Shunk, "Relationships between High-Stakes Testing Policies and Student Achievement after Controlling for Demographic Factors in Aggregated Data," *Educational Policy Analysis Archives*, 14(30), (2006): 1-31.

²⁵ Jaekyung Lee & Kenneth K. Wong, "The Impact of Accountability on Racial and Socioeconomic Equity: Considering Both School Resources and Achievement Outcomes," *American Educational Research Journal*, 41(4), (2004): 797-832.

²⁶ Merchant, Paulson, & Shunk, 23. See also: Sharon Nichols & David Berliner, "Testing the Joy out of Learning," *Educational Leadership*, 65(6), (2008): 14-18.; Sharon Nichols, "High-Stakes Testing: Does It Increase Achievement?" *Journal of Applied School Psychology*, 23(2), (2007): 47-64.; Jaekyung Lee, "Trick or treat: new ecology of education accountability system in the USA," *Journal of Education Policy*, 25(1), (2010): 73-93.

²⁷ Sharon L. Nichols & David C. Berliner, *Collateral Damage: How High-Stakes Testing Corrupts America's Schools* (Cambridge, MA: Harvard Education Press, 2007); Wayne Au, "High-Stakes Testing and Curricular Control: A Qualitative Metasynthesis," *Educational Researcher*, 36(5), (June 1, 2007): 258-267.

²⁸ Juilian Vasquez Heilig & Linda Darling-Hammond, "Accountability Texas-Style: The Progress and Learning of Urban Minority Students in a High-Stakes Testing Context," *Educational Evaluation and Policy Analysis*, 30(2), (2008): 75-110; Linda McSpadden McNeil, Eileen Coppola, Judy Radigan, & Julian Vasquez Heilig, "Avoidable Losses: High-stakes accountability and the dropout crisis," *Education Policy Archives*, 16 (2008): <http://epaa.asu.edu/ojs/article/view/28/154>; Brain A. Jacob, "Accountability, incentives and behavior: the impact of high-stakes testing in the Chicago Public Schools," *Journal of Public Economics* 89 (2005): 763; Randall Reback, "Teaching to the rating: School accountability and the distribution of student achievement," *Journal of Public Economics* 92(5-6), (2008): 1412; Michelle Fine, Reva Jaffe-Walter, Pedro Pedraza, Valerie Futch, & Brett Stoudt, "Swimming: On Oxygen, Resistance, and Possibility for Immigrant Youth Under Siege," *Anthropology & Education Quarterly*, 38(1), (2007): 76-96; Kris Sloan, "High-Stakes Accountability, Minority Youth, and Ethnography: Assessing the Multiple Effects," *Anthropology & Education Quarterly* 38(1) (2007): 24-41; Linda Valli & Marilyn Chambliss, "Creating Classroom Cultures: One Teacher, Two Lessons, and a High-Stakes Test,"

Anthropology & Education Quarterly 38(1), (2007): 57-75; Linda Valli & Daria Buese, "The changing roles of teachers in an era of high-stakes accountability," *American Educational Research Journal*, 44(3), (September 2007): 519-558.

²⁹ Barbara J. Thayer-Bacon, *Transforming Critical Thinking: Thinking Constructively* (New York: Teachers College Press, 2000).

³⁰ P. Sahlberg, "Education Reform for Raising Economic Competitiveness," *Journal of Educational Change*, 7(4), (December 1, 2006): 259-287; Tiina Itkonen & Markku Jahnukainen, "An Analysis of Accountability Policies in Finland and the United States," *International Journal of Disability, Development and Education*, 54(1), (2007): 5; Andreas Schleicher, *The Economics of Knowledge: Why Education is Key to Europe's Success* (Brussels, Belgium: The Lisbon Council, 2006), http://www.lisboncouncil.net/index.php?option=com_publications&task=view&id=1&Itemid=35; *OECD Briefing Note For Finland* (OECD, September 9, 2008), <http://www.oecd.org/dataoecd/31/46/41277828.pdf>.

³¹ OECD, *Program for International Student Assessment* (2009) www.oecd.org/dataoecd/51/27/37474503.pdf

³² Pasi Sahlberg, "Education Policies for Raising Student Learning: The Finnish Approach," *Journal of Education Policy*, 22(22), (2007): 147-171.

³³ Pasi Sahlberg, "Rethinking accountability in a knowledge society," *Journal of Educational Change*, 11(1), (2008): 55.

³⁴ Linda Darling-Hammond, "Standards, Accountability, and School Reform," *Teachers College Record*, 106(6), (2004): 1047-1085.

³⁵ Elizabeth DeBray-Pelot & Patrick McGuinn, "The New Politics of Education: Analyzing the Federal Education Policy Landscape in the Post-NCLB Era," *Educational Policy*, 23(1), (January 1, 2009): 15-42; Janelle Scott, "The Politics of Venture Philanthropy in Charter School Policy and Advocacy," *Educational Policy*, 23(1), (January 1, 2009): 106-136; Christopher Lubienski, Peter Weitzel, & Sarah Theule Lubienski, "Is There a 'Consensus' on School Choice and Achievement?: Advocacy Research and the Emerging Political Economy of Knowledge Production," *Educational Policy*, 23(1), (January 1, 2009): 161-193.

³⁶ Rob Roberts, "Entertainment Properties Trust invests \$48M in charter schools," *Kansas City Business Journal*, January 22, 2010, <http://www.bizjournals.com/kansascity/stories/2010/01/18/daily54.html>

³⁷ John Clarke, "Dissolving the Public Realm? The Logics and Limits of Neoliberalism," *Journal of Social Policy*, 33(1), (2004): 30; Michael Apple, "Creating Difference: Neo-Liberalism, Neo-Conservatism and the Politics of Educational Reform," *Educational Policy*, 18(1), (2004): 12-44; Michael Apple, "Between Neoliberalism and Neoconservatism: Education and Conservatism in a Global Context," in Nicholas C. Burbules & Carlos Alberto Torres, (Eds.), *Globalization and Education: Critical Perspectives* (New York: Routledge, 2000).

³⁸ Max Horkheimer & Theodor W. Adorno, *Dialectic of Enlightenment* (1st ed.). (Palo Alto, CA: Stanford University Press, 2007).

³⁹ Indeed, the factory model still with us today is itself a product of the industrial era from which it emerged.

⁴⁰ Antonio Gramsci, *Prison Notebooks, Volume 2* (New York: Columbia University Press, 1996), 200-201, 156-158, 52-53.

References

- Apple, M. (2004). Creating difference: Neo-liberalism, neo-conservatism and the politics of educational reform. *Educational Policy*, 18(1), 12-44.
- Apple, M. (2000). Between neoliberalism and neoconservatism: Education and conservatism in a global context. In N. C. Burbules & C. A. Torres (Eds.), *Globalization and education: Critical perspectives*. New York: Routledge.
- Au, W. (2007). High-stakes testing and curricular control: A qualitative meta-synthesis. *Educational Researcher*, 36(5), 258-267.
- Bully, M. R., & Valencia, S. W. (2002). Below the bar: Profiles of students who fail state reading assessments. *Educational Evaluation and Policy Analysis*, 24(3), 219-239.
- Clarke, J. (2004). Dissolving the public realm? The logics and limits of neo-liberalism. *Journal of Social Policy*, 33(1), 27-48.
- Darling-Hammond, L. (2004). Standards, accountability, and school reform. *Teachers College Record*, 106(6), 1047-1085.
- DeBray-Pelot, E., & McGuinn, P. (2009). The new politics of education: Analyzing the federal education policy landscape in the post-NCLB era. *Educational Policy*, 23(1), 15-42.
- Dewey, J. (1910). *How we think*. New York: D. C. Heath & Co.
- Ellison, S. (2010). In the shadow of Hegel: Toward a methodology appropriate to the sociological consciousness of philosophic inquiry. *Education and Culture*, 26(1), 44-66.
- Fine, M., Jaffe-Walter, R., Pedraza, P., Futch, V., & Stoudt, B. (2007). Swimming: On oxygen, resistance, and possibility for immigrant youth under siege. *Anthropology & Education Quarterly*, 38(1), 76-96.
- Foucault, M. (2007). *The politics of truth* (2nd ed.). Los Angeles: Semiotext(e).
- Gramsci, A. (1996). *Prison notebooks, Volume 2*. New York: Columbia University Press.
- Heilig, J. V., & Darling-Hammond, L. (2008). Accountability Texas-style: The progress and learning of urban minority students in a high-stakes testing context. *Educational Evaluation and Policy Analysis*, 30(2), 75-110.
- Hess, F. M. (2004). *Common sense school reform*. New York: Palgrave MacMillan.
- Ho, A. D. (2008). The problem with "proficiency": Limitations of statistics and policy under No Child Left Behind. *Educational Researcher*, 37(6), 351-360.
- Horkheimer, M. & Adorno, T. W. (2007). *Dialectic of enlightenment* (1st ed.). Palo Alto, CA: Stanford University Press.
- Itkonen, T., & Jahnukainen, M. (2007). An analysis of accountability Policies in Finland and the United States. *International Journal of Disability, Development, and Education*, 54(1), 5-23.
- Jacob, B. A. (2005). Accountability, incentives, and behavior: The impact of high-stakes testing in the Chicago Public Schools. *Journal of Public Economics*, 89, 761-796.
- Kearns, D. T., & Harvey, J. (2000). *A legacy of learning: Your stakes in standards and new kinds of public schools*. Washington, DC: Brookings Institution Press.
- Kosar, K. (2005). *Failing grades: The federal politics of education standards*. Boulder, CO: Lynne Rienner Publishers.

- Lee, J. (2010). Trick or treat: new ecology of education accountability system in the USA. *Journal of Education Policy*, 25(1), 73-93.
- Lee, J., & Wong, K. W. (2004). The impact of accountability on racial and socioeconomic equity: Considering both school resources and achievement outcomes. *American Educational Research Journal*, 41(4), 797-832.
- Linn, R. L. (2006). Validity of inferences from test-based educational accountability systems. *Journal of Personnel Evaluation in Education*, 19(1-2), 5-15.
- Lubienski, C., Weitzel, P., & Lubienski, S. T. (2009). Is there a “consensus” on school choice and achievement?: Advocacy research and the emerging political economy of knowledge production. *Educational Policy*, 23(1), 161-193.
- McNeil, L. M., Coppola, E., Radigan, J., & Heilig, J. V. (2008). Avoidable losses: High-stakes accountability and the dropout crisis. *Education Policy Archives*, 16, <http://epaa.asu.edu/ojs/article/view/28/154>
- Merchant, G. J., Paulson, S. E., & Shunk, A. (2006). Relationships between high-stakes testing policies and student achievement after controlling for demographic factors in aggregated data. *Educational Policy Analysis Archives*, 14(30), 1-31.
- Nichols, S. (2007). High-stakes testing: Does it increase achievement? *Journal of Applied School Psychology*, 23(2), 47-64.
- Nichols, S. & Berliner, D. (2008). Testing the joy out of learning. *Educational Leadership*, 65(6), 14-18.
- Nichols, S. & Berliner, D. C. (2007). *Collateral damage: How high-stakes testing corrupts America's schools*. Cambridge, MA: Harvard Education Press.
- Nichols, S., Glass, G. V., & Berliner, D. C. (2006). High-stakes testing and student achievement: Does accountability pressure increase student learning? *Educational Policy Analysis Archives*, 14(1), 1-175.
- Reback, R. (2008). Teaching to the rating: School accountability and the distribution of student achievement. *Journal of Public Economics*, 96(5-6), 1394-1415.
- Rupp, A. A., Ferne, T., & Choi, H. (2006). How assessing reading comprehension with multiple choice questions shapes the construct: A cognitive processing perspective. *Language Testing*, 23(4), 441-474.
- Rupp, A. A., & Lesaux, N. K. (2006). Meeting expectations? An empirical investigation of a standards-based assessment of reading comprehension. *Educational Evaluation and Policy Analysis*, 28(4), 315-333.
- Sahlberg, P. (2008). Rethinking accountability in a knowledge society. *Journal of Educational Change*, 11(1), 45-61.
- Sahlberg, P. (2007). Education policies for raising student learning: The Finnish approach. *Journal of Education Policy*, 22(22), 147-171.
- Sahlberg, P. (2006). Education reform for raising economic competitiveness. *Journal of Educational Change*, 7(4), 259-287.
- Schleicher, A. (2006). The economics of knowledge: Why education is key to Europe's success. Brussels, Belgium: The Lisbon Council.
- Scott, J. (2009). The politics of venture philanthropy in charter school policy and advocacy. *Educational Policy*, 23(1), 106-136.
- Sloan, K. (2007). High-stakes accountability, minority youth, and ethnography:

-
- Assessing the multiple effects. *Anthropology & Education Quarterly*, 38(1), 24-41.
- Sloan, F. C., & Kelly, A. E. (2003). Issues in high-stakes testing programs. *Theory Into Practice*, 42(1), 12-17.
- Thayer-Bacon, B. J. (2000). *Transforming critical thinking: Thinking constructively*. New York: Teachers College Press.
- Valli, L., & Buese, D. (2007). The changing roles of teachers in an era of high-stakes accountability. *American Educational Research Journal*, 44(3), 519-558.
- Valli, L., & Chambliss, M. (2007). Creating classroom cultures: One teacher, two lessons, and a high-stakes test. *Anthropology & Education Quarterly*, 38(1), 57-75.
- Walberg, H. J. (2005). Standards, testing, and accountability. In J. E. Chubb, *Within our reach: How America can educate every child*. Lanham, MD: Rowman & Littlefield.